GEOMATICS AND ENVIRONMENTAL ENGINEERING • Volume 18 • Number 1 • 2024

https://doi.org/10.7494/geom.2024.18.1.29

Prima Widayani[1], Abhista Fawwaz Sahitya[2], Agatha Andriantari Saputri[3]

Support Vector Machine for Susceptibility Modeling of Dengue Fever in Kendari, Southeast Sulawesi

Abstract: Dengue fever (DF) is an infectious disease that is still a problem in Indonesia. The total death rate due to DF was 705 people in 2021; in 2022, this increased to 1183 (Indonesian Ministry of Health, 2023). Seeing this fact, prevention efforts are still needed when handling DF cases in all of the regions of Indonesia. This research was conducted in the Kendari area of Southeast Sulawesi, where there are still cases of DF. The purpose of this study was to create a spatial model of dengue susceptibility using a support vector machine. Landsat 8 imagery was used to intercept data on building density, vegetation density, land use, and land surface temperatures. Rainfall and humidity variables were obtained from the Meteorological, Climatological, and Geophysical Agency (BMKG). Based on the modeling results, the districts of Wua-wua, Kadia, Barunga, Poasi, and Puuwatu are areas with high susceptibility. The results of testing the susceptibility model to dengue hemorrhagic fever (DHF) in Kendari obtained an area under the curve (AUC) of 0.75, meaning that this model was well-accepted.

Keywords: dengue, infectious disease, spatial model, support vector machine, susceptibility

Received: July 3, 2023; accepted: December 10, 2023

[1] Universitas Gadjah Mada, Faculty of Geography, Yogyakarta, Indonesia,
email: primawidayani@ugm.ac.id (corresponding author), https://orcid.org/0000-0003-0891-2074
[2] Universitas Gadjah Mada, Faculty of Geography, Yogyakarta, Indonesia.
[3] Universitas Gadjah Mada, Faculty of Geography, Yogyakarta, Indonesia

1. Introduction

The epidemiological theory states that the occurrence of a disease is influenced by three factors; namely, the agent, the host(s), and the environment [1]. The agent of dengue fever (DF) is the dengue virus, while the hosts include *Aedes aegypti* and *Aedes albopictus* mosquitoes as well as humans. The environment that influences the emergence of DF is an environment that is suitable as a mosquito habitat. This environment serves both as a breeding place and a resting place [2]. As a country with a tropical climate, Indonesia is a suitable habitat for various mosquitoes, including *Aedes* sp. mosquitoes; as such, DF is still an active threat in several regions in Indonesia. The Indonesian Ministry of Health recorded the number of DF cases in Indonesia as 138,127 back in 2019, which then dropped to 108,303 in 2020 and decreased further to 73,518 in 2021. However, this number rose significantly in 2022 – to the point of 131,265 cases.

Dengue fever cases in Indonesia have been reported to have occurred in 451 regencies/cities spread across 34 provinces [3]. The cases that were recorded in the data from the Ministry of Health included DF cases in Kendari. According to data from the Kendari City Health Office, Kendari is the region with the highest number of DF cases as compared to the other districts/cities throughout Southeast Sulawesi. A graph of the DF cases in Kendari (Fig. 1) shows an increase from 93 cases in 2017 to 111 cases in 2018, with a significant rise in 2019 (450 cases). By 2020, the number of cases gradually decreased to 307, then to 211 cases in 2021; these rose again slightly in 2022 (230 cases).

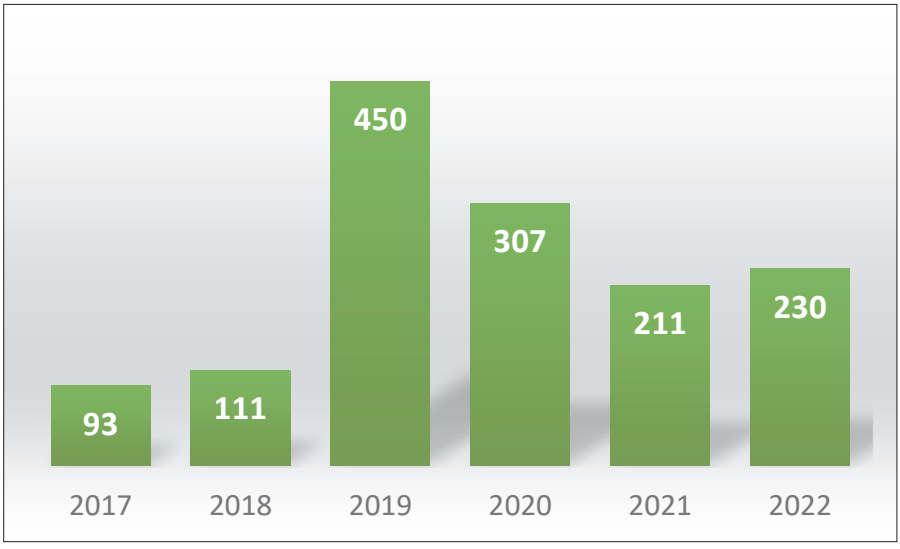

Fig. 1. Dengue fever cases in Kendari, Southeast Sulawesi (2017–2022)
Source: Kendari Municipality in Figures, https://kendarikota.bps.go.id

Various efforts had been made by the government in collaboration with the community through various activities, such as (1) mosquito net installations on beds and windows, (2) 3M (drain, close, and recycle) program implementation, (3) increased consumption of vitamin C; and (4) doses of DF vaccine with doctors' consultations [4]. The World Mosquito Program (WMP) gave an additional effort by spreading out *Wolbachia* mosquitoes, which were tested in some parts of Yogyakarta. Based on the results of the study, it was found that, in areas where the *Wolbachia* mosquitoes were spread, the DF rate decreased by 77.1%, and the hospitalization rate due to DF decreased by 86.1%. It turns out that this intervention possesses a better effectivity than the dengue vaccine [3]. The results of this research were quite good in eliminating DF cases; however, not all of the regions in Indonesia have the opportunity to have *Wolbachia* mosquitoes spread. Therefore, prevention efforts must continue to be pursued in any form, including the mapping of those areas that are susceptible to DF.

Law No. 24 of 2007 defined susceptibility as geographical, geological, biological, hydrological, socio-cultural, political, economic, and technological conditions in an area for a certain period of time that decreases the ability to prevent, reduce, and achieve readiness and respond to the adverse effects of certain hazards [5]. A comprehension of the vulnerable areas is the first step in reducing DF cases as well as being an input in reducing the risks that arise when DF cases spread. A model is a simplification of the real world, while one of the models that is often used by geographers is a spatial model. The spatial model is manifested as a map (among other things). Through maps, geographers attempt to understand the features that are on the surface of the Earth as spatial phenomena. In this case, spatial phenomena can be seen such as distribution (what, where, and how wide) as well as the relationships among objects in space (influence, adjacency, and accessibility) [6]. The use of spatial models in an effort to understand the influence of the environment on a disease is necessary in the effort to control DF cases. Producing a susceptibility model requires data input in the form of variables that influence the incidence of DF; namely, environmental and climatic conditions. A large area and difficult terrain will render this variable difficult to attain. Therefore, another approach is needed besides field surveys, and this approach is to use remote-sensing data. Various types of remote-sensing data are available, both as free and paid access; this is a perfect opportunity to utilize remote-sensing data in various fields, including the field of health studies.

The development of remote-sensing-data processing has also experienced rapid progress – ranging from the simplest to the most automated methods. Visual, digital, and machine-learning interpretation methods are alternatives that can be used and adapted to the necessary means. The spatial modeling of susceptibility, which is carried out using a weighted tiered method, has become the most widely used method; however, this does not rule out the possibility of trying other approaches such as a machine-learning approach to model susceptibility. In this research, the spatial modeling of DF susceptibility will be carried out using one of the machine-learning algorithms – the support vector machine (SVM). The classification performance

using SVM for dengue cases in Malaysia was superior to the other techniques, as it obtained the highest prediction accuracy (85%) [7]. Hamdani et al. [8] classified dengue using five classification methods, including C.45, decision tree (DT), *k*-nearest neighbor (KNN), random forest (RF), and support vector machine (SVM). The most optimal evaluation results were obtained when using SVM (with an accuracy value of 99.1%). Mountrakis et al. [9] stated that SVM was suitable for object classification using remote-sensing data; this was due to SVM's ability to generalize well – even with limited training data. This research was not the first to model DF susceptibility using SVM, but it was hoped that it could further strengthen the results of previous research by attempting it in other areas under different environmental conditions.

2. Materials and Methods

2.1. Study Area

This research covers the area of Kendari, Southeast Sulawesi (Fig. 2). The reason to go for this location is that, among other cities/districts in Southeast Sulawesi, Kendari experiences the highest number of DF cases. In addition, research on the spatial model of DF susceptibility is mostly carried out in the western part of Indonesia, so it was deemed necessary to conduct research in the eastern part as well.

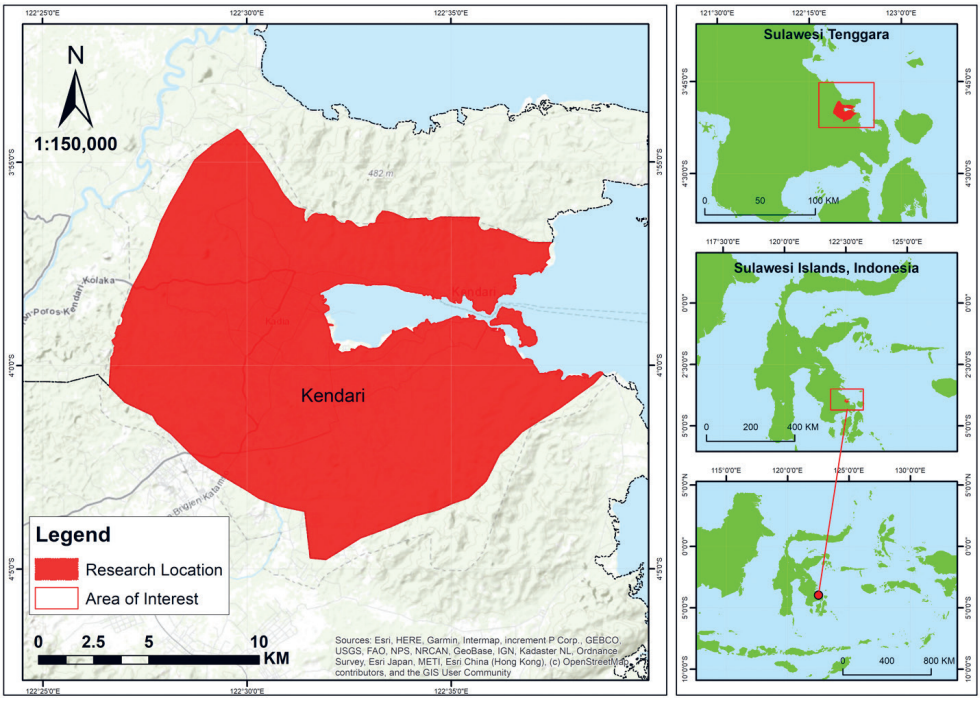

Fig. 2. Map of Kendari based on base map of Rupa Bumi Indonesia

2.2. Data

The following research materials and tools were used:

- Landsat 8 imagery that was taken on January 24, 2022 with 25.88% cloud cover in the TIF format, Bands 2 through 7, obtained from the United States Geological Survey's (USGS) official website https://earthexplorer.usgs.gov/. Geometric and radiometric corrections were done using the semi-automatic classification plugin (SCP) in QGIS for further processing.
- High-resolution imagery (Google Satellite is applicable) from 2023.
- Base map of Rupa Bumi Indonesia, with a scale of 1:25,000 that was obtained from the Geospatial Information Agency (BIG) as a source for extracting administrative boundaries in the form of a shapefile.
- Annual rainfall and humidity data obtained from the Meteorology, Climatology, and Geophysical Agency (BMKG) for 2020 from four stations (located in Kendari, Betoambari Bau-Bau, South Konawe, and Sangia Ni Bandera Kolaka).
- A computer with RStudio, ENVI, and QGIS software.

2.3. Method

Multispectral Classification

Land-cover and land-use data was obtained from a supervised multispectral classification using the maximum likelihood algorithm [10, 11]. This algorithm was chosen due to it being considered statistically the most stable as compared to some others (such as parallelepiped or minimum distance to mean) [12]. Another advantage of this method is its simplicity as a parametric algorithm; in such a case where the probability distributions are valid, this algorithm would yield a better result than newer algorithms [13]. The classification scheme used SNI 7645-1:2014 [14], which was modified according to the needs of the study theme and can be seen in Table 1.

Table 1. Land-cover and land-use classification scheme

No.	Classification
1	Upland forest
2	Lowland forest
3	Mangrove forest
4	Garden and mix crop
5	Dryland farming
6	Paddy field
7	Grassland
8	Woodland
9	Shrub
10	Cultivated open land
11	Settlement
12	Water bodies
13	Natural open land

Source: SNI 7645-1:2014 [14] (with modifications)

This classification scheme requires several composites to work with in order to gain the best result. The Natural Color composite (Band 4-3-2) was utilized to determine range-of-interest (ROI) samples for classes that are easier to recognize such as settlements, water bodies, and natural open lands. On the other hand, those classes that mostly consist of vegetation need a more specific composite to assist in determining each class's ROI samples. Settlements and natural open lands as well as some other classes that require vegetation-density identification sample determinations can also be assisted by the Short-Wave Infrared composite (Band 7-6-4) to be more accurate. For cropland classes such as paddy fields, cultivated open lands, dryland farms, and gardens and mix crops, the Agriculture composite (Band 6-5-2) can help a lot. Based on the confusion matrix, the resulting overall accuracy was 77.3358% (with a Kappa coefficient of 0.7552). The mangrove forest class had the best accuracy, while gardens and mixed crops had the worst. The results of the multispectral classification were then further tested for accuracy using high-resolution images. Based on this method, an accuracy value of 90% was obtained.

Index Transformation

Index transformation was performed to obtain information on vegetation density and built-up land density, which were represented in the normalized difference vegetation index (NDVI) and the built-up land feature-extraction index (BLFEI). The calculation of the vegetation-density index referred to the equation that was used by Ganie and Nusrath [15] (as follows):

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{RED})}{(\mathrm{NIR} + \mathrm{RED})} \tag{1}$$

where:

NDVI – vegetation density index,
NIR – near-infrared band (Band 5 of Landsat 8, with 0.851–0.879 μm wavelength),
RED – red band (Band 4 of Landsat 8, with 0.636–0.673 μm wavelength).

To obtain the vegetation-density data, the NDVI values needed to be correlated with the density values from the field surveys. The regression equation that was obtained regarding the NDVI value and the density in the field will be used to convert the NDVI value into a vegetation-density value throughout the image coverage.

The index that was used to show the appearance of the built-up land and open land and water levels in this study referred to the formula that was proposed by Bouhennache et al. [16], which was as follows:

$$\mathrm{BLFEI} = \frac{\frac{\mathrm{OLI3} + \mathrm{OLI4} + \mathrm{OLI7}}{3} - \mathrm{OLI6}}{\frac{\mathrm{OLI3} + \mathrm{OLI4} + \mathrm{OLI7}}{3} + \mathrm{OLI7}} \tag{2}$$

where:

BLFEI – built-up land object-extraction index,
OLI3 – Band 3 of Landsat 8 OLI (0.533–0.590 μm green band),
OLI4 – Band 4 of Landsat 8 OLI (0.636–0.673 μm red band),
OLI6 – Band 6 of Landsat 8 OLI (1.566–1.651 μm SWIR 1 band),
OLI7 – Band 7 of Landsat 8 OLI (2.107–2.294 μm SWIR 2 band).

A threshold value was needed to separate the pixels that were considered to be built-up areas from those areas that were considered to be non-built-up areas (vegetation, water, and open land). The calculation of this threshold value referred to Equation (3):

$$T_h = \text{ArgMax}_{\alpha \le t \ge \beta}(\sigma^2) \tag{3}$$

Meanwhile, the variance (σ) was calculated using the following general variance calculation equation:

$$\sigma = \sqrt{P_b \cdot (M - M_b)^2 + P_n \cdot (M - M_n)^2} \tag{4}$$

with M as the average value of all of the images calculated through Equation (5):

$$M = P_b \cdot M_b + P_n \cdot M_n \text{ and } P_b + P_n = 1 \tag{5}$$

where:

T_h – Otsu's optimal threshold value,
σ – variance,
t – threshold value,
M – average value of all images,
P_b and P_n – pixel probability of being classified as built-up or non-built-up,
M_b and M_n – average value of built-up and non-built-up classes.

The calculation of land surface temperature (LST) referred to the equation that was issued in the Landsat 8 handbook by USGS [17]:

$$T_{\text{rad}} = \frac{K_2}{\ln\left(\frac{K_1}{L_\lambda} + 1\right)} - 273.15 \tag{6}$$

where:

T_{rad} – top of atmosphere brightness temperature or radiance temperature [K],
L_λ – TOA spectral radiance or corrected spectral radiance value [W/(m²·sr·μm)],
K_1 – spectral radiance calibration constant of Band 10 obtained from meta data (K1_CONSTANT_BAND_x, where x is number of band);
K_2 – absolute temperature calibration constant of Band 10 [K] obtained from meta data (K2_CONSTANT_BAND_x, where x is number of band).

Based on the description of Equation (6), it was known that calculating the T_{rad} information on the corrected radiance value was needed; this was then calculated based on the equation that was proposed by Coll et al. [18]:

$$L_{\text{sensor},\,\lambda} = \frac{L_\lambda - L^{\uparrow}_{\text{atm}}}{\tau_\lambda} - \frac{1-\varepsilon_\lambda}{\varepsilon_\lambda} L^{\downarrow}_{\text{atm}} \tag{7}$$

where:

$L_{\text{sensor},\lambda}$ – at surface radiance value or corrected spectral radiance [W/(m²·sr·μm)],
L_λ – spectral radiance of Band 10,
ε_λ – object emissivity value,
$L^{\downarrow}_{\text{atm}}$ – atmospheric downwelling radiance value [W/(m²·sr·μm)],
$L^{\uparrow}_{\text{atm}}$ – atmospheric upwelling radiance value [W/(m²·sr·μm)],
τ_λ – atmospheric transmittance value.

The downwelling, upwelling, atmospheric-radiance, and atmospheric-transmittance values can be obtained through http://atmcorr.gsfc.nasa.gov/, which is a website that was developed by NASA that is specifically intended for calculating atmospheric correction parameters. Meanwhile, spectral radian value was calculated using the following equation in the Landsat 8 handbook:

$$L_\lambda = M_L Q_{\text{cal}} + A_L \tag{8}$$

where:

L_λ – spectral radiance at sensor [W/(m²·sr·μm)],
M_L – rescaling constant (RADIANCE_MULT_BAND_X, where x is Band 10),
Q_{cal} – pixel value (DN),
A_L – additional constant (RADIANCE_ADD_BAND_X, where x is Band 10).

Based on Equation (8), it is then known that, in order to obtain the corrected spectral radiance value $L_{\text{sensor},\lambda}$, the emissivity value of object ε_λ is required; this was then calculated using the equation that was proposed by Valor and Caselles [19]:

$$\varepsilon = \varepsilon_v P_v + \varepsilon_s \left(1 - P_v\right) + 0.06 P_v \left(1 - P_v\right) \tag{9}$$

where:

ε – object emissivity,
ε_v – vegetation emissivity,
P_v – vegetation fraction,
ε_s – soil emissivity.

Calculating the emissivity value of the object in Equation (9) requires the value of the vegetation fraction where vegetation fraction P_v can be derived through NDVI:

$$P_v = \left[\frac{\text{NDVI} - \text{NDVI}_s}{\text{NDVI}_v - \text{NDVI}_s} \right]^2 \tag{10}$$

where:

P_v – vegetation fraction,
NDVI – normalized difference vegetation index,
NDVI_s – normalized difference vegetation index value for open land,
NDVI_v – normalized difference vegetation index value for 100% vegetation fraction.

Climate Data Processing

The mapping of the rainfall and humidity was done using the inverse distance weighting (IDW) interpolation method. The IDW method was chosen based on two main reasons: the topographical conditions of the study area, and the number of rainfall and humidity observation stations. The IDW formula that was used referred to Kurniadi et al. [20] as follows:

$$Z_0 = \frac{\sum_{i=1}^{s} Z_i \frac{1}{d_i^k}}{\sum_{i=1}^{s} \frac{1}{d_i^k}} \tag{11}$$

where:

Z_0 – approximate value at point 0,
Z_i – z value at control point i,
d_i – distance between point i and point 0,
k – influence of the neighboring points,
s – number of S points used.

Susceptibility Spatial Model with SVM

The classification of the area susceptibility against DF was carried out using the SVM algorithm. In order to classify the high-complexity data, a kernel function was used to move the data into a higher dimensional space. The training data was denoted as $TD(x_i y_i)_{i=1^N}$, where $x_i \in R^n$ is the input vector, N is the number of samples of the training data, n is the dimension of the training data, and $y_i \in (-1,+1)$ denotes the class-label classification. The vector input consisted of six variables, including the land-use land cover, vegetation density index (NDVI), built-up land index (BLFEI), surface temperature (LST), rainfall, and air humidity. The susceptibility classification consisted of susceptible and non-susceptible area classes. The goal of the SVM model was to find the best decision function (which we call a hyperplane – Equation (12)) that can separate the sample data into two classes. If the input

vector ($x_i \in R^n$) was lower or higher than the hyperplane, then it would be classified as −1 (non-susceptible class) and +1 (DF-susceptible class), respectively.

$$f(x) = \text{sign } (\sum y_i \alpha_i K(x_i x_j) + b) \tag{12}$$

where:

b – the offset value,
α_i – Lagrange multiplier,
K – kernel function.

It was known that the capability of the SVM model depended on the kernel functions that were used; therefore, the radial basis function (RBF) was chosen as the kernel function in this study, as it has advantages over the other functions in terms of its predictive ability in various fields of work [21]. Due to the RBF-SVM kernel function that was used, the ability of this algorithm was influenced by two main parameters: regularization (C), and kernel width (γ). These needed to be chosen appropriately [22]. The selection was done through the commonly used trial-and-error method.

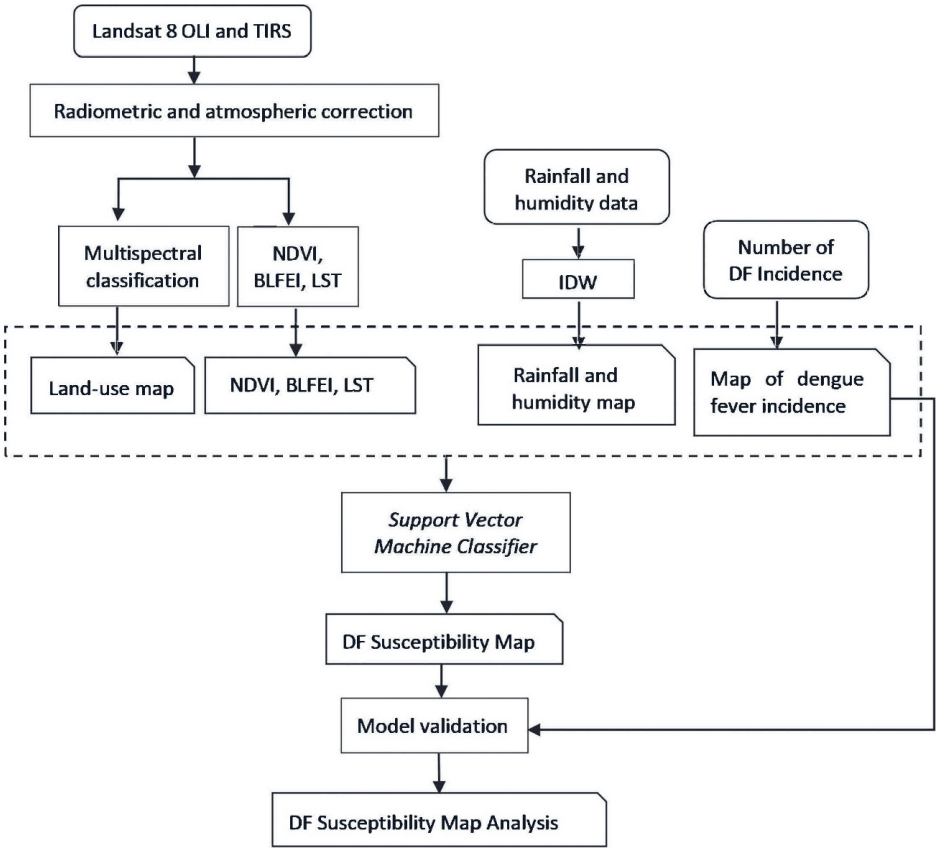

Fig. 3. Research flowchart

Dengue Fever Susceptibility Spatial Model Testing

An accuracy test of the regional susceptibility model against DF was carried out using receiver operating characteristic (ROC) curve analysis. The calculation of the ROC curve was based on a graph of sensitivity and specificity with a certain threshold, which could be done by involving a confusion matrix table that showed the number of pixels that were mistakenly classified as areas with dengue fever or areas without dengue fever. The area under the ROC curve (AUC) was used to compare the overall capabilities based on the statistics from the resulting model. Consequently, the AUC-ROC curve described the ability of the model algorithm in predicting areas that were susceptible or non-susceptible to dengue fever based on some summary of the statistical values. The qualitative relationship between AUC and accuracy was described by Motevalli et al. [21] in the following value ranges: 0.9–1.0 (perfect), 0.8–0.9 (very good), 0.7–0.8 (good), 0.6–0.7 (moderate), and 0.5–0.6 (poor). The full course of the research is presented in the following research flowchart (Fig. 3).

3. Results and Discussion

3.1. Dengue Susceptibility Variable Mapping

The policies for preventing dengue fever require accurate and real-time information – remote-sensing data is one of the alternative solutions for providing environmental data as a habitat for dengue fever-carrying vectors. The utilization of remote-sensing data for monitoring the presence of *Aedes aegypti* mosquitoes in Northern Argentina was conducted by Scavuzzo et al. [23]. Remote-sensing data was used to obtain several environmental variables through the transformations of NDVI, NDWI, LST night-time, LST daytime, and TRMM-GPM rain from 2012 through 2016 as predictive variables. In their systematic review, Louis et al. [24] selected 26 journals that mostly used remote sensing to provide land-use, vegetation density, building density, and wetlands data.

Figure 4 presents the variables that were used to model dengue fever susceptibility. Figure 4a is a land-use map in Kendari; there are ten land-use categories, where residential land dominates the land-use in Kendari. Settlements are the preferred habitats for *Aedes aegypti* mosquitoes, with dense settlements in the Wua-wua, Kadia, Barunga, Poasia, Puuwatu, and Kendari Barat sub-districts.

Figure 4b is the result of NDVI transformation; from the figure in the areas with very low vegetation densities (with a minimum index value of −0.24), there were built-up land areas, while the highest NDVI value (0.85) was in the highland forest and lowland forest areas. NDVI represents vegetation density; the relationship between vegetation density and dengue habitats is that areas with dense vegetation are not habitats for *Aedes aegypti* mosquitoes. As previously mentioned, a form of *Aedes* mosquito container is a natural container that is closely related to the presence of vegetation. Several researchers (such as Widayani et al. [25]) tried to conduct spatial

modeling that involved vegetation density (NDVI) and building density (VGNIR-BI) variables in order to assess the effect of changes in these two variables on dengue fever incidences. The research that was conducted by Widayani et al. [25] showed that there were consecutive correlations of 68–74% to the incidences of dengue fever. In parallel, Palaniyandi [26] showed that vegetation density/NDVI values could represent clues for delineating areas that were susceptible to the spread of dengue fever.

The NDVI values of <0.4 that were obtained in the study of Palaniyandi [26] were influenced by the presence of actively photosynthesizing vegetation such as rubber plantations and pineapple plants (which influenced the increasing number of *Aedes* mosquitoes), while the coconut shells that were used to collect sap in rubber plantations were good artificial containers for the growth of *Aedes* mosquitoes. In addition, the water that collected on pineapple leaf plates was also a good natural container for *Aedes aegypti* and *Aedes albopictus* mosquitoes for laying their eggs in the breeding process.

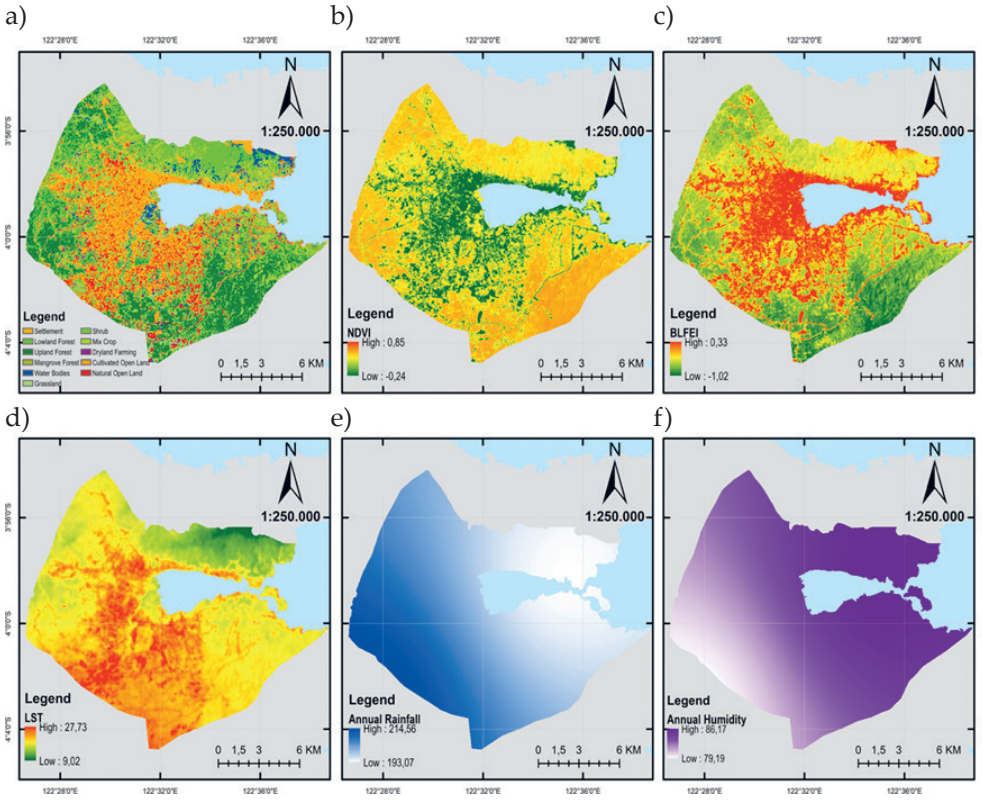

Fig. 4. Distribution of each variables used to model dengue fever susceptibility:
a) land use; b) normalized difference vegetation index;
c) built-up land feature-extraction index; d) land surface temperature;
e) annual rainfall; f) annual humidity

The results of the Landsat image processing for building density are shown in Figure 4c using BLFEI transformation. The image shows that the area around Teluk Kendari is dominated by dense buildings to the west and southwest. The built-up land in Kendari is dominated by settlements, commerce, and public facility buildings. The BLFEI index value ranges from −1.33 to 0.33. Areas with high density have an index value of 0.00–0.33.

Climate variables in the form of land surface temperature were obtained from the processing of Channel 10 or the TIR channel of Landsat 8. Land surface temperatures influence air temperatures. Figure 4d presents the results of the land surface-temperature processing. In those areas with land use in the form of built-up land and little vegetation, the land surface temperature was higher than the areas with forest land-use or agricultural areas. The use of climate variables to look at dengue susceptibility was also conducted by Tjaden et al. [27], Yin et al. [28], Davis et al. [29], and Dickin and Wallace [30]. Climate variables that have relationships with the breeding of *Aedes aegypti* mosquitoes are temperature, humidity, and rainfall. The *Aedes aegypti* larval and pupal stages take an average of 31–32 days at a temperature of 15°C, 13.9–16.6 days at 20°C, 10.2–11.8 days at 25°C, and 7.4–8.2 days at 30°C; *Aedes aegypti* mosquitoes do not develop at temperatures of 8.2–10.6°C. *Aedes albopictus* takes 12.6–15.1 days at 20°C, 10.0–11.7 days at 25°C, and 7.6–8.4 days at 30°C; the larvae and pupae of *Aedes albopictus* do not develop at all at temperatures of 4.2–8.2°C.

At 25°C and 80% humidity, female mosquitoes live twice as long and produce 40% more eggs as compared to the conditions of 35°C and 80% relative humidity. However, 45% of female mosquitoes experience oviposition inhibition at 35°C and 60% relative humidity, and only 15% of female mosquitoes are able to lay more than a hundred eggs. It was further observed by Costa et al. [31] that 60% humidity accompanied by increased temperatures would lead to a decrease in egg fertility (although this effect was not observed under conditions where the relative humidity was 80% with surface temperatures between 25–30°C.

Despite the temperature and relative humidity, another climatic factor that may contribute to the rapid growth of *Aedes* mosquitoes is rainfall. Figure 4f shows the results of the data processing of the average annual rainfall from the BMKG Kendari data. The IDW interpolation results showed that the average rainfall was high in the western area of Kendari. In the central area of Kendari, the average annual rainfall is moderate at around 200 mm^3/year. Areas with moderate rainfall are more suitable for mosquito development; this is because there is a lag time for eggs to hatch. However, these climatic factors do not necessarily correlate with the larvae-free number (LFN). Some research that was performed on Java Island (such as Sleman and Surabaya) showed that LFNs have no effect on the incidences of dengue fever [32, 33]. Valdez et al. [34] predicted that higher variability in the timing distribution of rainfall would lead to the decreased development of *Aedes aegypti* during the dry season. Furthermore, Iriani [35] looked at the relationship between rainfall

patterns and the increase in the incidences of dengue fever. The relationship began one month before the peak of the rainfall ($r = 0.332$; $p = 0.001$), then the relationship increased during the peak of the rainfall ($r = 0.353$; $p = 0.000$) and then decreased one month after the peak of the rainfall ($r = 0.262$; $p = 0.008$); in other words, the strongest correlation between dengue fever and the rainfall factors occur during the peak of the rainfall.

In addition, humidity also affects mosquito development. According to Mukhopadhyay et al. [36], *Aedes aegypti* mosquitoes can survive at 60–89% humidity. Based on the data from BMKG Kendari, the average annual air humidity is as low as 79% and as high as 89%. The distribution of the humidity in Kendari can be seen in Figure 4f.

3.2. Modeling Dengue Fever Susceptibility using Support Vector Machine

The variables that were previously mentioned were then used to create a spatial modeling of the dengue fever susceptibility. Spatial modeling using a support vector machine has been proven to be an effective method for describing areas that are susceptible or not susceptible to dengue fever. The reliability of the SVM algorithm in the spatial modeling of dengue was proven in a study by Scavuzzo et al. [23], who used machine-learning techniques with fully accessible open-source tools. The produced model had the advantage of being non-parametric and able to describe the nonlinear among between variables. According to Mountrakis et al. [9], SVM can be applied very well in the field of remote sensing because of its ability to generalize well – even with limited training samples.

The previously mentioned parameters were input to RStudio and combined using a raster stack. The data of the dengue fever presence was also included, while its absence was generated using random points. The values from each of these points was then extracted as a basis for considering dengue fever's presence. Since an SVM sees data as binary, the value of dengue fever's absence would be considered to be −1, while its presence would be 1. This information was added to the object's table as a new column (both dengue fever's presence as well as its absence). The binary value of each parameter for each pixel would then determine its own susceptibility class; the more parameters that are read as 1, the higher their probabilities of be regarded as susceptible pixels. This algorithm was trained using 70% training data and 30% test data in order to gain more-accurate results. Based on trial and error, a kernel width of 0.01 and a kernel regularization of 2 were considered to be adequate.

The pixel values of each training and test data point for each variable became the basis on which each pixel could be determined as either presence or absence. The algorithm would recognize which values for each parameter were most likely to raise the probability of dengue fever's presence and which were not.

The calculation for this classification determination would be either:

$$w \cdot x_i + b < 0$$

for an absence classification (which would yield a hypothesis of −1) or:

$$w \cdot x_i + b \geq 0$$

for a presence classification (which would yield a hypothesis of 1).

A hyperplane was created through the result of each point based on the presence and absence of dengue fever. The variables of each training and the test data values determined their locations in each plane. The boundary of this hyperplane could be obtained via Equation (12); the calculation was done by inserting a sigma value (σ) of 0.01 and a *C* of 2 in the script, and RStudio would then create an optimal hyperplane boundary based on existing data that has been included. The hyperplane would make an adjustment when there was new data inserted.

Based on the results of the modeling of dengue fever susceptibility using SVM in Kendari, there are three classes of susceptibility; namely, high, medium, and low. High vulnerability is shown in orange in Figure 5, while medium vulnerability is depicted in yellow color and low vulnerability in gray.

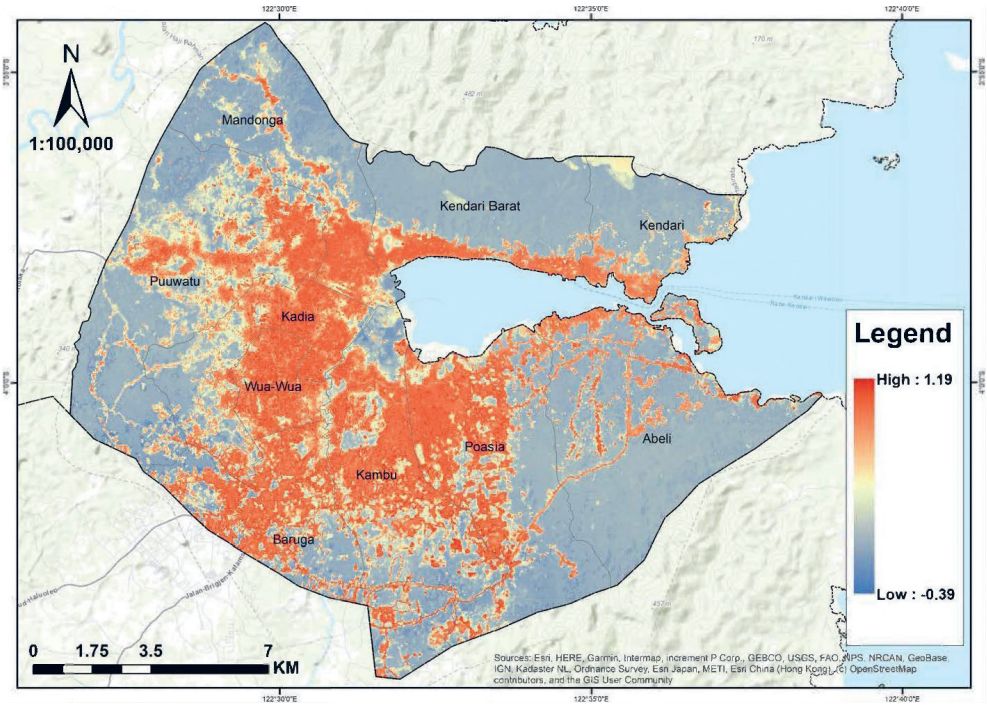

Fig. 5. Susceptibility map of dengue fever in Kendari based on the analysis using SVM

The Wua-wua, Kadia, Barunga, Poasi, and Puuwatu sub-districts are high-susceptibility areas, the Kendari Barat and Ambu sub-districts are in the medium-susceptibility class, and the Abelia, Mandonga, Kendari, and Nambo sub-districts are low-susceptibility areas. These results were in line with the average data of the DHF cases during the period of 2015–2021 from the Kendari City Health Office (as can be seen in Figure 6).

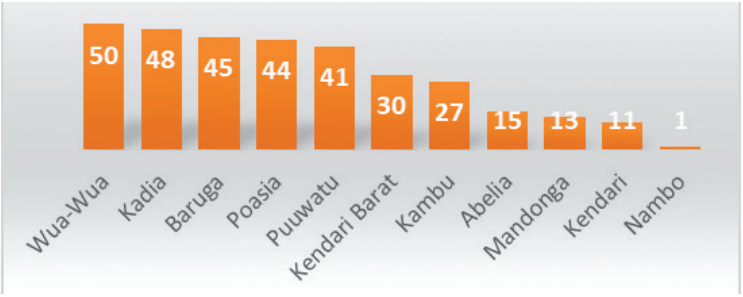

Fig. 6. Average DF cases in Kendari in 2015–2021

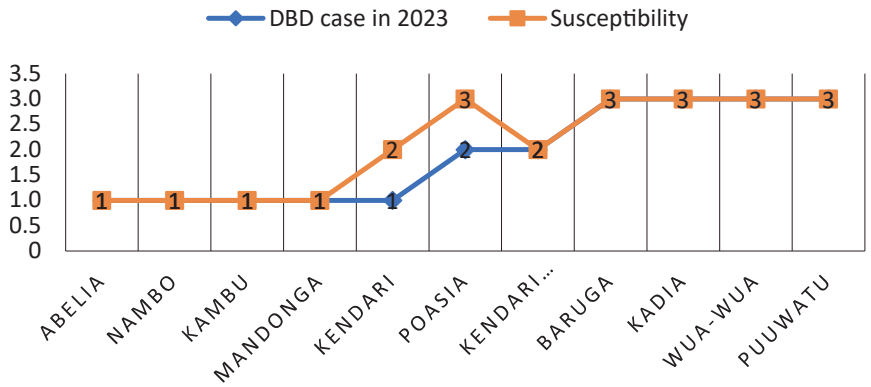

Fig. 7. Susceptibility classification vs. DHF case classification in 2023

Figure 7 is a graph that shows the results of the susceptibility modeling with the classification of DHF cases in Kendari in 2023. In the graph, there are two areas that were not suitable; namely, the Kendari and Poasia sub-districts. Based on this graph, the accuracy of the susceptibility modeling vs. the dengue cases in 2023 was 81.81%. This was a good result, as it proved that SVM can be used to produce some sort of forewarnings for local people to be more vigilant – especially for those who live in highly susceptible areas. The government (be it in Kendari or other cities) can also use this method to analyze its own region. The SVM algorithm's hyperplane adaptability is quite commendable as well, opening up the possibilities to add more variables in order to obtain the best results.

3.3. Model Test

The area under the curve indicates the accuracy of the prediction model and was calculated using a method called area under curve (AUC); an AUC is a square-shaped area whose value is always between 0 and 1. The result test of the dengue fever susceptibility model in Kendari obtained an AUC of 0.75 (see Fig. 8). An AUC has the following diagnostic value levels: accuracy levels of 0.90–1.00 = excellent classification; 0.80–0.90 = good classification, 0.70–0.80 = fair classification, 0.60–0.70 = poor classification, and 0.50–0.60 = failure classification. Based on the accuracy classifications, the model is in the fair classification class.

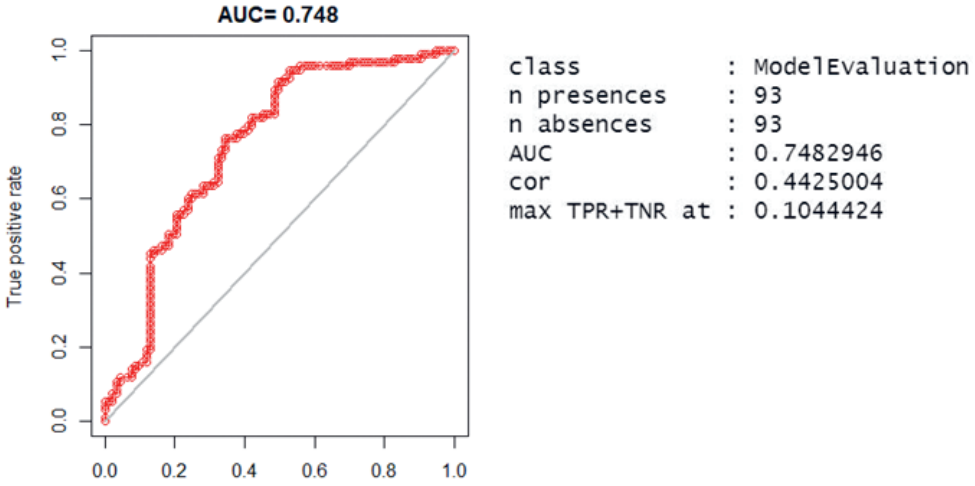

Fig. 8. Validation model using area under curve

Siddiq [37] concluded that, when modeling the predictions of dengue fever cases from the tested models, the support vector machine (SVM) model had the best performance (achieving a 76% prediction accuracy), while linear regression, random forest regression, and decision tree regression achieved prediction accuracies of 52, 55, and 57%, respectively. Mizan and Widayani [38] compared the uses of SVM and RF algorithms for DHF susceptibility; the results indicated that the SVM algorithm had a better ability than the random forest algorithm (with an AUC value of 1 and a correlation of 0.95). The AUC test results in the research of Mizan and Widayani [38] were better when compared to this study, as Mizan's research used data on DHF cases before the COVID-19 pandemic; the data on DHF cases was more accurate before the COVID-19 pandemic. This study used data on DHF cases from 2017 to 2023. In 2020 and 2021, DHF symptoms may have been mixed with COVID-19 symptoms; this is what caused the AUC value of this study to be lower when compared to the research by Mizan and Widayani [38].

3.4. Future Work

The next research plan is to create a spatial modeling of the risk of dengue fever; risk modeling considers vulnerability, susceptibility and capacity. The data that is used can be developed by adding information from the social environment. The use of image data can also be developed by using images with high spatial resolutions. The number of cases of dengue fever are still high in Indonesia, so opportunities for research into the spatial modeling of dengue fever using various methods are still wide open. The most important thing after conducting the research will be implementing the results into real activities. The results of this research can be used to support local governmental efforts in controlling dengue fever.

4. Conclusions

Dengue fever susceptibility modeling was carried out using physical land variables and climate. Determining these variables was based on epidemiological theories about the occurrence of a disease. The Landsat 8 OLI TIRS satellite images were able to provide data on land physical variables in the forms of building density, vegetation density, and land use through NDVI, BLFEI transformation processing, and supervised maximum-likelihood classification. The Landsat imagery was also able to provide land surface-temperature variables through the calculation of the LST algorithm. Average annual rainfall and humidity were obtained from the Meteorological, Climatological, and Geophysical Agency station measurement data. The use of machine learning with the SVM algorithm was able to properly separate the classes of regional susceptibility to dengue fever in the city of Kendari. The model was tested using AUC, which was a square-shaped area whose value was always between 0 and 1; the test results showed an AUC value of 0.75. The results of the dengue fever susceptibility spatial modeling showed that the sub-districts of Wua-wua, Kadia, Barunga, Poasi, and Puuwatu were areas with high susceptibility. The West Kendari and Ambu sub-districts were placed in the moderate susceptibility class, while the Abelia, Mandonga, Kendari, and Nambo sub-districts were areas with low susceptibility. These results were in line with the average data on dengue cases for the period of 2015–2021 from the Kendari Health Office. The spatial modeling of DF susceptibility is expected to assist in future efforts to prevent, handle, and eradicate DF cases in Kendari.

Author Contributions

Prima Widayani: conceptualization, methodology, formal analysis, supervision, writing – original draft.

Abhista Fawwaz Sahitya: formal analysis, visualization, writing – editing.

Agatha Andriantari Saputri: data curation, software, validation, writing – review & editing.

References

[1] Hewa S.: *Theories of disease causation: Social epidemiology and epidemiological transition*. Galle Medical Journal, vol. 20(2), 2015, pp. 26–32. https://doi.org/10.4038/gmj.v20i2.7936.

[2] World Health Organization: *Dengue and severe dengue*. March 17, 2023. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue [access: 11.02.2023].

[3] Kementerian Kesehatan Republik Indonesia: *Wolbachia, Inovasi Baru Cegah Penyebaran DBD*. July 22, 2022. https://sehatnegeriku.kemkes.go.id/baca/umum/20220722/3340692/wolbachia-inovasi-baru-cegah-penyebaran-dbd/ [access: 11.02.2023].

[4] Pusat Krisis Kesehatan Kementrian Kesehatan RI: *Langkah Pencegahan Demam Berdarah Dengue*, March 8, 2022. https://penanggulangankrisis.kemkes.go.id/4-langkah-pencegahan-demam-berdarah-dengue [access: 11.02.2023].

[5] Indonesia, Pemerintah Pusat: *Undang-Undang Nomor 24 Tahun 2007 Tentang Penanggulangan Bencana*. Lembaran Negara Republik Indonesia Tahun 2007 Nomor 66, Sekeretariat Negara, Jakarta 2007. https://peraturan.bpk.go.id/Details/39901/uu-no-24-tahun-2007.

[6] Skidmore A.K.: *Environmental Modelling with GIS and Remote Sensing*. Taylor & Francis, London 2002.

[7] Nordin N.I., Sobri N.M., Ismail N.A., Zulkifli S.N., Razak N.F.A., Mahmud M.: *The classification performance using support vector machine for endemic dengue cases*. Journal of Physics: Conference Series, vol. 1496, 012006. http://doi.org/10.1088/1742-6596/1496/1/012006.

[8] Hamdani H., Hatta H.R., Puspitasari N., Septiarini A., Henderi H.: *Dengue classification method using support vector machines and cross-validation techniques*. IAES International Journal of Artificial Intelligence, vol. 11(3), 2022, pp. 1119–1129. http://doi.org/10.11591/ijai.v11.i3.pp1119-1129.

[9] Mountrakis G., Im J., Ogole C.: *Support vector machines in remote sensing: A review*. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 66(3), 2011, pp. 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001.

[10] Erbek F.S., Özkan C., Taberner M.: *Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities*. International Journal of Remote Sensing, vol. 25(9), 2022, pp. 1733–1748. https://doi.org/10.1080/0143116031000150077.

[11] Liang F., Zhang X., Li H., Yu H., Lin Q., Jiang M., Zhang J.: *Land use classification based on maximum likelihood method*. [in:] Pan J.S., Balas V.E., Chen C.M. (eds.), *Advances in Intelligent Data Analysis and Applications*, Smart Innovation, Systems and Technologies, vol. 253, Springer, Singapore 2022, pp. 133–139. https://doi.org/10.1007/978-981-16-5036-9_15.

[12] Danoedoro P.: *Pengantar Penginderaan Jauh Digital*. Penerbit Andi, Yogyakarta 2012.
[13] Chakraborty A., Sachdeva K., Joshi P.K.: *A Reflection on Image Classifications for Forest Ecology Management: Towards Landscape Mapping and Monitoring*. [in:] Samui P., Sekhar S., Balas V.E. (eds.), *Handbook of Neural Computation*, Elsevier, Amsterdam 2017, pp. 67–85. https://doi.org/10.1016/B978-0-12-811318-9.00004-1.
[14] Badan Standardisasi Nasional: *Klasifikasi penutup lahan – Bagian 1: Skala kecil dan menengah* (SNI 7645-1:2014). BSN, Jakarta 2014. https://202.4.179.213/uploadsfile/sni-7645-1-2014.pdf.
[15] Ganie M.A., Nusrath A.: *Determining the vegetation indices (NDVI) from Landsat 8 satellite data*. International Journal of Advanced Research, vol. 4(8), 2016, pp. 1459–1463. https://doi.org/10.21474/ijar01/1348.
[16] Bouhennache R., Bouden T., Taleb-Ahmed A., Cheddad A.: *A new spectral index for the extraction of built-up land features from Landsat 8 satellite imagery*. Geocarto International, vol. 34(14), 2018, pp. 1531–1551. https://doi.org/10.1080/10106049.2018.1497094.
[17] Landsat Missions: *Landsat 8 (L8) Data Users Handbook*. Department of the Interior U.S. Geological Survey, Reston 2018.
[18] Coll C., Galve J.M., Sanchez J.M., Caselles V.: *Validation of Landsat-7/ETM+ thermal-band calibration and atmospheric correction with ground-based measurements*. IEEE Transactions on Geoscience and Remote Sensing, vol. 48(1), 2010, pp. 547–555. https://doi.org/10.1109/TGRS.2009.2024934.
[19] Valor E., Caselles V.: *Mapping land surface emissivity from NDVI: Application to European, African, and South American areas*. Remote Sensing of Environment, vol. 57(3), 1996, pp. 167–184. https://doi.org/10.1016/0034-4257(96)00039-9.
[20] Kurniadi H., Aprilia E., Utomo J.B., Kurniawan A., Safril A.: *Perbandingan METODE IDW Dan Spline dalam Interpolasi Data Curah Hujan (Studi Kasus Curah Hujan Bulanan Di Jawa Timur Periode 2012–2016)*. Prosiding Seminar Nasional GEOTIK 2018, pp. 213–220.
[21] Motevalli A., Pourghasemi H.R., Zabihi M.: *Assessment of GIS-based Machine Learning Algorithms for Spatial Modeling of Landslide Susceptibility: Case Study in Iran*. [in:] Huang B. (ed.), *Comprehensive Geographic Information Systems*, Elsevier, Amsterdam 2018, pp. 258–280. https://doi.org/10.1016/B978-0-12-409548-9.10461-0.
[22] Cho M.Y., Hoang T.T.: *Feature selection and parameters optimization SVM using particle swarm optimization for fault classification in power distribution systems*. Computational Intelligence and Neuroscience, vol. 2017(3), 2017, 4135465. https://doi.org/10.1155/2017/4135465.
[23] Scavuzzo J.M., Trucco F., Espinosa M., Tauro C.B., Abril M., Scavuzzo C.M., Frery A.C.: *Modeling dengue vector population using remotely sensed data and machine learning*. Acta Tropica, vol. 185, 2018, pp. 167–175. https://doi.org/10.1016/j.actatropica.2018.05.003.

[24] Louis V.R., Phalkey R., Horstick O., Ratanawong P., Smith A.W., Tozan Y., Dambach P.: *Modeling tools for dengue risk mapping – a systematic review*. International Journal of Health Geographics, vol. 13, 2014, 50. https://doi.org/10.1186/1476-072X-13-50.

[25] Widayani P., Yanuar S.R., Yogi H.A.: *Relationship analysis of environmental factor change on the evidence of dengue fever diseases using image transformation (case study: Surakarta City)*. IOP Conference Series: Earth and Environmental Science, vol. 169, 2018, 012061. https://doi.org/10.1088/1755-1315/169/1/012061.

[26] Palaniyandi M.: *The environmental aspects of dengue and chikungunya outbreaks in India: GIS for epidemic control*. International Journal of Mosquito Research, vol. 1(2), 2014, pp. 35–40.

[27] Tjaden N.B., Caminade C., Beierkuhnlein C., Margarete S.T.: *Mosquito borne diseases: Advances in modelling climate-change impacts*. Trends in Parasitology, vol. 34(3), 2018, pp. 227–245. https://doi.org/10.1016/j.pt.2017.11.006.

[28] Yin S., Ren C., Shi Y., Hua J., Yuan H.-Y., Tian L.-W.A.: *A systematic review on modeling methods and influential factors for mapping dengue-related risk in urban settings*. International Journal of Environmental Research and Public Health, vol. 19(22), 2022, 15625. https://doi.org/10.3390/ijerph192215265.

[29] Davis C., Murphy A.K., Bambrick H., Devine G., Frentiu F., Yakob L., Huang X., Li Z., Yang W., Williams G., Hu W.: *A regional suitable conditions index to forecast the impact of climate change on dengue vectorial capacity*. Environmental Research, vol. 195, 2021, 110849. https://doi.org/10.1016/j.envres.2021.110849.

[30] Dickin S.K., Wallace C.J.: *Assessing changing vulnerability to dengue in northeastern Brazil using a water-associated disease index approach*. Global Environmental Change, vol. 29, 2014, pp. 155–164. https://doi.org/10.1016/j.gloenvcha.2014.09.007.

[31] Costa E.A.P.A., Santos E.M.M., Correia J.C., Albuquerque C.M.R.: *Impact of small variations in temperature and humidity on the reproductive activity and survival of Aedes aegypti (Diptera, Culicidae)*. Medical and Veterinary Entomology, vol. 54(3), 2010, pp. 488–493. https://doi.org/10.1590/S0085-56262010000300021.

[32] Kesetyaningsih T.W., Andarini S., Sudarto, Pramoedyo H.: *Correlation between Larvae Free Number with DHF Incidence in Sleman, Yogyakarta, Indonesia*. [in:] *The 2nd International Conference of Medical Health & Health Sciences and the Life Sciences Conference 2016*. http://repository.umy.ac.id/handle/123456789/12653.

[33] Yudhastuti R., Satyabakti P., Basuki H.: *Climate conditions, larvae free number, DHF incidence in Surabaya Indonesia*. Journal of US-China Public Administration, vol. 10(11), 2013, pp. 1043–1049.

[34] Valdez L.D., Sibona G.J., Condat C.A.: *Impact of rainfall on Aedes aegypti populations*. Ecological Modelling, vol. 385, 2018, pp. 96–105. https://doi.org/10.1016/j.ecolmodel.2018.07.003.

[35] Iriani Y.: *Hubungan antara Curah Hujan dan Peningkatan Kasus Demam Berdarah Dengue Anak di Kota Palembang*. Sari Pediatri, vol. 13(6), 2012, pp. 378–383. https://doi.org/10.14238/sp13.6.2012.378-83.

[36] Mukhopadhyay A.K., Tamizharasu W., Satya Babu P., Chandra G., Hati A.K.: *Effect of common salt on laboratory reared immature stages of Aedes aegypti (L)*. Asian Pacific Journal of Tropical Medicine, vol. 3(3), 2010, pp. 173–175. https://doi.org/10.1016/S1995-7645(10)60002-8.

[37] Siddiq A., Shukla N., Pradhan B.: *Predicting dengue fever transmission using machine learning methods*. [in:] *IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2021, Singapore, December 13–16, 2021. IEEE 2021*, IEEE, Piscataway 2021, pp. 21–26. https://doi.org/10.1109/IEEM50564.2021.9672977.

[38] Mizan R.A., Widayani P., Farda N.M.: *Assessment and comparison of machine learning algorithm capability in spatial modeling of dengue fever vulnerability based on Landsat Image 8 Oli/Tirs*. Jurnal Geografi, vol. 13(2), 2021, pp. 211–224. https://doi.org/10.24114/jg.v13i2.21019.