Knowledge Engineering and Data Science (KEDS)
Vol 7, No 1, April 2024, pp. 40–57

pISSN 2597-4602
eISSN 2597-4637

Optimizing Malaria Control: Granular and Cost-Effective Mosquito Habitat Index in Endemic Areas Through Satellite Imagery

Nur Ainun Daulay [a,1], Salwa Rizqina Putri [a,2], Arie Wahyu Wijayanto [a,3,*], Ika Yuni Wulansari [b,4]

[a] *Department of Statistical Computing, Politeknik Statistika STIS*
Jl. Otto Iskandardinata 64C, DKI Jakarta, 13330, Indonesia
[b] *School of Mathematical and Physical Sciences, University of Technology Sydney*
15 Broadway, Ultimo, Sydney, New South Wales 2007, Australia
[1] *222011464@stis.ac.id;* [2] *salwa@stis.ac.id;* [3] *ariewahyu@stis.ac.id*;* [4] *ikayuni.wulansari@uts.edu.au*
** corresponding author*

ARTICLE INFO

Article history:
Received 10 March 2024
Revised 18 April 2024
Accepted 03 May 2024
Published online 27 May 2024

Keywords:
MHSI
Malaria
K-Means
Remote Sensing

A B S T R A C T

Malaria, classified as a tropical disease under the Sustainable Development Goals (SDGs) indicator 3.3, remains a significant global health challenge. In this study, by taking advantage of multiple spectral composite indexes of multisource satellite imagery to capture various geospatial features relevant to the suitability of marsh mosquito habitat, we introduced the Mosquito Habitat Suitability Index (MHSI) to assess potential Anopheles mosquito breeding sites in terms of the vegetation density, water bodies, environment temperature, and humidity in any particular areas. The MHSI integrates the publicly accessible granular level of the normalized difference vegetation index, water index, land surface temperature, and moisture index from cost-effective low and medium-resolution optical satellite data. We focus on West Papua Province, Indonesia, known for diverse ecological conditions and varying malaria prevalence, as a case study area. From the built index, the risk zone map is then formed with the K-Means algorithm. One key finding is the elevated risk in Fakfak Regency, demanding particular attention, as its high-risk area represents 45% of its total. This research aids localized decision-making to combat malaria's unique challenges in West Papua Province which are relevant for implementation in other regions, contributing to SDG-aligned interventions for malaria eradication by 2030.

I. Introduction

Malaria continues to be a critical worldwide health issue, causing significant morbidity and mortality [1]. The World Health Organization (WHO) approximates that around 229 million cases of malaria occurred, resulting in 409.000 deaths in the year 2019 [2]. This vector-borne disease represents a severe obstacle to reaching Sustainable Development Goal 3 (SDG 3): promoting health and well-being for all [3]. Malaria is referred to as a neglected tropical illness on SDG indicator 3.3, and it is expected to end in 2030 [4].

Malaria is most likely to be found in tropical and subtropical countries [5]. This disease is caused by the Plasmodium parasite, transmitted mainly through the bite of an infected female Anopheles mosquito [6]. The Anopheles mosquito thrives in tropical regions due to higher temperatures and humidity [7]. Indonesia is one of the tropical countries that is plagued by malaria [8]. According to the WHO, Indonesia is the second highest contributor to the worldwide malaria case count [9]. Even though the government's efforts to eradicate malaria have materialized in most parts of Indonesia (94%) [10], cases of malaria are still high in eastern Indonesia, for example, the Provinces of West Papua [11]. This indicates that the eradication of Malaria has not been evenly distributed. According to data from Statistics Indonesia (BPS), West Papua had a relatively high malaria incidence rate of 7.380 cases per 1000 inhabitants in 2019 [12]. After Papua, West Papua has the second highest malaria incidence rate in Indonesia [12].

https://doi.org/10.17977/um018v7i12024p40-57

To meet the Indonesia Ministry of Health's target of eliminating malaria by 2030, especially in Papua [13], various efforts have been made. This endeavor includes a variety of strategies, including local leaders' advocacy, mosquito net distribution, monitoring of their use, the availability of antimalarial medications, early disease identification, enhancing health professionals' competency, and cross-program coordination [14]. However, according to the publication of the National Action Plan for Accelerating Malaria Elimination 2020-2026, one of the shortcomings in those efforts is the low receptivity mapping [source]. Yet, geographical variables significantly contribute to the growth of Anopheles mosquitoes, thus these efforts need to be optimized with a better understanding of malaria risk zones [15]. However, to produce a good mapping, direct exploration of the entire West Papua region is required, which is considered less effective in terms of time and cost.

Remote sensing, on the other hand, can be a useful method for mapping possible mosquito habitat zones [16]. The advantages of remote sensing include its ability to provide comprehensive insights into environmental conditions without the need to touch it [17], facilitate the identification of potential breeding sites [18], and help optimize targeted control strategies for mosquito-borne diseases [19]. Recent studies have found that mosquito disease transmission is influenced by rainfall, humidity, and temperature [20]. Significant rain causes a lot of standing water [21] stated that the Normalized Differenced Water Index (NDWI) could detect the surface of the earth that contains water, which can be a breeding ground for mosquitoes [21]. NDWI can also detect flood inundation [22], mangrove distribution [23], agricultural land drought detection [23], and so on. The use of NDWI has been implemented in various countries, such as Indonesia [24], India [25], Iraq [26], and Nepal [27]. Apart from NDWI, satellite imagery can also provide a Normalized Difference Moisture Index (NDMI), which is used to identify the humidity of an area [28]. The moisture of an area has a significant relationship to malaria cases [29][30] have used NDMI as a variable to determine areas where mosquitoes breed. The humidity of an area is also affected by the vegetation in that area [31]. Vegetation density can be detected by the Normalized Difference Vegetation Index (NDVI) [32]. NDWI, NDMI, and NDVI can be obtained from Sentinel-2 satellite imagery provided by Copernicus. Remote sensing can also detect the temperature of an area [33]. One of the satellites that can be utilized is the Moderate Resolution Imaging Spectroradiometer (MODIS) [33], which provides a particular band of Land Surface Temperature (LST) [33].

Of the various advantages offered, remote sensing data has the potential to support the eradication of malaria in Indonesia [34]. To the best of our knowledge, research discussing this is still rare in Indonesia, especially the combination of NDWI, NDMI, NDVI, and LST, which has not been widely explored in Indonesia, especially in West Papua. Therefore, we propose an index that can identify potential mosquito breeding areas formed from several indices with weighted summation. Precisely, this study will (1) calculate the Mosquito Habitat Suitability Index (MHSI) based on multi-source remote sensing representing mosquito breeding areas in the case study area, using the aggregation method of averaging at a 1 km grid level in 2020–2022; (2) provides a map of areas at risk as mosquito breeding sites with a spatial resolution of 1 km based on data for 2020-2022. By establishing a more granular risk zone, it is expected that mapping receptive areas can be generated without the need for direct visits to each area, allowing resources to be allocated more effectively, and targeted measures can be implemented to curb malaria transmission. The results of this research are expected to contribute to efforts to achieve SDG 3 targets regarding the reduction and even eradication of malaria in Indonesia.

II. Methods

A. Study Area

Indonesia is a tropical country with the potential to be a habitat for the breeding of mosquitoes that serve as disease vectors [35]. Particularly in the eastern regions of Indonesia, the prevalence of malaria still presents concerning [11]. Data collected by Statistics Indonesia (BPS) in 2019 revealed that Papua Province had the highest prevalence rate, reaching 64,03 as in Figure 1. This was followed by West Papua with a prevalence rate of 7,38, and East Nusa Tenggara Province (locally known as "Nusa Tenggara Timur" or NTT) ranking third with a prevalence rate of 2,37 [12]. The significant disparities in these prevalence rates confirm the diversity of situations among provinces, where most other provinces exhibit much lower malaria prevalence rates, even below 1.

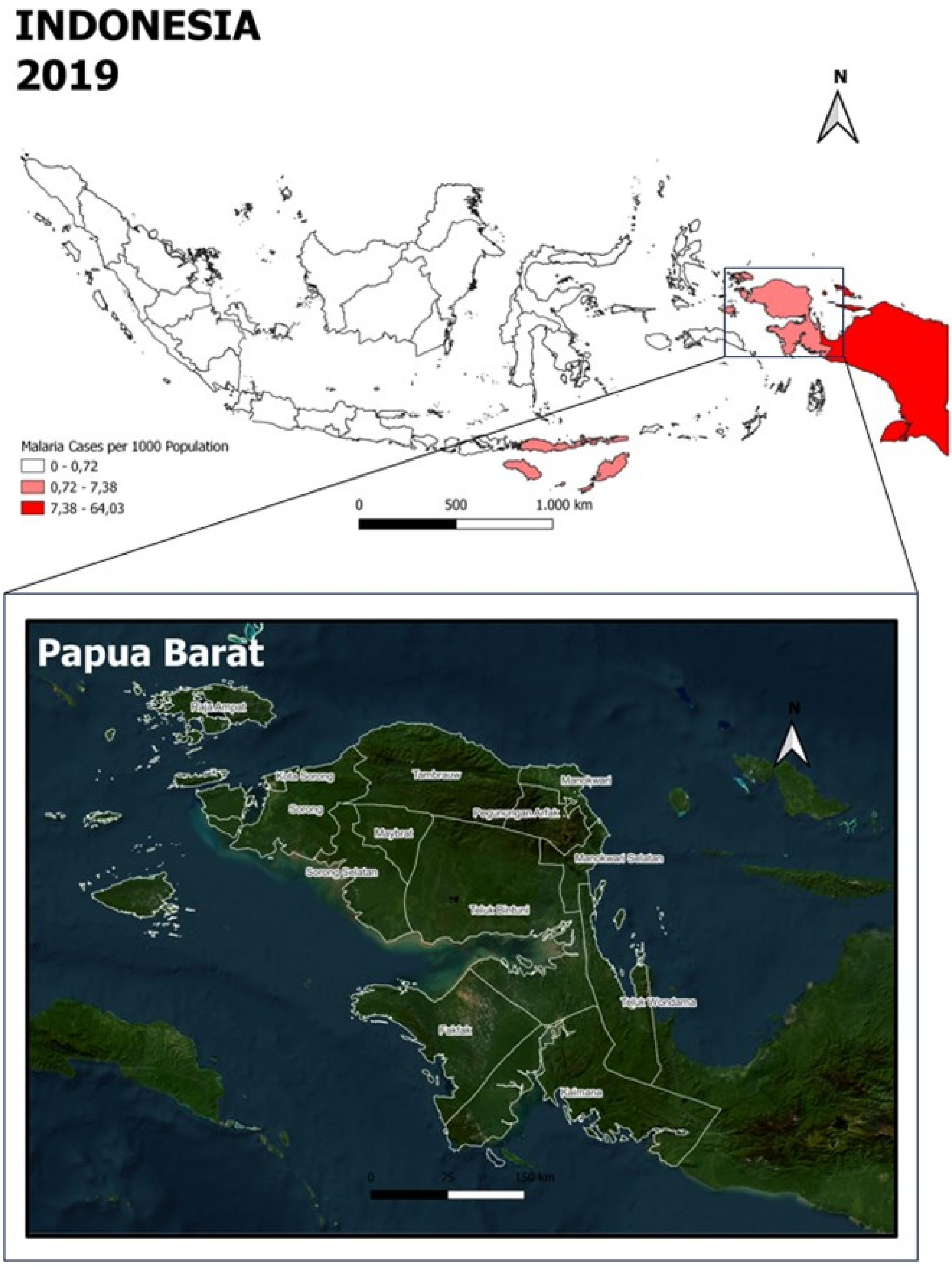

Fig. 1. West Papua, Indonesia as the case study area and its malaria cases (per 1000 population)

B. *Data Used in This Study*

In formulating the MHSI with a multisource approach, our research relies on utilizing satellite imagery encompassing several significant variables, namely NDWI, NDMI, NDVI, and LST. NDWI, as one of the variables, plays a role in identifying water bodies present on the Earth's surface [36]. On the other hand, NDMI plays a crucial role in detecting the level of moisture in a given area [37]. Meanwhile, NDVI is employed to assess vegetation density, aiding in depicting the density of vegetation in specific areas [38]. Furthermore, to comprehend the dynamics of surface temperature, LST information is utilized, which can reveal the temperature of regions in a more detailed manner [39].

All of these variables were obtained and analyzed using Google Earth Engine (GEE), a cloud computing-based platform that is accessible for free [40]. In this approach, the main focus was directed towards our primary study area, West Papua Province, Indonesia. Data collection was conducted across three distinct periods, encompassing the years 2020, 2021, and 2022, in order to provide a comprehensive perspective on mosquito development within the region. NDWI, NDMI, and NDVI were derived from Sentinel-2 satellite imagery, which boasts relatively high resolutions ranging from 10 m to 20 m [41]. On the other hand, LST data was acquired from MODIS satellite imagery [42]. Each dataset obtained from these satellites was scaled by a specific factor [41][42]. Further details on the characteristics of each variable are elaborated in Table 1.

Table 1. Summary of variables

Source	Spatial Resolution	Variable	Band Use	Year Data Analysis	Units	References
Sentinel-2 [41]	10 meter - 20 meter	NDWI	B3 (Green) and B8 (NIR)	The mean value of 2315 cloud masked images	Index	[43][44]
		NDMI	B8 (NIR) and SWIR (B11)		Index	[45][46]
		NDVI	B4 (Red) and B8 (NIR)		Index	[44][46]
MODIS [42]	1000 meter	LST	LST_Day_1 km	The average of 365 images with cloud masking	Kelvin (K)	[45][47]

C. Methodology

The methodology employed in this study traverses a series of meticulous and systematic stages, aimed at formulating and developing the MHSI as a relevant analytical tool for identifying and mapping potential areas as habitats for disease vector mosquitoes. In this section, we will elaborate in detail on the approach taken, including data sources, satellite image processing, and computational methods employed. Data collection was carried out using GEE, while analysis and visualization were conducted using Google Colaboratory and QGIS 3.20.2, resulting in the expected output of a 1 km x 1 km MHSI Map. The executed steps adhere to contemporary methodological guidelines and recognized frameworks within remote sensing and spatial analysis. Thus, the methodology we elucidate here establishes a robust foundation for achieving the objectives of this research. The research framework is systematically illustrated in Figure 2.

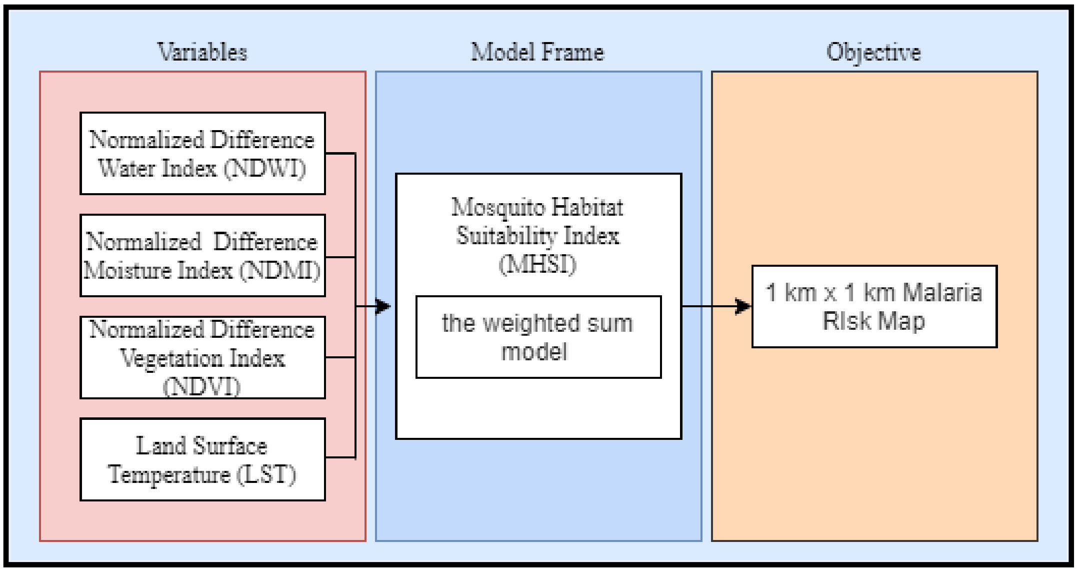

Fig. 2. Research framework

The data was collected from the aforementioned sources. Subsequently, the data underwent preprocessing as a crucial stage [40]. Generally, data preprocessing is conducted to ensure the cleanliness of the data and enhance the quality of the analysis. Satellite image data used in this study was collected over the span of a year in 2020, 2021, and 2022. Each image data underwent 5 preprocessing stages, namely cloud selection, cloud masking, mean reduction, missing value imputation, and band compositing. In the event of missing values within the collected data, we employed the K-Nearest Neighbors (KNN) Imputer for imputation. To obtain NDWI, NDMI, and NDVI, we utilized the following formulas as in (1) to (3).

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \tag{1}$$

$$NDMI = \frac{NIR - SWIR}{NIR + SWIR} \tag{2}$$

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{3}$$

In the methodology of this research, we employ zonal statistical analysis as a central approach. We initiate by subdividing the study area into spatial units of 1 km x 1 km using a grid shapefile. The subsequent process involves implementing zonal statistical analysis on each utilized variable, including NDWI, NDMI, NDVI, and LST. In this step, each variable is scrutinized in detail based on the 1 km x 1 km grid. The outcomes of this zonal statistical analysis offer a comprehensive understanding of the characteristics of each variable within each spatial unit. This approach provides further opportunities for interpretation regarding the distribution and variability of the variables across the entire study area.

To construct the MHSI, transformations are applied to the variables to ensure they share the same range. This uniform range eliminates the dominance of any single variable, thereby enhancing the quality of the obtained analysis. The transformation method employed is MinMaxScaler. MinMaxScaler is a normalization technique that scales data to have a range of 0 to 1 [48]. The following is the MinMaxScaler formula [48] as in (4) to (5), where X_{max} and X_{min} are the maximum and minimum values of each variable.

$$X_{std} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4}$$

$$X_{scaled} = X_{std} * (X_{max} - X_{min}) + X_{min} \tag{5}$$

The MHSI is constructed by overlaying variables that represent the spatial factors of mosquito habitat. We implement a weighted summation to combine the utilized variables. Weighted summation has been commonly applied in various spatial analyses [40]. Equation (6) is the formula, in which p stands for the quantity of overlay variables applied, w_i signifies the weight, and x_i represents the normalized variable value. This study assigns weights for NDWI, NDMI, NDVI, and LST as 12%, 22%, 33%, and 33% respectively [49].

$$MHSI = \sum_{i=1}^{p} w_i \, x_i \tag{6}$$

We clustered MHSI data for 2020-2021 to obtain a risk level for malaria mosquito breeding areas. Clustering uses the K-Means method because it has been widely used in other research, such as who clustered COVID-19 cases [50], who clustered nutritional status [51], who clustered image data [52], who carried out classification and detection of malarial parasite in blood samples using K-Means clustering algorithm [53], carried out clustering of plasmodium falciparum genes to their functional roles using K-Means [54], and carried out identification of Giemsa stains of malaria using K-Means clustering segmentation technique [55].

III. Results and Discussion

This study utilizes the NDWI, NDMI, NDVI, and LST variables as factors for mosquito breeding. The NDWI has values ranging from -1, indicating non-aqueous surfaces, to 1, representing water surfaces. Similarly, the NDMI also has values ranging from -1 to 1, indicating that higher NDMI values correspond to higher humidity in the area. Like NDWI and NDMI, the NDVI also ranges from -1 to 1, where higher NDVI values signify greater vegetation in the area. In contrast to these three variables, the LST is presented in Kelvin units, although we converted it to Celsius. As present in Figure 3, all variables are compared for each period of this study, namely 2020, 2021, and 2022. Figure 4 until Figure 7 illustrate the distribution of each variable for each year. We classified each variable value into three categories - low, medium, and high - using the Natural Breaks Jenks method.

```
Function convertKelvinToCelsius(LST_kelvin):
    LST_celsius = LST_kelvin - 273.15
    Return LST_celsius

Function applyNaturalBreaksJenks(data):
    breakpoints = JenksNaturalBreaks(data, 3)
    Return breakpoints

Procedure main():
    For each variable (NDWI, NDMI, NDVI, LST) and year (2020, 2021, 2022):
        variable_data = getVariableData(variable, year)

        If variable == "LST":
            variable_data = convertKelvinToCelsius(variable_data)

        breakpoints = applyNaturalBreaksJenks(variable_data)

        classified_data = classifyData(variable_data, breakpoints)

        Output classified data
```

Fig. 3. Pseudocode for distributing each variable

Figure 4 is classified using the Natural Breaks Jenks method and produces 3 colors: white (low), light blue (medium), and dark blue (high). Based on the NDWI data obtained in 2020, the classification results show the range of Low (-0.8180 to -0.5873), Medium (-0.5874 to -0.2514), and High (-0.2515 to 0.8677). The 2021 data shows the classification range of Low (-0.8617 to -0.5917), Medium (-0.5918 to -0.2488), and High (-0.2489 to 0.9747). The 2022 data yields classification ranges of Low (-0.8405 to -0.5648), Medium (-0.5649 to -0.2159), and High (-0.2160 to 0.9618). It is noticeable that in some areas of the Arfak Mountains, the NDWI values are increasing from year to year. In 2020, there were only 2 dark blue areas, which are Lake Anggi Giji and Lake Anggi Gida. Then, from 2021 to 2022, the number of light blue and dark blue areas increases. This indicates that there are more water puddles in the Arfak Mountains region. The lower the NDWI value, the less water there is on the surface; conversely, the higher the NDWI value, the more water there is on the surface.

Figure 5 is classified using the Natural Breaks Jenks method and produces 3 colors: dark blue (low), gray (medium), and yellow (high). Based on the NDMI data obtained in 2020, the classification results show the range of Low (-0.5722 to 0.1322), Medium (0.1323 to 0.3050), and High (0.3051 to 0.8120). In contrast, the 2021 data shows the classification range of Low (-0.8447 to 0.1284), Medium (0.1285 to 0.2998), and High (0.2999 to 0.8344). Similarly, the 2022 data yields classification ranges of Low (-0.4922 to 0.1268), Medium (0.1269 to 0.3046), and High (0.3047 to 0.7556). It is noticeable that in some areas of Fakfak, there are changes in NDMI values. The NDMI also has values ranging from -1 to 1, indicating that higher NDMI values correspond to higher humidity in the area.

Figure 6 is classified using the Natural Breaks Jenks method and produces 3 colors: white (low), light green (medium), and dark green (high). Based on the NDVI data obtained in 2020, the classification results show the range of Low (-0.7918 to 0.2844), Medium (0.2845 to 0.6768), and High (0.6769 to 0.9239). In contrast, the 2021 data shows the classification range of Low (-0.8304 to 0.2863), Medium (0.2864 to 0.6833), and High (0.6834 to 0.9310). Similarly, the 2022 data yields classification ranges of Low (-0.9142 to 0.2459), Medium (0.2460 to 0.6473), and High (0.6474 to 0.9304). Figure 6 shows a decrease in NDVI values in some areas of the Arfak Mountains from year to year. The higher the NDVI value, the higher the vegetation density in that area.

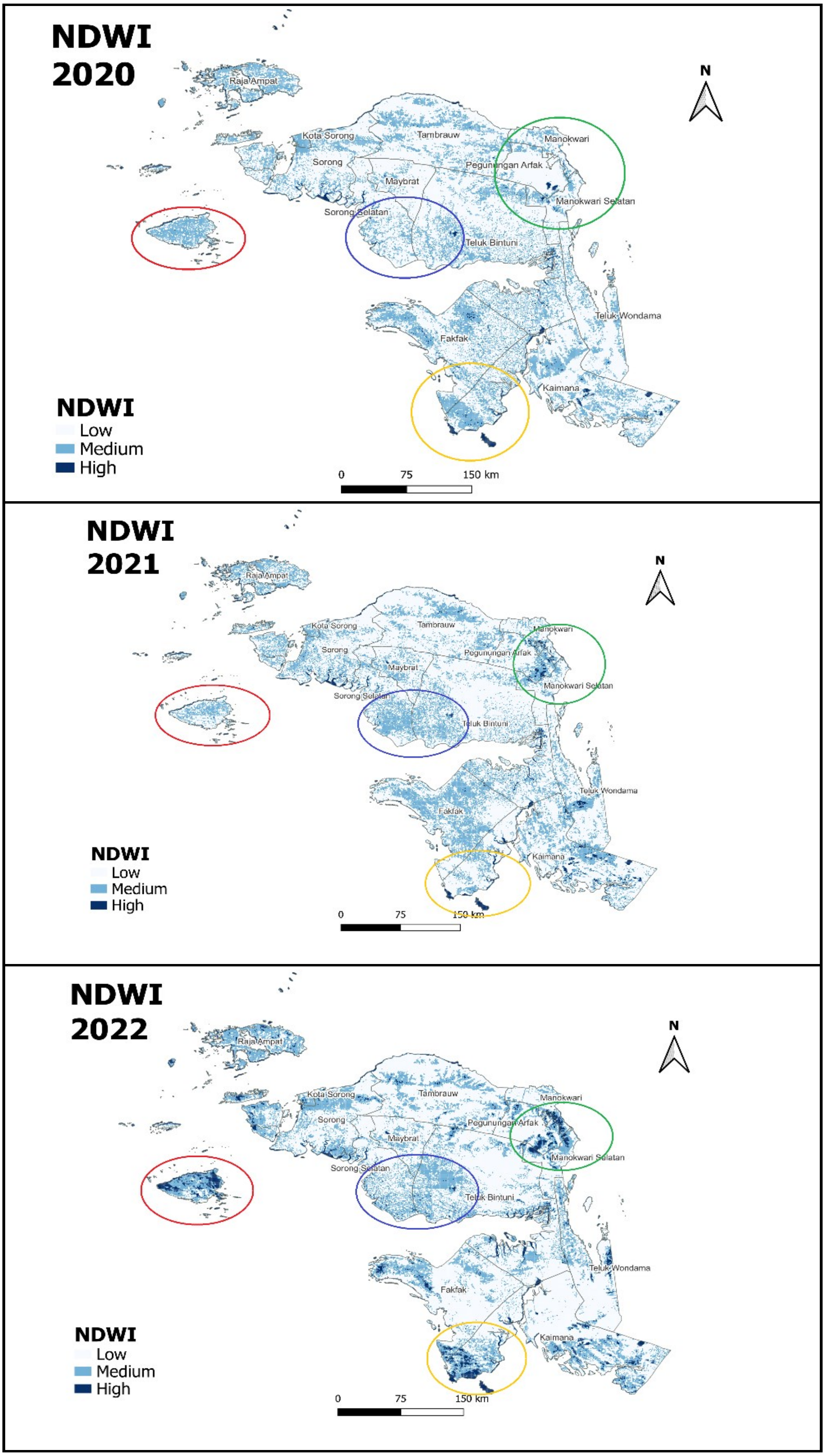

Fig. 4. Spatial mapping of humidity level in West Papua Province measured by the NDWI

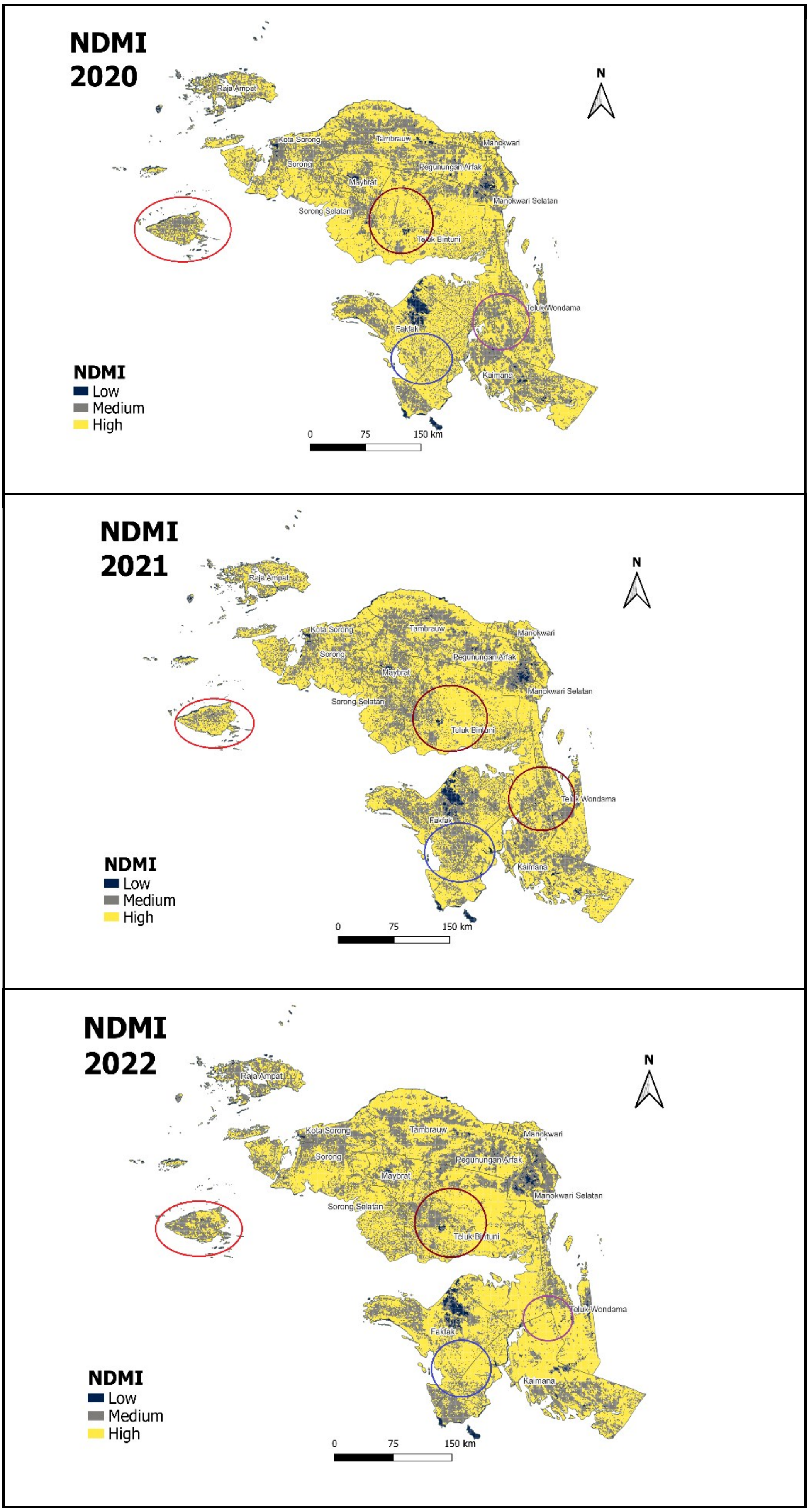

Fig. 5. Spatial mapping of moisture level in West Papua Province measured by the NDMI

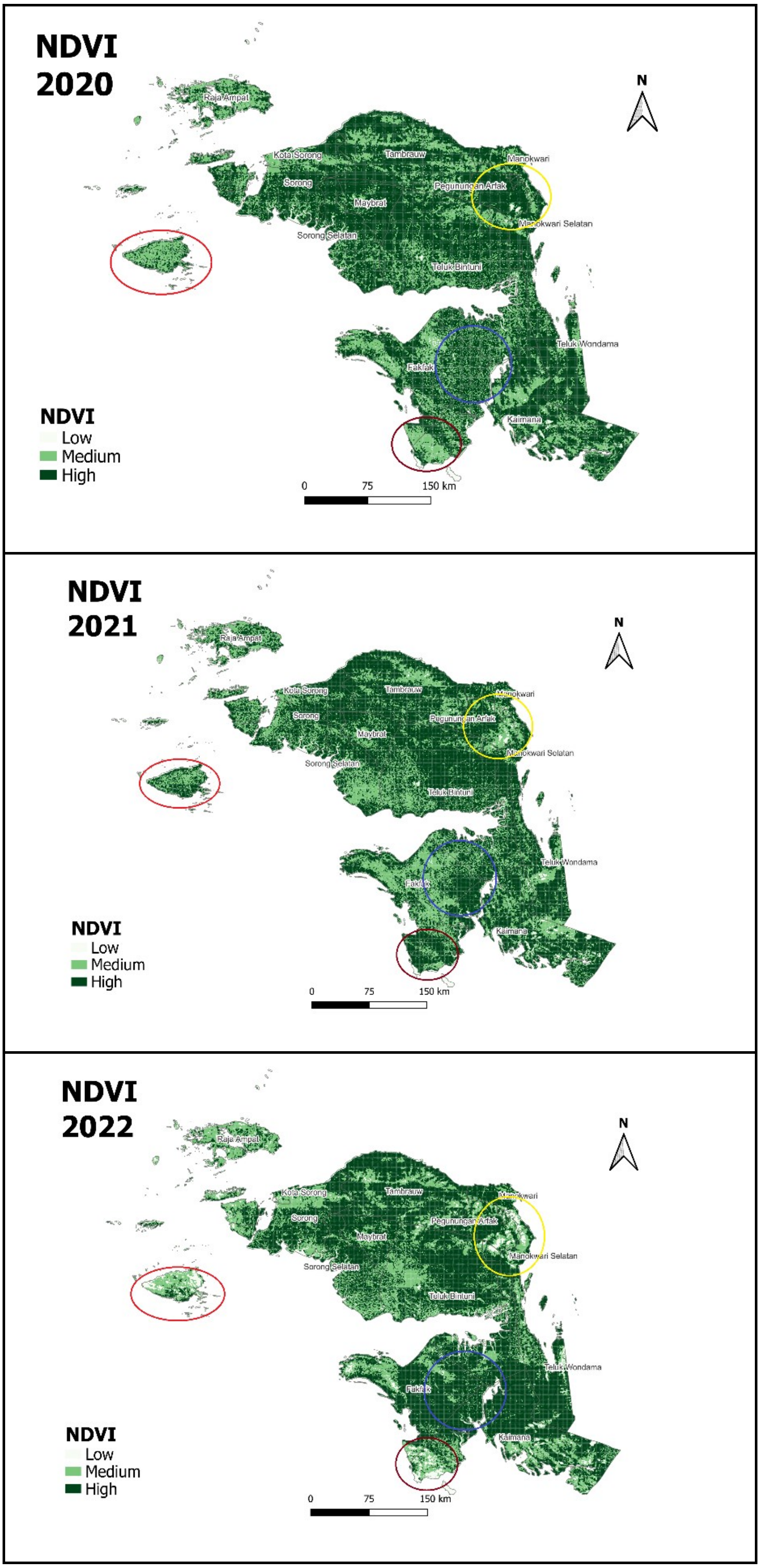

Fig. 6. Spatial mapping of vegetation density in West Papua Province measured by the NDVI

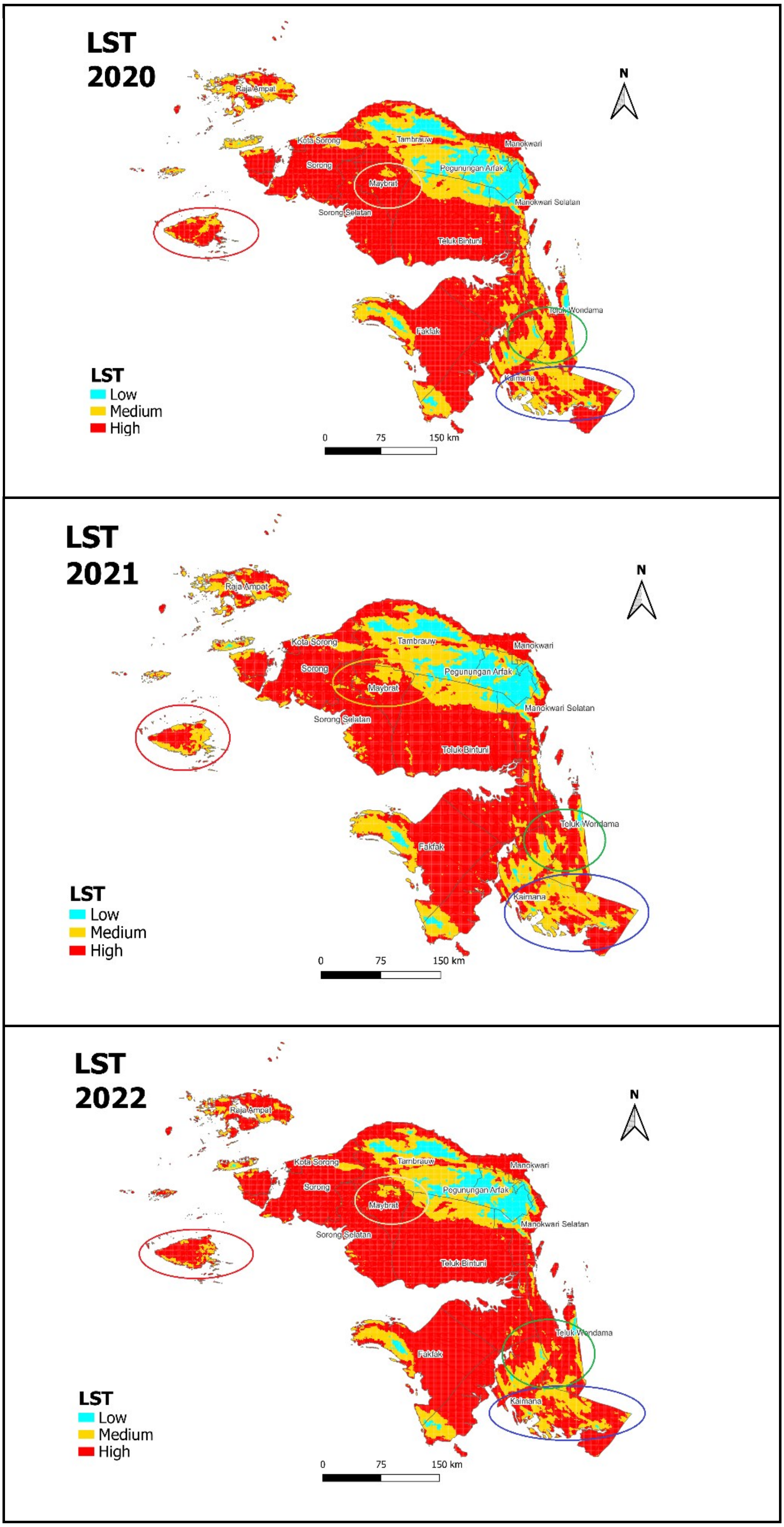

Fig. 7. Spatial mapping of area temperature in West Papua Province measured by the LST

Figure 7 is classified using the Natural Breaks Jenks method and produces 3 colors in Celsius: blue (low), yellow (medium), and red (high). Based on the LST data obtained in 2020, the classification results show the range of Low (13.5943 to 22.9751), Medium (22.9752 to 26.3774), and High (26.3775 to 34.6979). In contrast, the 2021 data shows the classification range of Low (11.9936 to 22.8823), Medium (22.8824 to 26.3163), and High (26.3164 to 34.7249). Similarly, the 2022 data yields classification ranges of Low (13.9372 to 22.5171), Medium (22.5172 to 25.8498), and High (25.8499 to 33.8600). Overall, the temperature in West Papua has not changed much from 2020 to 2022. Looking more closely at Figure 7, there are temperature changes in the Maybrat region from year to year, although not significant.

To establish the MHSI, MinMaxScaler transformations were applied to the three variables: NDWI, NDMI, and NDVI. Upon transformation, these three variables would have a range of 0 to 1. In contrast to these three variables, we rendered the LST variable binary. LST is assigned a value of 1 when the temperature falls within the range of 23 to 29 degrees Celsius, while it takes the value of 0 otherwise. This approach is taken due to the optimal breeding temperature for mosquitoes lying within the 23 to 29 degrees Celsius range [49]. Subsequently, the MHSI is constructed using weighted summation, with weights assigned to NDWI, NDMI, NDVI, and LST being 12%, 22%, 33%, and 33%, respectively [49]. So far, there have been no established guidelines for calculating a malaria vulnerability index based on environmental aspects in Indonesia. Additionally, the presented malaria prevalence data is still very limited. Therefore, the weight selection in this study refers to [49], which calculates the weight of environmental variables based on remote sensing data to construct predictions of malaria risk areas by considering the distribution of mosquitoes in a given region. Pseudocode for forming MHSI can be seen in Figure 8.

```
Function minMaxScaler(data):
    scaled_data = (data - min(data)) / (max(data) - min(data))
    Return scaled_data

Function temperatureCoding(LST_data):
    For each temperature in LST_data:
        If temperature is between 23 and 29 degrees Celsius:
            code = 1
        Else:
            code = 0
    Return coded_data

Function weightedSum(NDWI, NDMI, NDVI, LST):
    weighted_sum = (0.12 * NDWI) + (0.22 * NDMI) + (0.33 * NDVI) + (0.33 *
LST)
    Return weighted_sum

Function classifyMHSI(weighted_sum):
    // Classification thresholds can be predefined or determined dynamically
    If weighted_sum < threshold_low:
        Return "Low"
    Else If weighted_sum < threshold_medium:
        Return "Medium"
    Else:
        Return "High"

Procedure main():
    For each year (2020, 2021, 2022):
        // Step 1: Min-Max scaling for NDWI, NDMI, and NDVI
        scaled_NDWI = minMaxScaler(NDWI_data_2020)
        scaled_NDMI = minMaxScaler(NDMI_data_2020)
        scaled_NDVI = minMaxScaler(NDVI_data_2020)

        // Step 2: Temperature coding for LST
        coded_LST = temperatureCoding(LST_data_2020)

        // Step 3: Weighted sum calculation
        weighted_sum_2020 = weightedSum(scaled_NDWI, scaled_NDMI,
scaled_NDVI, coded_LST)

        // Step 4: Classify weighted sum into MHSI categories
        classified_MHSI = classifyMHSI(weighted_sum_2020)

        // Repeat steps 1-4 for 2021 and 2022 with respective data

        // Output MHSI map for each year
        Output MHSI_map_2020
        Output MHSI_map_2021
        Output MHSI_map_2022
```

Fig. 8. Pseudocode for forming MHSI

Figure 9 is classified using the Natural Breaks Jenks method and produces 3 colors: white (low), pink (medium), and brown (high). Based on the MHSI values obtained in 2020, the classification results show the range of Low (0.1872 to 0.5473), Medium (0.5474 to 0.7253), and High (0.7254 to 0.8447). In contrast, the 2021 data shows the classification range of Low (0.2968 to 0.5666), Medium (0.5667 to 0.7528), and High (0.7529 to 0.8569). Similarly, the 2022 data yields classification ranges of Low (0.1567 to 0.5628), Medium (0.5629 to 0.7390), and High (0.7391 to 0.8518). The higher the MHSI value, the more potential the area is for mosquito breeding.

Figure 9 illustrates a strikingly similar pattern, indicating that a majority of the areas with low MHSI values are situated in the Pegunungan Arfak Regency. The Pegunungan Arfak region is positioned at an elevation of 800 to 3,000 meters above sea level [56]. As an area's elevation increases, malaria cases tend to decrease [57][58]. Furthermore, the visualization portrays that the mosquito breeding habitat is progressively diminishing from year to year, leading to a reduction in national malaria cases [59]. Notably, a striking trend emerges as we progress from 2020 to 2022, with MHSI values showing a consistent decline over this period. The year 2022, in particular, highlights a significant reduction in MHSI values, reflecting a diminishing suitability for mosquito breeding habitats. This decline in habitat suitability aligns with the broader trends in Indonesia regarding malaria cases. As reported, Indonesia has been making substantial progress in combating malaria, with declining cases noted nationwide [56].

Malaria risk mapping was conducted in West Papua Province based on the MHSI obtained from 2020 to 2022. We labeled areas that never had a high index in all three years of the research period as "BLOCK BOUNDARY" [49]. Then, other areas were clustered using the K-Means method for ease of interpretation. We utilized a parameter K set at 5, leading to the identification of 5 distinct risk levels. Figure 10 shows the pseudocode for forming risk map.

```
Procedure classifyRiskLevel(MHSI_map):
    // Procedure to classify risk levels based on MHSI values
    For each grid in MHSI_map:
        If grid is never classified as "high" for all three years
consecutively:
            risk_level = "block boundary"
        Else:
            Perform K-Means clustering with K=5 on the MHSI values of
all three years
            // Implementation of K-Means algorithm with 5 clusters

            If grid belongs to cluster 5:
                risk_level = "highest risk"
            Else if grid belongs to cluster 4:
                risk_level = "high risk"
            Else if grid belongs to cluster 3:
                risk_level = "moderate risk"
            Else if grid belongs to cluster 2:
                risk_level = "low risk"
            Else:
                risk_level = "lowest risk"

    Return classified_risk_map
```

Fig. 10. Pseudocode for forming risk map

Figure 11 shows the distribution of areas according to the risk level of malaria vector mosquito breeding. Supporting our previous discussion, Figure 8 indicates that the Arfak Mountains are not included in the 5 risk levels. Areas with the highest risk are spread across most of West Papua. However, the purple color (symbolizing the highest risk level) is concentrated in the western and southwestern parts of West Papua. This aligns with the fact that the elevation of the western and southwestern regions is not higher than the northern part of West Papua. The higher the altitude of an area, the lower the number of malaria cases [57][56].

Figure 12 shows the topography and elevation of West Papua obtained from Google Maps. The Figure reveals that the Arfak Mountains District (not labeled on Google Maps but adjacent to Manokwari District), Tindawi in Manokwari District, and Koor or Kwor in Tambrauw District are areas with higher elevation than others.

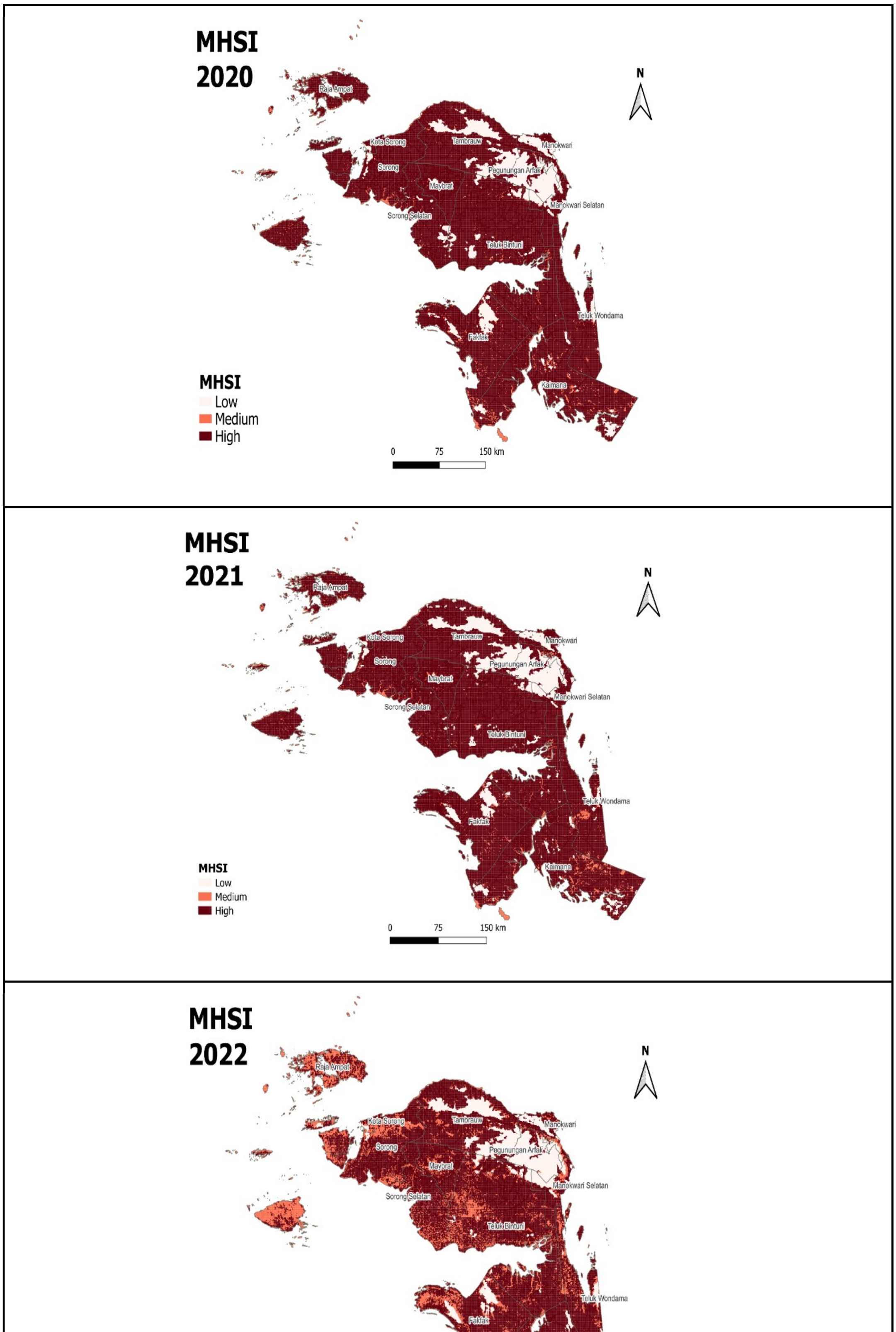

Fig. 9. Spatial distribution of the MHSI in West Papua Province

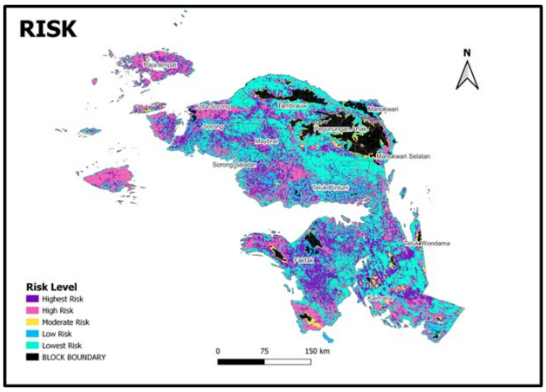

Fig. 11. Mapping the risk zones

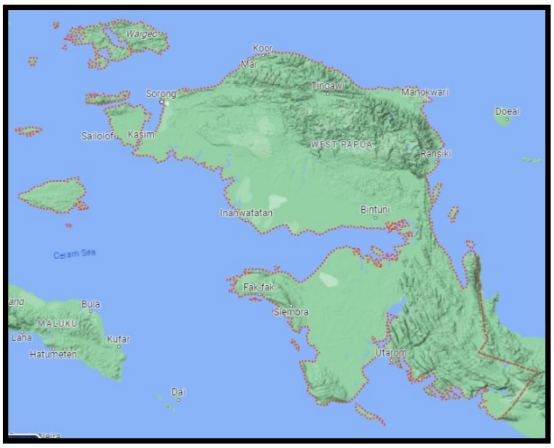

Fig. 12. Topography and elevation map of West Papua

Additionally, we calculated the area of each district based on the risk level with computational capability. Figure 13 shows that most areas with the highest risk are in the districts of Teluk Bintuni, Fakfak, and Kaimana. The area of Teluk Bintuni with the highest risk is 8498 km², followed by Fakfak and Kaimana, with areas of 6462 km² and 6247 km², respectively. Although Teluk Bintuni has the largest area with the highest risk, Fakfak requires more attention. This is because the total area of Fakfak is only 14320 km², meaning that the percentage of the area with the highest risk in Fakfak is 45% of the total Fakfak area. In contrast, Teluk Bintuni and Kaimana have total areas of 20840.33 km² and 16241.84 km², respectively, indicating that the percentage of the area with the highest risk is 41% and 38%.

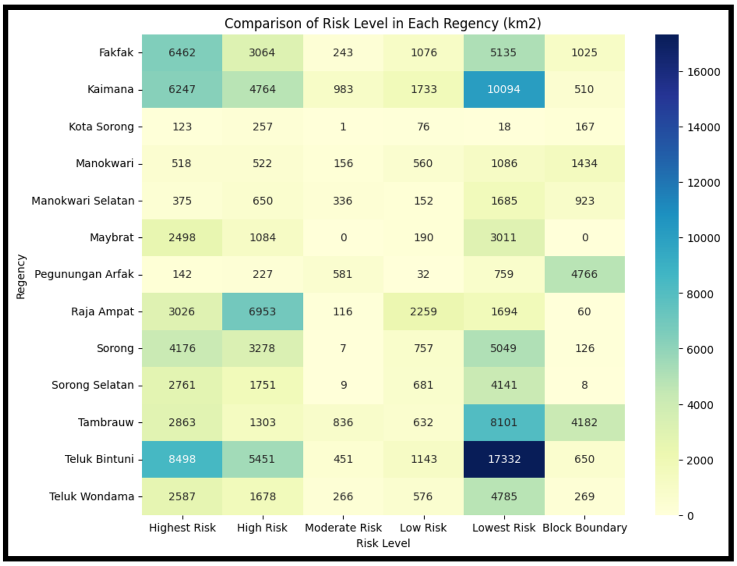

Fig. 13. Comparison of risk level in each regency (km^2)

Using spatial variables at varying resolutions can have a notable impact on analytical outcomes. Differences in spatial resolution can lead to varying information, which, in turn, affects the precision of mapping results [60][61]. To illustrate, when it comes to mapping forest stock volume, the accuracy of mapping is often lower when high spatial resolution optical imagery (e.g., one meter or less) is used compared to medium-resolution imagery (greater than 10 meters), even when the same features and methods are applied [60]. To address this limitation, researchers in the field have utilized GF-2 imagery, which was adjusted to spatial resolutions ranging from 1 to 30 meters. This adjustment allowed them to explore the connection between feature spatial resolution and the accuracy of forest stock volume mapping [60]. Their findings highlighted the substantial impact of feature spatial resolution on the performance of the modeling used to estimate forest stock volume [60]. Furthermore, another study has emphasized the influence of spatial resolution choices on model outcomes. This study has demonstrated how the selection of spatial resolution and scale can affect both mathematical and statistical models [61].

With the abundant potential of remote sensing technology, especially satellite imagery analytics [62][63][64] government efforts to achieve SDGs in numerous real-world use cases are highly supported [65][66][67]. The results provided show that there is potential for implementing malaria risk mapping with remote sensing in other endemic areas in Indonesia or even other regions. The advantage of precise characterizations of land uses, sea surfaces, and land coverings that are inexpensive and quickly updated enables researchers to provide solutions not only in agriculture, urban studies, demographic, socio-economic, environments, but also in health monitoring such we propose in this study [68][69][70][71].

IV. Conclusions

This paper examines the potential of remote sensing and spatial analysis in mapping the malaria risk in West Papua Province, Indonesia. By formulating the MHSI and applying the K-Means clustering method, we successfully identified regions with Highest Risk, High Risk, Moderate Risk, Low Risk, and Lowest Risk for the breeding of malaria vector mosquitoes. The key findings indicate that most of the high-risk areas are located around the western and southwestern regions of West Papua, characterized by lower elevations.

This outcome is consistent with the fact that malaria vector mosquitoes tend to thrive in areas with warmer climatic conditions. Additionally, our analysis also measured the extent of high-risk areas in each district. While the Teluk Bintuni District has the largest area with the highest risk, the Fakfak District deserves special attention due to its relatively higher percentage of high-risk areas compared to other regions. The results of this study can serve as a crucial foundation for policymakers in their efforts to control and prevent malaria in West Papua. The use of remote sensing data and spatial analysis provides a more comprehensive picture of malaria risk distribution, which can be utilized to direct limited resources to areas in need of further intervention. Thus, this study contributes to the global goal of eradicating malaria by 2030.

Declarations

Author contribution

All authors contributed equally as the main contributor of this paper. All authors read and approved the final paper.

Funding statement

The authors are grateful for the support provided by Politeknik Statistika STIS.

Conflict of interest

The authors declare no known conflict of financial interest or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

Reprints and permission information are available at http://journal2.um.ac.id/index.php/keds.

Publisher's Note: Department of Electrical Engineering and Informatics - Universitas Negeri Malang remains neutral with regard to jurisdictional claims and institutional affiliations.

References

[1] J. Hwang, K. A. Cullen, S. P. Kachur, P. M. Arguin, and J. K. Baird, “Severe morbidity and mortality risk from malaria in the United States, 1985-2011,” Open Forum Infect Dis, vol. 1, no. 1, Mar. 2014.

[2] T. Fatima, S. Ali, M. Zehra, and N. Raufi, “Revolutionizing malaria prevention: RSTS vaccine takes center stage in Afghanistan: correspondence,” International Journal of Surgery: Global Health, vol. 6, no. 4, Jul. 2023.

[3] United Nations, “Goal 3: Ensure healthy lives and promote well-being for all at all ages.” (Accessed: Mar. 05, 2024)

[4] World Health Organization, “SDG Target 3.3 End the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases.” (Accessed: Mar. 05, 2024)

[5] D. Rawal, “An overview of natural history of the human malaria,” Int J Mosq Res, vol. 7, no. 2, pp. 8–10, 2020.

[6] S. Piedrahita, N. Álvarez, N. Naranjo-Díaz, S. Bickersmith, J. E. Conn, and M. M. Correa, “Anopheles blood meal sources and entomological indicators related to Plasmodium transmission in malaria endemic areas of Colombia,” Acta Trop, vol. 233, 2022.

[7] S. Zhou and Z. Li, “Malaria Risk and Control,” In Prevention and Control of Infectious Diseases in BRI Countries, pp. 111–117, 2021.

[8] M. H. Herawati et al., “Service availability and readiness of malaria surveillance information systems implementation at primary health centers in Indonesia,” PLoS One, vol. 18, no. 4 April, pp. 1–13, 2023.

[9] The Lancet Regional Health – Southeast Asia, “2030 – Countdown to malaria elimination in India and southeast Asia,” Lancet Reg. Heal. - Southeast Asia, vol. 2, p. 100033, Jul. 2022.

[10] J. T. Bandzuh et al., “Knowledge, attitudes, and practices of Anopheles mosquito control through insecticide treated nets and community-based health programs to prevent malaria in East Sumba Island, Indonesia,” PLOS Global Public Health, vol. 2, no. 9, p. e0000241, 2022.

[11] M. Ipa, M. Widawati, A. D. Laksono, I. Kusrini, and P. W. Dhewantara, “Variation of preventive practices and its association with malaria infection in eastern Indonesia: Findings from community-based survey,” PLoS One, vol. 15, no. 5, pp. 1–18, 2020.

[12] Badan Pusat Statistik, “Kejadian Malaria Per 1000 Orang, 2019-2020.” (Accessed: Mar. 05, 2024)

[13] Rokom, “Kejar Target Bebas Malaria 2030, Kemenkes Tetapkan 5 Regional Target Eliminasi,” Sehat Negeriku, 2022. (Accessed Mar. 05, 2024)

[14] Kementerian Kesehatan Indonesia, “Kasus Malaria di Indonesia Menurun, NTT Jadi Provinsi Pertama di Kawasan Timur Berhasil Eliminasi Malaria.” 2021. (Accessed Mar. 05, 2024)

[15] World Health Organization, “Malaria eradication : benefits , future scenarios A report of the Strategic Advisory Group,” 2020.

[16] Z. Kenyeres et al., “Cost–benefit analysis of remote sensing data types for mapping mosquito breeding sites,” Spat. Inf. Res., vol. 31, no. 4, pp. 419–428, Aug. 2023.

[17] A. M. Lechner, G. M. Foody, and D. S. Boyd, “Applications in Remote Sensing to Forest Ecology and Management,” One Earth, vol. 2, no. 5, pp. 405–412, 2020.

[18] M. Govindaraju, J. F. Banu, K. Sridhar, P. Vignesh, and K. Rangarajan, “Identification and Mapping of Mosquito Breeding Habitats in Tiruchirappalli City Using Remote Sensing and GIS Technologies,” in Genetically Modified and other Innovative Vector Control Technologies, Singapore: Springer Singapore, 2021, pp. 379–387.

[19] T. Catry et al., “Wetlands and Malaria in the Amazon: Guidelines for the Use of Synthetic Aperture Radar Remote-Sensing,” Int. J. Environ. Res. Public Health, vol. 15, no. 3, p. 468, Mar. 2018.

[20] A. S. M. M. Kamal et al., “Relationship between Urban Environmental Components and Dengue Prevalence in Dhaka City—An Approach of Spatial Analysis of Satellite Remote Sensing, Hydro-Climatic, and Census Dengue Data,” Int. J. Environ. Res. Public Health, vol. 20, no. 5, p. 3858, Feb. 2023.

[21] S. McFeeters, “Using the Normalized Difference Water Index (NDWI) within a Geographic Information System to Detect Swimming Pools for Mosquito Abatement: A Practical Approach,” Remote Sens., vol. 5, no. 7, pp. 3544–3561, Jul. 2013.

[22] A. G. Ramadhan, H. H. Handayani, and M. R. Darminto, “Analisis Peta Rawan Banjir Metode Pembobotan dan Peta Genangan Banjir Metode NDWI terhadap Kejadian Banjir (Studi Kasus: Kabupaten Sidoarjo),” Geoid, vol. 17, no. 2, p. 232, Apr. 2022.

[23] Y. A. Singgalen and D. Manongga, “Monitoring of Mangrove Ecotourism Area using NDVI, NDWI, and CNRI in Dodola Island, Morotai Island Regency, Indonesia,” J. Ilmu dan Teknol. Kelaut. Trop., vol. 14, no. 1, pp. 95–108, Apr. 2022.

[24] I. P. G. N. Nandaka, “Pemanfaatan Citra Satelit-2 Dan Google Earth Engine Untuk Identifikasi Badan Air Permanen (Studi Kasus : Sungai Bengawan Solo Hilir Tahun 2018-2022),” Institut Teknologi Sepuluh Nopember, 2023.

[25] U. Chirala and B. Pedada, “Hydrogeomorphology, NDWI and, NDVI of the Meghadrigedda Sub-Watersheds for Optimal Utilization of Resources, Visakhapatnam District, Andhra Pradesh-India Using Landsat Data 2000 and sentinel Data 2020,” Int. J. Geosci., vol. 12, no. 06, pp. 584–604, 2021.

[26] I. F. Ibraheem and M. Al-hadithi, “Remote Sensing Utilization for the Modelling of Land Surface Temperature for Sustainable City Development,” vol. 8, no. 7, pp. 95–106, 2022.

[27] A. Subedi and T. Dev Acharya, “Small Water Bodies Detection and Evaluation using Normalized Difference Water Index (NDWI) from Landsat Image in Western Terai, Nepal Exploring the Status and Distribution of Owls and its Conservation Initiatives in the Raghuganga Gaupalika of Myagdi District, Nepal View project Urban Landcover and their change in Nepal View project,” no. July, pp. 88–96, 2021.

[28] Y. S. Sari, F. Aina Rizky, I. Ranggadara, N. R. Kurnianda, I. Prihandi, and Suhendra, “Comparison of Feature Extraction to Test Dryness and Moisture Levels in Burned Restoration Areas Using Linear Discriminant Analysis,” in 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Feb. 2023, pp. 800–804.

[29] M. Mohammadkhani, N. Khanjani, B. Bakhtiari, S. M. Tabatabai, and K. Sheikhzadeh, "The Relation Between Climatic Factors and Malaria Incidence in Sistan and Baluchestan, Iran," SAGE Open, vol. 9, no. 3, p. 215824401986420, Jul. 2019.
[30] S. Román-Pérez, R. Aguirre-Gómez, J. E. Hernández-Ávila, L. B. Íñiguez-Rojas, R. Santos-Luna, and F. Correa-Morales, "Identification of Risk Areas of Dengue Transmission in Culiacan, Mexico," ISPRS Int. J. Geo-Information, vol. 12, no. 6, p. 221, May 2023.
[31] W. Peng, T. Kuang, and S. Tao, "Quantifying influences of natural factors on vegetation NDVI changes based on geographical detector in Sichuan, western China," J Clean Prod, vol. 233, pp. 353–367, Oct. 2019.
[32] B. Boiarskii, "Comparison of NDVI and NDRE Indices to Detect Differences in Vegetation and Chlorophyll Content," J. Mech. Contin. Math. Sci., vol. spl1, no. 4, Nov. 2019.
[33] W. Zhao, H. Wu, G. Yin, and S.-B. Duan, "Normalization of the temporal effect on the MODIS land surface temperature product using random forest regression," ISPRS J. Photogramm. Remote Sens., vol. 152, pp. 109–118, Jun. 2019.
[34] F. Kogan, Remote Sensing for Malaria. Cham: Springer International Publishing, 2020.
[35] A. Ratnasari, A. R. Jabal, N. Rahma, S. N. Rahmi, M. Karmila, and I. Wahid, "The ecology of aedes aegypti and aedes albopictus larvae habitat in coastal areas of South Sulawesi, Indonesia," Biodiversitas J. Biol. Divers., vol. 21, no. 10, Sep. 2020.
[36] S. Singh and M. L. Kansal, "Chamoli flash-flood mapping and evaluation with a supervised classifier and NDWI thresholding using Sentinel-2 optical data in Google earth engine," Earth Sci. Informatics, vol. 15, no. 2, pp. 1073–1086, Jun. 2022.
[37] O. Strashok, M. Ziemiańska, and V. Strashok, "Evaluation and Correlation of Normalized Vegetation Index and Moisture Index in Kyiv (2017–2021)," J. Ecol. Eng., vol. 23, no. 9, pp. 212–218, Sep. 2022.
[38] K. M. Al-Kindi, R. Al Nadhairi, and S. Al Akhzami, "Dynamic Change in Normalised Vegetation Index (NDVI) from 2015 to 2021 in Dhofar, Southern Oman in Response to the Climate Change," Agriculture, vol. 13, no. 3, p. 592, Feb. 2023.
[39] S. Nasseri, B. Farhadi Bansouleh, and A. Azari, "Estimation of land surface temperature in agricultural lands using Sentinel 2 images: A case study for sunflower fields," Irrig. Drain., vol. 72, no. 3, pp. 796–806, Jul. 2023.
[40] S. R. Putri, A. W. Wijayanto, and A. D. Sakti, "Developing Relative Spatial Poverty Index Using Integrated Remote Sensing and Geospatial Big Data Approach: A Case Study of East Java, Indonesia," ISPRS Int. J. Geo-Information, vol. 11, no. 5, p. 275, Apr. 2022.
[41] Google, "Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A." (Accessed Mar. 05, 2024)
[42] Google, "MOD11A1.061 Terra Land Surface Temperature and Emissivity Daily Global 1km." (Accessed Mar. 05, 2024)
[43] M. Schmieder, F. Holl, M. L. Fotteler, M. Ortl, E. Buchner, and W. Swoboda, "Remote sensing and on-site characterization of wetlands as potential habitats for malaria vectors – A pilot study in southern Germany," in 2020 IEEE Global Humanitarian Technology Conference (GHTC), Oct. 2020, pp. 1–4.
[44] O. E. Malahlela, C. Adjorlolo, and J. M. Olwoch, "Mapping the spatial distribution of Lippia javanica (Burm. f.) Spreng using Sentinel-2 and SRTM-derived topographic data in malaria endemic environment," Ecol. Modell., vol. 392, pp. 147–158, Jan. 2019.
[45] M. C. Wimberly, D. M. Nekorchuk, and R. R. Kankanala, "Cloud-based applications for accessing satellite Earth observations to support malaria early warning," Sci. Data, vol. 9, no. 1, p. 208, May 2022.
[46] S. A. Ali and A. Ahmad, "Spatial susceptibility analysis of vector-borne diseases in KMC using geospatial technique and MCDM approach," Model. Earth Syst. Environ., vol. 5, no. 3, pp. 1135–1159, Sep. 2019.
[47] E. Parselia et al., "Satellite Earth Observation Data in Epidemiological Modeling of Malaria, Dengue and West Nile Virus: A Scoping Review," Remote Sens., vol. 11, no. 16, p. 1862, Aug. 2019.
[48] B. Deepa and K. Ramesh, "Epileptic seizure detection using deep learning through min max scaler normalization," Int. J. Health Sci. (Qassim)., pp. 10981–10996, May 2022.
[49] S. Kumar and S. Agrawal, "Prevention of vector-borne disease by the identification and risk assessment of mosquito vector habitats using GIS and remote sensing: a case study of Gorakhpur, India," Nanotechnol. Environ. Eng., vol. 5, no. 2, p. 19, Aug. 2020.
[50] Z. Nabila, A. R. Isnain, and P. Permata, "Mining Data Analysis for Clustering of Covid-19 Case in Lampung Province Using K-Means Algorithm," 2021.
[51] S. S. Nagari and L. Inayati, "Implementation of Clustering using K-Means Method to Determine Nutritional Status," J. Biometrika dan Kependud., vol. 9, no. 1, p. 62, Jun. 2020.
[52] V. K. Dehariya, S. K. Shrivastava, and R. C. Jain, "Clustering of Image Data Set Using K-Means and Fuzzy K-Means Algorithms," in 2010 International Conference on Computational Intelligence and Communication Networks, Nov. 2010, pp. 386–391.
[53] R. Mohana Priya, L. K. Hema, V. Vanitha, and R. Karthikeyan, "Classification and Detection of Malarial Parasite in Blood Samples Using K-Means Clustering Algorithm and Support Vector Machine Classifier," in Lecture Notes in Networks and Systems, 2021, pp. 423–428.
[54] V. C. Osamor, E. F. Adebiyi, and S. Doumbia, "Clustering Plasmodium falciparum Genes to their Functional Roles Using k-means," vol. 2, no. 2, 2010.
[55] E. V. H. S., M. Y. Mashor, A. S. A. Nasir, and Z. Mohamed, "Identification of Giemsa Staind of Malaria Using K-Means Clustering Segmentation Technique," in 2018 6th International Conference on Cyber and IT Service Management (CITSM), Aug. 2018, pp. 1–4.

[56] E. Maturahmah and S. Prafiadi, “Inventarisasi Tumbuhan Obat Dan Kearifan Lokal Masyarakat Suku Mandacan Dalam Memanfaatkan Tanaman Obat Di Desa Anggi Gida, Kabupaten. Pegunungan Arfak, Provinsi Papua Barat,” Nusant. J. Ilmu Pengetah. Sos., vol. 7, no. 2, pp. 408–420, 2020.
[57] R. Suyono, J. A. R. Salmun, and H. I. Ndoen, “Analisis Spasial Tempat Perindukan Nyamuk, Kepadatan Larva dan Indeks Habitat dengan Kejadian Malaria di Kecamatan Waigete Kabupaten Sikka,” Media Kesehat. Masy., vol. 3, no. 1, pp. 1–11, Apr. 2021.
[58] A. Asnifatima, “Pola Kecenderungan Spasial Kejadian Malaria (Studi Kasus ; Di Kabupaten Kepulauan Selayar Tahun 2011-2013),” HEARTY, vol. 5, no. 1, Feb. 2017.
[59] J. L. Pasau, A. Asmirah, and A. Burchanuddin, “Tindakan Sosial Masyarakat dalam Pencegahan Penyakit Malaria di Kota Jayapura,” J. Sosiol. Kontemporer, vol. 3, no. 1, pp. 31–36, Jun. 2023.
[60] Z. Liu et al., “Combination Strategies of Variables with Various Spatial Resolutions Derived from GF-2 Images for Mapping Forest Stock Volume,” Forests, vol. 14, no. 6, p. 1175, Jun. 2023.
[61] J. L. Servadio, G. Machado, J. Alvarez, F. E. de Ferreira Lima Júnior, R. Vieira Alves, and M. Convertino, “Information differences across spatial resolutions and scales for disease surveillance and analysis: The case of Visceral Leishmaniasis in Brazil,” PLoS One, vol. 15, no. 7, p. e0235920, Jul. 2020.
[62] N. Afira and A. W. Wijayanto, “Mono-temporal and multi-temporal approaches for burnt area detection using Sentinel-2 satellite imagery (a case study of Rokan Hilir Regency, Indonesia),” Ecol. Inform., vol. 69, p. 101677, Jul. 2022.
[63] A. W. Wijayanto, N. Afira, and W. Nurkarim, “Machine Learning Approaches using Satellite Data for Oil Palm Area Detection in Pekanbaru City, Riau,” in 2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), Jun. 2022, pp. 84–89.
[64] S. R. Putri, A. W. Wijayanto, and S. Pramana, “Multi-source satellite imagery and point of interest data for poverty mapping in East Java, Indonesia: Machine learning and deep learning approaches,” Remote Sens. Appl. Soc. Environ., vol. 29, p. 100889, Jan. 2023.
[65] T. Devara and A. W. Wijayanto, “Machine Learning Applied To Sentinel-2 and Landsat-8 Multispectral and Medium-Resolution Satellite Imagery for the Detection of Rice Production Areas in Nganjuk, East Java, Indonesia,” Int. J. Remote Sens. Earth Sci., vol. 18, no. 1, p. 19, Sep. 2021.
[66] Y. Nurmasari and A. W. Wijayanto, “Oil Palm Plantation Detection in Indonesia Using Sentinel-2 and Landsat-8 Optical Satellite Imagery (Case Study: Rokan Hulu Regency, Riau Province),” Int. J. Remote Sens. Earth Sci., vol. 18, no. 1, p. 1, Sep. 2021.
[67] Y. C. Putra and A. W. Wijayanto, “Automatic detection and counting of oil palm trees using remote sensing and object-based deep learning,” Remote Sens. Appl. Soc. Environ., vol. 29, p. 100914, Jan. 2023.
[68] A. W. Wijayanto, D. Wahyu Triscowati, and A. H. Marsuhandi, “Maize field area detection in East Java, Indonesia: An integrated multispectral remote sensing and machine learning approach,” in 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), Oct. 2020, pp. 168–173.
[69] Y. C. Putra, A. W. Wijayanto, and G. A. Chulafak, “Oil palm trees detection and counting on Microsoft Bing Maps Very High Resolution (VHR) satellite imagery and Unmanned Aerial Vehicles (UAV) data using image processing thresholding approach,” Ecol. Inform., vol. 72, p. 101878, Dec. 2022.
[70] N. A. Utami, A. W. Wijayanto, S. Pramana, and E. T. Astuti, “Spatially granular poverty index (SGPI) for urban poverty mapping in Jakarta metropolitan area (JMA): a remote sensing satellite imageries and geospatial big data approach,” Earth Sci. Informatics, vol. 16, no. 4, pp. 3531–3544, Dec. 2023.
[71] A. W. Wijayanto and S. R. Putri, “Estimating Rice Production using Machine Learning Models on Multitemporal Landsat-8 Satellite Images (Case Study: Ngawi Regency, East Java, Indonesia),” in 2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), Jun. 2022, pp. 280–285.