

Article

High-Temporal-Resolution Prediction of Malaria Transmission Risk in South Kivu, Democratic Republic of the Congo, Based on Multi-Criteria Evaluation Using Geospatial Data

Ryunosuke Komura ^{1,2,*} and Masayuki Matsuoka ^{3,†} 

¹ Information Engineering Department, Faculty of Engineering, Mie University, 1577 Kurimamachiya, Tsu 514-8507, Mie, Japan

² FIXER Inc., Seavans South Building, Top Floor 1-2-3 Shibaura, Minato-ku 105-0023, Tokyo, Japan

³ Graduate School of Engineering, Mie University, 1577 Kurimamachiya, Tsu 514-8507, Mie, Japan; matsuoka@info.mie-u.ac.jp

* Correspondence: ryukomura@icloud.com or komura.ryunosuke23@fixer.co.jp

† These authors contributed equally to this work.

Abstract: Malaria is a major public health concern, and accurate mapping of malaria risk is essential to effectively managing the disease. However, current models are unable to predict malaria risk with high temporal and spatial resolution. This study describes a climate-based model that can predict malaria risk in South Kivu, Democratic Republic of the Congo, daily at a resolution of 2 km using meteorological (relative humidity, precipitation, wind speed, and temperature) and elevation data. We used the multi-criteria evaluation technique to develop the model. For the weighting of factors, we employed the analytical hierarchy process and linear regression techniques to compare expert knowledge-driven and mathematical methods. Using climate data from the prior 2 weeks, the model successfully mapped regions with high malaria case numbers, enabling accurate prediction of high-risk regions. This research may contribute to the development of a sustainable malaria risk forecasting system, which has been a longstanding challenge. Overall, this study provides insights into model development to predict malaria risk with high temporal and spatial resolution, supporting malaria control and management efforts.

Keywords: malaria; Democratic Republic of the Congo; multi-criteria evaluation; geographic information system



Citation: Komura, R.; Matsuoka, M. High-Temporal-Resolution Prediction of Malaria Transmission Risk in South Kivu, Democratic Republic of the Congo, Based on Multi-Criteria Evaluation Using Geospatial Data. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 489. <https://doi.org/10.3390/ijgi12120489>

Academic Editors: Christos Chalkias, Demosthenes Panagiotakos and Wolfgang Kainz

Received: 28 August 2023

Revised: 19 November 2023

Accepted: 23 November 2023

Published: 6 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Malaria is one of the longest-known diseases affecting humankind [1], yet the numbers of malaria cases and deaths remain high, with an estimated 1.7 billion cases and 10.6 billion deaths recorded from 2000 to 2020 [2]. Clarifying the global impact of malaria is crucial, as it remains one of the most severe public health problems, with its burden predominantly borne by poorer regions, particularly in Africa. In 2020 alone, an estimated 241 million malaria cases and 627,000 deaths occurred, with substantial economic costs exceeding \$12 billion annually. Despite major progress, including a 36% reduction in malaria mortality from 2010 to 2020, challenges remain, particularly in sub-Saharan Africa, where reductions in incidence and mortality rates have slowed. These statistics underscore the urgent need for improved malaria control and eradication strategies supported by crucial political and scientific initiatives [3,4]. The World Health Organization (WHO) has set the goal of a world free of malaria by 2030, defined as reductions in the mortality rate and incidence of malaria by at least 90% compared to 2015 [2]. To achieve this goal, the development of a high-quality and sustainable malaria risk forecasting system is essential [5].

Numerous researchers have attempted to map malaria risk using either geographic information system (GIS)-based spatial analysis or machine learning techniques. For

example, Ferrao et al. developed a model for GIS-based spatial analysis and produced a map of malaria risk in Mozambique using climatic, sociodemographic, and clinical data [6]. Wieland et al. used machine learning to combine climate data with mosquito samples collected in the field to model mosquito habits [7].

We report the creation of a model using the multi-criteria evaluation (MCE) method for predictions with high temporal and spatial resolution. MCE is a standard method used with GIS for various purposes, including the mapping of malaria risk. Rakotoarison et al. used this method to estimate malaria risk in Madagascar using land-cover classifications based on satellite images, evaluation data, temperature, rainfall, and population density [8]. Adeola et al. used it in combination with the normalized difference vegetation index (NDVI), the normalized difference water index (NDWI), and land surface temperature (LST) data from satellite images of South Africa [9]. Researchers have extensively studied the estimation of malaria risk, considering various factors, including temperature, humidity, hydrogeomorphology, land usage, and precipitation. Notably, Asgarian et al. highlighted the complex influence of meteorological parameters on mosquito abundance, emphasizing a strong positive correlation between temperature and mosquito populations [10]. Minale and Alemu developed a detailed malaria risk map for Bahir Dar, Ethiopia, considering factors like rainfall, temperature, and land use, which are crucial determinants of malaria hazard levels in different areas [11]. Furthermore, Bhatt and Joshi utilized an integrated geospatial and MCE approach to identify malaria risk zones in Vadodara district, demonstrating the utility of geospatial tools in public health planning for malaria [12]. However, little research has focused on the use of meteorological data with the MCE model, despite meteorological factors having strong impacts on mosquitoes [13].

Malaria cases depend on the number of female *Anopheles* mosquitoes. According to the U.S. Centers for Disease Control and Prevention, these insects fly no further than 2 km and generally live for 2–4 weeks (1–2 weeks of the egg, larva, and pupa stages and 1–2 weeks of the adult stage [14,15]). Therefore, a model that can predict those time intervals is needed. Kim et al. developed a model to forecast these stages from 1 to 16 weeks in advance, while Panzi et al. forecast malaria morbidity up to the year 2036, and attempts have been made to develop a model that can make such predictions for a short time [5,16]. These studies have achieved good accuracy but the spatial resolution of the predictions remains quite low.

In this study, we aim to address the research question: How can we improve the temporal and spatial resolution of malaria risk prediction models using meteorological data? The motivation behind these efforts is not only to enhance malaria mapping but also to enable more targeted and effective malaria control interventions, optimize resource allocation, and ultimately contribute to the global goal of malaria eradication. By increasing the temporal and spatial resolution of the predictions, we can better characterize the dynamics of mosquito populations and their habitats, leading to more precise identification of high-risk areas and facilitating timely and targeted interventions to reduce the burden of malaria.

Therefore, the objective of this research is to develop a sustainable prediction model based on mosquito habitat and life span information obtained from meteorological data that have higher spatial resolution and shorter time intervals than existing methods.

2. Materials

2.1. Study Area

This study focused on a region within the Democratic Republic of the Congo (DRC), which has experienced high malaria incidence. According to the 2021 World Malaria Report by the WHO, around 12% of malaria cases worldwide occur in the DRC, the second-highest contribution of any nation [2]. A 2018 survey conducted in the DRC revealed that South Kivu had the highest number of individuals undergoing malaria testing [17]. The DRC is divided into 26 provinces, each containing multiple health zones, with a total of 519 health zones throughout the country. The target province of this study, South Kivu, encompasses 34 health zones. Figure 1 shows the locations of the DRC and South Kivu.

Geographically, South Kivu is located in the eastern region of the DRC, bordering Rwanda, Burundi, and Tanzania. It has internal borders with the provinces of North Kivu, Maniema, and Katanga. The provincial capital is Bukavu, and other major cities are Baraka and Uvira. The topography of South Kivu consists of mountains in the eastern and northern parts, the Central Basin in the western part, and a vast plain in the southeastern part. Vegetation cover in the region includes upland forests, savannah grasslands, bamboo woods, and dense forests [18].

In terms of climate conditions, monthly average values for the past 20 years indicate that November has the largest number of rainy days at 18, while July has the fewest. The study region experiences a rainy season from October through the end of May. This is interrupted by a short drier period in January and February. Average monthly relative humidity over the past two decades was highest in April at 83% and lowest in August at 63%, with generally stable values throughout the year [19,20].

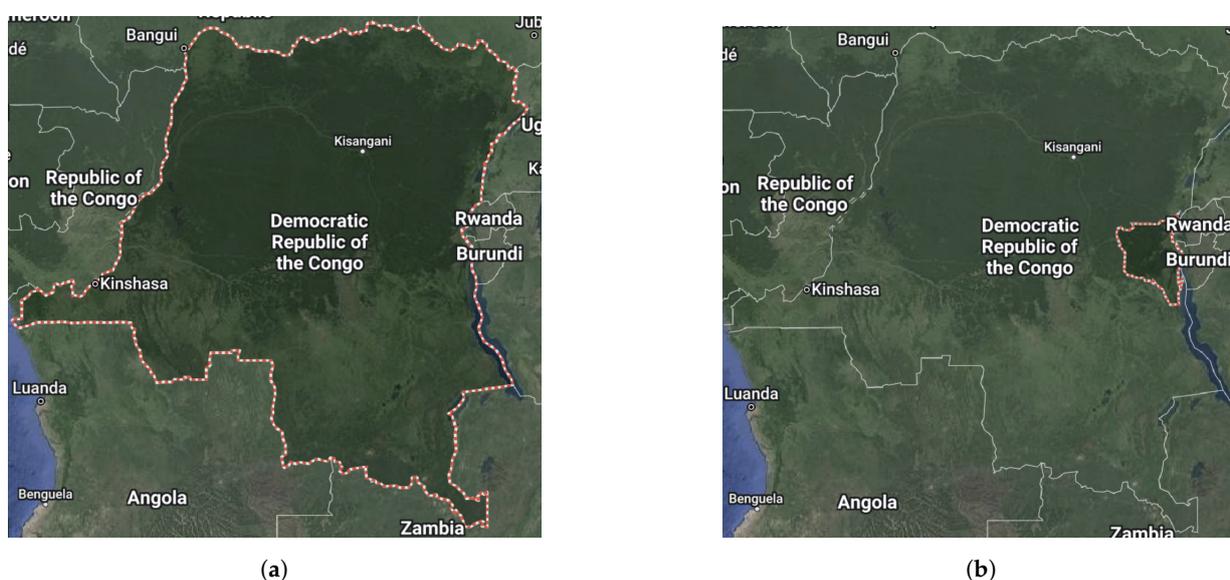


Figure 1. Map of Democratic Republic of the Congo (a), and South Kivu (b). Reprinted with permission from Ref. [21]. Copyright 2023 Google.

2.2. Datasets

Malaria incidence and population density data were used to calculate the number of malaria cases at the selected resolution. To calculate the risk of becoming infected with malaria, elevation and meteorological data were used. Table 1 shows the data used in this study, including the names of datasets and data providers.

Table 1. Descriptions of the datasets used.

Data Type	Name of Dataset	Data Provider	Resolution
Malaria cases		Directorate of Epidemiological Surveillance (DES)	Per health zone
Elevation	Shuttle Radar Topography Mission (SRTM)	the National Aeronautics and Space Administration (NASA)	30 km × 30 km
Meteorological	ECMWF Reanalysis v5 (ERA5) hourly data on single levels from 1959 to the present	The European Center for Medium-Range Weather Forecasts (ECMWF)	1-arc-second × 1-arc-second (approximately 30 m × 30 m)
Population density	High-Resolution Population Density Maps + Demographic Estimates	Meta and Center for International Earth Science Information Network (CIESIN) at Columbia Climate School of Columbia University	30 m × 30 m

Malaria datasets were provided by the Directorate of Epidemiological Surveillance (DES). Before data acquisition, our research protocol was thoroughly reviewed and approved by the DES. The datasets include the number of cases within each health zone. The data did not include any information that can identify individuals. The datasets for South Kivu (total across 34 health zones) in the time period of 1 January 2018 to 31 December 2021 were used.

For elevation data, Shuttle Radar Topography Mission (SRTM) data from the National Aeronautics and Space Administration (NASA) were used. Elevation data were obtained from raw C-band radar signals spaced at intervals of 1 arc-second (approximately 30 m) through processing by NASA's Jet Propulsion Laboratory. The data were edited by the National Geospatial Intelligence Agency to delineate and flatten water bodies, improve definition, remove spikes and wells, and fill small voids. These data were collected on 1 February 2005 and are accessible to the public online [22].

Meteorological information was obtained from the European Center for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) hourly data from 1959 to the present. Further information is provided in Table 2. This meteorological dataset is openly accessible online [23].

For population density, "Democratic Republic of the Congo: High-Resolution Population Density Maps + Demographic Estimates" collected by Meta and the Center for International Earth Science Information Network (CIESIN) at Columbia Climate School of Columbia University were used. This dataset is available online and can be freely accessed, showcasing the number of people living within 30 m grid tiles calculated from a combination of satellite imagery (Sentinel-2A) and census information. These data were collected on 20 May 2019 [24].

Table 2. ERA5 meteorological variables and their definitions.

Variables	Unit	Definition
10 m u-component of wind	m/s	The eastward component of the 10 m wind. It is the horizontal speed of air moving towards the east, at the height of ten meters above the surface of the Earth, in meters per second.
10 m v-component of wind	m/s	The northward component of the 10 m wind. It is the horizontal speed of air moving towards the north, at the height of ten meters above the surface of the Earth, in meters per second
2 m dewpoint temperature	K	Temperature to which the air at 2 m above the surface of the Earth would need to be cooled for saturation to occur; a measure of the humidity of the air
2 m temperature	K	Temperature of air at 2 m above the surface of land, sea or inland waters. Calculated by interpolating between the lowest model level and the Earth's surface, taking account of the atmospheric conditions.
Total precipitation	m	Accumulated liquid and frozen water that falls to the Earth's surface, including rain and snow

3. Methods

3.1. Research Outline

The research methodology of this study is illustrated in Figure 2. Following data collection, preprocessing steps were undertaken, including conversion of file types if required, resampling, and splitting into grids of the selected size. In this study, the term malaria risk refers to an index ranging from 0 to 1, with higher values indicating a greater risk of malaria occurrence. We developed this index to quantify and analyze the malaria risk within each grid. The number of malaria cases within each grid was computed using the population distribution data and total malaria cases in each health zone.

The MCE technique employed in this study includes the analytical hierarchy process (AHP) and multiple linear regression (LR method). AHP is a structured technique for organizing and analyzing complex decisions based on mathematics and psychology. This method was developed by Saaty [25] and is particularly useful for group decision-making.

AHP can capture both subjective and objective aspects of a decision, and allows the intuitive judgments of the decision-maker to be quantified, ensuring consistency in comparisons among factors [26].

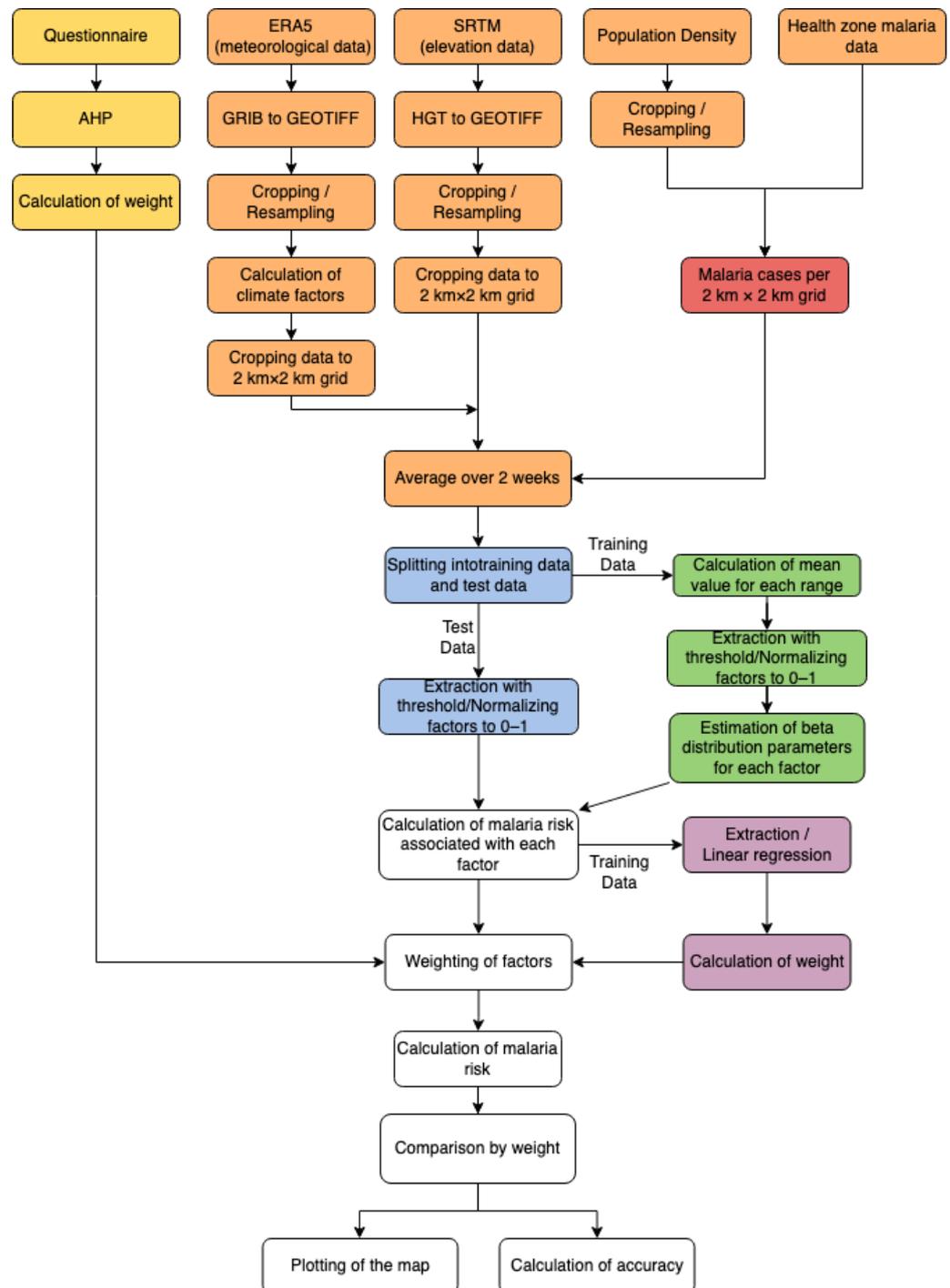


Figure 2. Research workflow: This figure illustrates the sequential steps of the research process. The orange segment represents the preprocessing phase, encompassing data preparation. The red segmented phase involves the computation of malaria case numbers. The blue segment represents that data from the period of 1 January 2018, to 31 December 2020, was used as the training period, while the period from 1 January 2021, to 31 December 2021, was selected as the validation period. Subsequently, the green segment calculates malaria risks. The purple phase denotes the application of the LR method, while the yellow phase corresponds to the AHP method, responsible for factor weighting. Finally, the white phase involves a comparative analysis of the two methods.

Meanwhile, the LR method is a statistical technique that allows a dependent variable to be predicted based on the values of multiple independent variables. This method extends the simple linear regression model to include multiple predictors and its general form is represented by Equation (1) [27], as follows:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + e_i \quad (1)$$

In our study, the MCE technique was used to determine the malaria risk in each grid by summing the weighted variables that indicate the risk for malaria associated with each factor. The risk for each factor was calculated using the probability density of each factor, which was estimated using the beta distribution. The weight of each factor was determined using either the AHP or LR method for comparison of the results of manual and mathematical methods. The accuracy of the model was evaluated through a comparison of the training and testing periods as well as a comparison of models that used the AHP technique versus the LR technique.

3.2. Preprocessing Data

The complete preprocessing workflow for the data described in this section is illustrated in the orange section of Figure 2. Distributed ERA5 data (meteorological data) were in the form of a GRIB file and SRTM data (elevation data) were in an HGT file. Both the GRIB file and the HGT file were translated into GeoTIFF files. CIESIN population density data were distributed as a GeoTIFF file, and this file was used without translation. After translation of the necessary files, resampling was conducted as outlined in Table 3; this process was selected to simplify the subsequent splitting of the data into 2 km × 2 km grids. Resampling was conducted using a bilinear method with the translator library GDAL [28]. Meteorological data were collected hourly, and daily data for each pixel were calculated as the average hourly values over each day. After resampling, the data were split into 2 km × 2 km grids.

Table 3. Resampling data.

Data Type	Before Resampling	First Resampling	Final Resampling
ERA5	30 km × 30 km	100 m × 100 m	2 km × 2 km
SRTM	1-arc-second × 1-arc-second (approximately 30 m × 30 m)	20 m × 20 m	2 km × 2 km
Population	30 m × 30 m	20 m × 20 m	2 km × 2 km

The meteorological and elevation data were processed by averaging the values of all pixels within each 2 km × 2 km grid. Humidity data were calculated using the August–Roche–Magnus formula (Equation (2)), which incorporates the 2 m temperature (T), 2 m dewpoint temperature (TD), coefficient C_a (equal to 17.625), coefficient C_b (equal to 243.05), and the value of K_{zero} (equal to 273.15 K, which is equivalent to 0 °C) [29]. This formula was used to calculate relative humidity (RH).

$$RH = 100 \times \frac{\exp\left(\frac{C_a \times (TD - K_{zero})}{C_b + (TD - K_{zero})}\right)}{\exp\left(\frac{C_a \times (T - K_{zero})}{C_b + (T - K_{zero})}\right)} \quad (2)$$

Wind speed was calculated from the 10 m u and v_s components of wind using Equation (3), where WS is wind speed (m/s), U is the 10 m u component of wind (m/s), and V is the 10 m v_s component of wind (m/s). Then, 2-week average factor values were calculated for each day. To obtain population density data, the total population within each 2 km × 2 km grid was calculated.

$$WS = \sqrt{U^2 + V^2} \quad (3)$$

3.3. Calculation of Malaria Data per 2 km × 2 km Grid

The workflow described here is depicted in red in Figure 2. The malaria case number for each 2 km × 2 km grid was calculated using the total number of malaria cases in the health zone where the grid is located. Population percentage was calculated using the total population within the 2 km × 2 km grid and the total health zone population, which was calculated by summing the population density data of each health zone (Equation (4)), where a_k is the malaria case number in grid k , M is health zone M , H_M is the total malaria case number in health zone M , P_k is the total population of location k , and P_M is the total population of health zone M

$$a_k = H_M \times \frac{P_k}{P_M} \quad (4)$$

3.4. Outline of the Model

We developed a model, shown in Equation (5), using the MCE method [30], where S is the risk for malaria with a range of 0 to 1 for the 2 km × 2 km grid at location k , C_i is constraint i , n is the number of constraints, ω_j is the weight of factor j , f_j is a function that indicates the risk for malaria for factor j , x_{jk} is the normalized factor j value in location k , and m is the number of factors.

$$S_k = \prod_{i=1}^n C_i \left(\sum_{j=1}^m \omega_j f_j(x_{jk}, \alpha_j, \beta_j) \right) \quad (5)$$

3.5. Estimation of the Parameters in Function $f_j(x_j, \alpha_j, \beta_j)$

3.5.1. Calculation of the Probability Distribution

The green section in Figure 2 depicts the process detailed in this section. The f function is derived using Equation (6), wherein b_j is the probability density function of factor j (Equation (6)), and α_j and β_j are beta distribution parameters of factor j . To derive the parameters α and β , the open-source Python library SciPy maximum likelihood estimation (MLE) algorithm (version 1.9.3) was used with data for the training period (1 January 2018–31 December 2020).

$$f_j(x_j, \alpha_j, \beta_j) = b_j(x_j, \alpha_j, \beta_j) \times \frac{1}{\max_{0 \leq x \leq 1} (b_j(x, \alpha_j, \beta_j))} \quad (6)$$

$$b_j(x_j, \alpha_j, \beta_j) = \frac{\Gamma(\alpha + \beta) x_j^{\alpha-1} (1 - x_j)^{\beta-1}}{\Gamma(\alpha) \Gamma(\beta)} \quad (7)$$

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

After each beta distribution parameter was estimated, the maximum of each beta distribution within the range 0–1 was calculated and the beta distribution was normalized to 0–1, creating the f function (Equation (6)). This expresses the malaria risk associated with each factor over the range 0–1.

3.5.2. Calculation of the Mean Value of Each Range

To determine the parameter values (α and β) of the f function, we calculated the average number of malaria cases per 2 km × 2 km grid. Table 4 presents selected parameter values for each factor. For example, to analyze the impact of elevation, the mean number of malaria cases was computed for each interval of 10^{-1} m. The workflow described here is indicated by the green portions of Figure 2.

Table 4. Parameter.

Factor's Name	Range of Step Size
Elevation	10^{-1} m
Humidity	$10^{-5}\%$
Precipitation	10^{-6} m
Temperature	10^{-5} K
Wind speed	10^{-7} m/s

3.5.3. Normalizing Factors and Deriving Constraints

Each factor was adjusted using a threshold and then normalized to the range 0–1. The threshold value for each factor is listed in Table 5. The thresholds were selected based on locations where mosquitoes can survive and the relationships between malaria case numbers and the factors described in Section 4.1 [31]. Data falling outside of the threshold line were normalized to 0 and the constraint (C_j in Equation (5)) was set to 0; meanwhile, for values within the range, the constraint was set to 1. Using this normalized factor and the mean number of malaria cases, we estimated the f function parameter values using the MLE algorithm. The workflow described in this section is shown in green in Figure 2.

Table 5. Thresholds.

Data Name	Minimum Threshold	Maximum Threshold
Elevation (m)	0	3000
Humidity (%)	60	100
Precipitation (m)	0	0.003
Temperature (K)	293.15 (20 °C)	313.15 (40 °C)
Wind speed (m/s)	0	3

3.6. Calculation of Weights

The weight (ω) of each factor was calculated using two methods, one based on the answers to a questionnaire from eight specialists on malaria or mosquitoes working in the United States and Japan (AHP method), and the other based on the LR method. These two methods were used to compare manual and mathematical derivation methods.

3.6.1. AHP Method

For the method based on specialist knowledge, the questionnaire was collected using Saaty's Continuous Rating Scale [25]. All experts were asked to compare the importance of two factors to the occurrence of mosquitoes and the resulting scores for each factor are presented in Table 6 with respect to all combinations of paired factors. In the MCE process, the primary criterion was the professional expertise of the respondents. All participants in the study were required to be affiliated with universities or research institutes, ensuring that their responses were grounded in specialized knowledge and experience in the field. The data collected in the questionnaire were input into the AHP formula to obtain weights. Then, those weights were normalized to the range 0–1 such that the sum of the weights equals 1, thereby enabling comparison with the LR method. The workflow steps noted in this section are marked in yellow in Figure 2. The answers received from each expert are shown as the Table A1.

Table 6. Saaty’s Continuous Rating Scale.

Intensity of Importance	Definition
9 (More important)	Extremely strong
7	Very strong
5	Strong
3	Moderately strong
1	Equally important
1/3	Moderately weak
1/5	Weak
1/7	Very weak
1/9 (Less important)	Extremely weak

3.6.2. LR Method

To determine factor weights for the model, we extracted meteorological and elevation data associated with the top 1–20% of malaria cases. Then, we utilized the LR method, applying Equation (8), where a_k represents the number of malaria cases. Subsequently, we derived the weight (ω) by normalizing the explanatory variables (Ω) to the range 0–1 using Equation (9). The top 1–20% was selected to prioritize cases with higher percentages, emphasizing areas where malaria poses a significant risk over areas of lower risk. The workflow outlined in this section is visualized in the purple portion of Figure 2.

$$a_k = \sum_{j=1}^m \Omega_j f(x_j, \alpha_j, \beta_j) \quad (8)$$

$$\omega_j = \frac{\Omega_j}{\sum_{i=1}^m \Omega_i} \quad (9)$$

3.7. Accuracy Evaluation

We evaluated the accuracy of our model from three perspectives. First, we compared the selected ranges of both models. Second, we evaluated model performance by selecting the top 1, 5, 10, and 20 percentiles of malaria cases for each risk range and comparing results between the training and validation periods. Finally, we compared the plotted risk and malaria case numbers on two dates, representing the rainy and dry seasons.

To compare the models using AHP and LR, we calculated the top 1–5% of malaria infection numbers for every 0.1 increment of risk and plotted those values on the grid. This process allowed us to determine whether an overall increase in malaria cases occurs with increasing risk.

For the percentile comparisons, a linear function was derived from the data points corresponding to each percentile. Finally, for the seasonal comparison, we compared the plotted risk and malaria cases across all 34 health zones on two dates, 1 April 2021 and 21 June 2021, representing the rainy and dry seasons, respectively. We selected these dates because they align with distinct seasons and the DES provided malaria case counts for all 34 health zones on these dates.

4. Result and Discussion

4.1. Relationship between Malaria Case Numbers and Environmental Factors

Figure 3 shows the plotted relationships between 2 and week average values of each factor and malaria cases for 2018 to 2020. At some points, the incidence of malaria tends to be high for each factor. Overall, cases tend to decrease with high wind speed, high temperatures, low humidity, heavy precipitation, and at high elevations. These results clearly indicate that malaria cases depend on the habitat of mosquitoes, as reported previously [31]. While our research assumes that these environmental factors are optimal predictors for our model, further investigation is warranted to confirm this assumption and explore whether other factors might yield more accurate predictions.

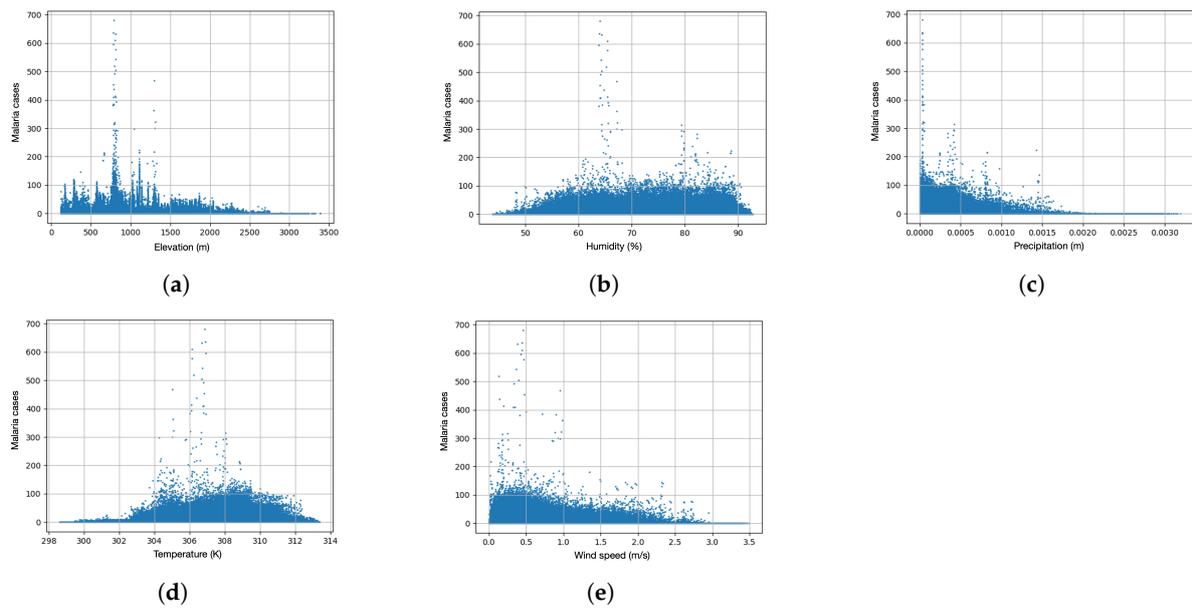


Figure 3. Relationships of 2-week average values of each factor with malaria risk over the period of 1 January 2018–31 December 2020. (a) Elevation (b) Humidity (c) Precipitation (d) Temperature and (e) Wind speed.

4.2. Estimation of f Function Parameters

The values of each parameter and corresponding maximum values are presented in Table 7. Figure 4 shows the relationship between normalized factor values, malaria cases, and the f function. The vertex of the f function is situated near the peak of malaria cases. In addition, the data demonstrate that the f function is a well-fitted regression curve of malaria case numbers.

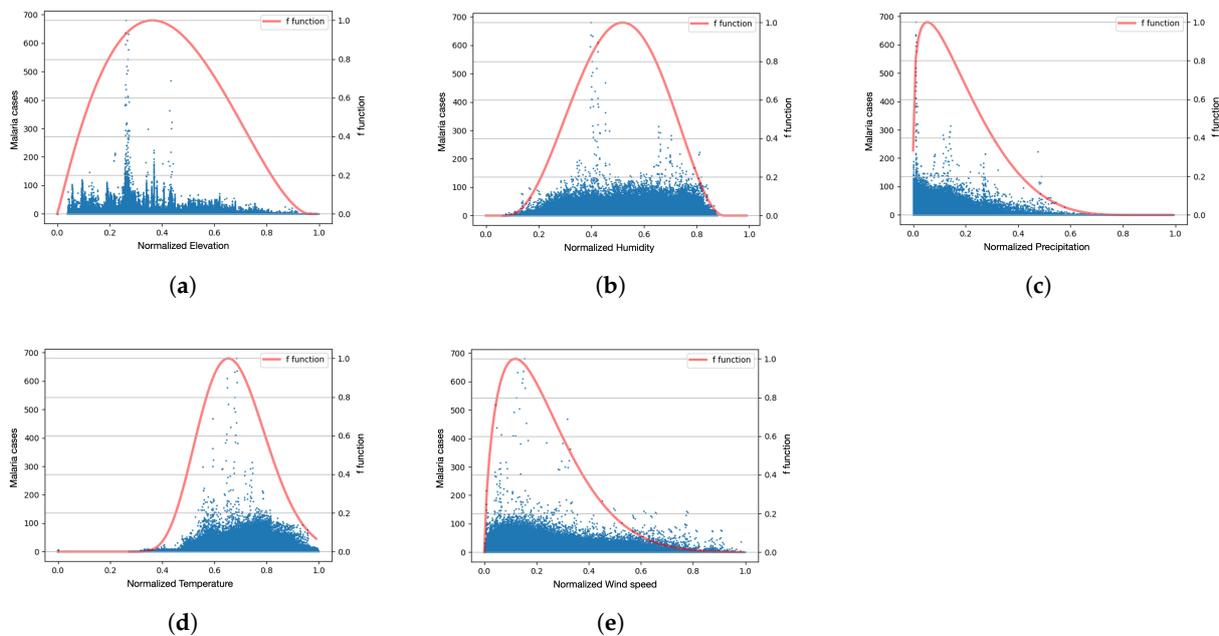


Figure 4. Relationship of malaria case number with the normalized factor value for the training period (left axis) and representation with the f function (right axis). (a) Normalized elevation (b) Normalized humidity (c) Normalized precipitation (d) Normalized temperature (e) Normalized wind speed.

Table 7. Parameter values of the f function.

Factors	α	β	Loc	Scale	$\max_{0 \leq x \leq 1}(f(x, \alpha, \beta))$
Elevation	2.046102	2.729151	0.000768	0.961757	1.763140
Humidity	3.374001	3.055518	0.074224	0.830404	2.348990
Precipitation	1.237038	4.485245	-0.00020	0.847968	3.545051
Temperature	7.197822	16.91633	0.265858	1.381764	3.112984
Wind speed	1.715887	8.811624	0.000163	1.409230	2.978210

4.3. Calculated Weights

Table 8 shows the weights calculated using the AHP and LR methods. We found a significant difference in the weights between the two methods. For the AHP method, elevation was the weakest factor and humidity was the second weakest; for the LR method, humidity was the greatest factor, followed by elevation. We mapped the population distribution and elevation of the study region (Figure 5). The population is concentrated in locations at lower elevations within the region. As malaria transmission occurs when humans are present in an area, elevation is one of the most important factors, and this importance is reflected in its weighting. In South Kivu, temperature remains relatively constant throughout the year. Therefore, temperature had the third weakest influence on LR. For precipitation, the number of rainy days per month can range from 1 to 18 throughout the year with dramatic shifts, and thus its light weight was surprising [19].

Table 8. Weight calculated using AHP and LR methods.

Data Name	AHP	LR
Elevation	0.06976	0.34538
Humidity	0.12989	0.40588
Precipitation	0.28761	0.09751
Temperature	0.30844	0.16747
Wind speed	0.20429	-0.01625

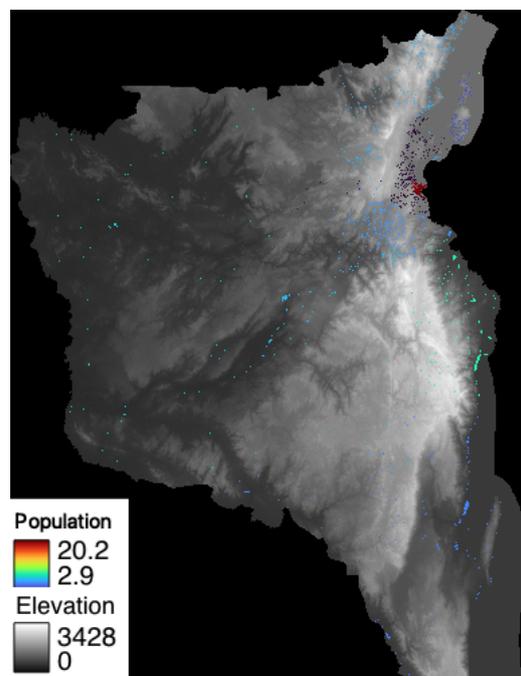


Figure 5. Map of population (colored dots indicate locations with more than 0 people within a 20 m × 20 m grid) and elevation.

4.4. Accuracy Evaluation

4.4.1. Comparison of Selected Risk Ranges

Figure 6 shows the relationships between estimated malaria risk and observed cases obtained using each method. LR shows a stronger trend of association, although the difference is small.

Figure 7 shows the results of plotting the top 1–5% of malaria case counts per 0.1 increment of risk. Both methods indicate increased risk with an increase in cases. However, for the AHP method, a high number of cases is associated with a low risk, indicating an underestimation of risk. In contrast, the LR method tends to overestimate risk at low case numbers. The observed discrepancy may be attributable to two potential causes: first, in the LR method, the top 1–20% of malaria cases were used to weight each factor, resulting in more densely populated areas having a greater impact than less populated areas with high risk due to meteorological and elevation factors; second, the model does not account for the number of individuals within a selected area when calculating risk. These limitations may have contributed to the model's improved performance at high case levels, but not at low case levels.

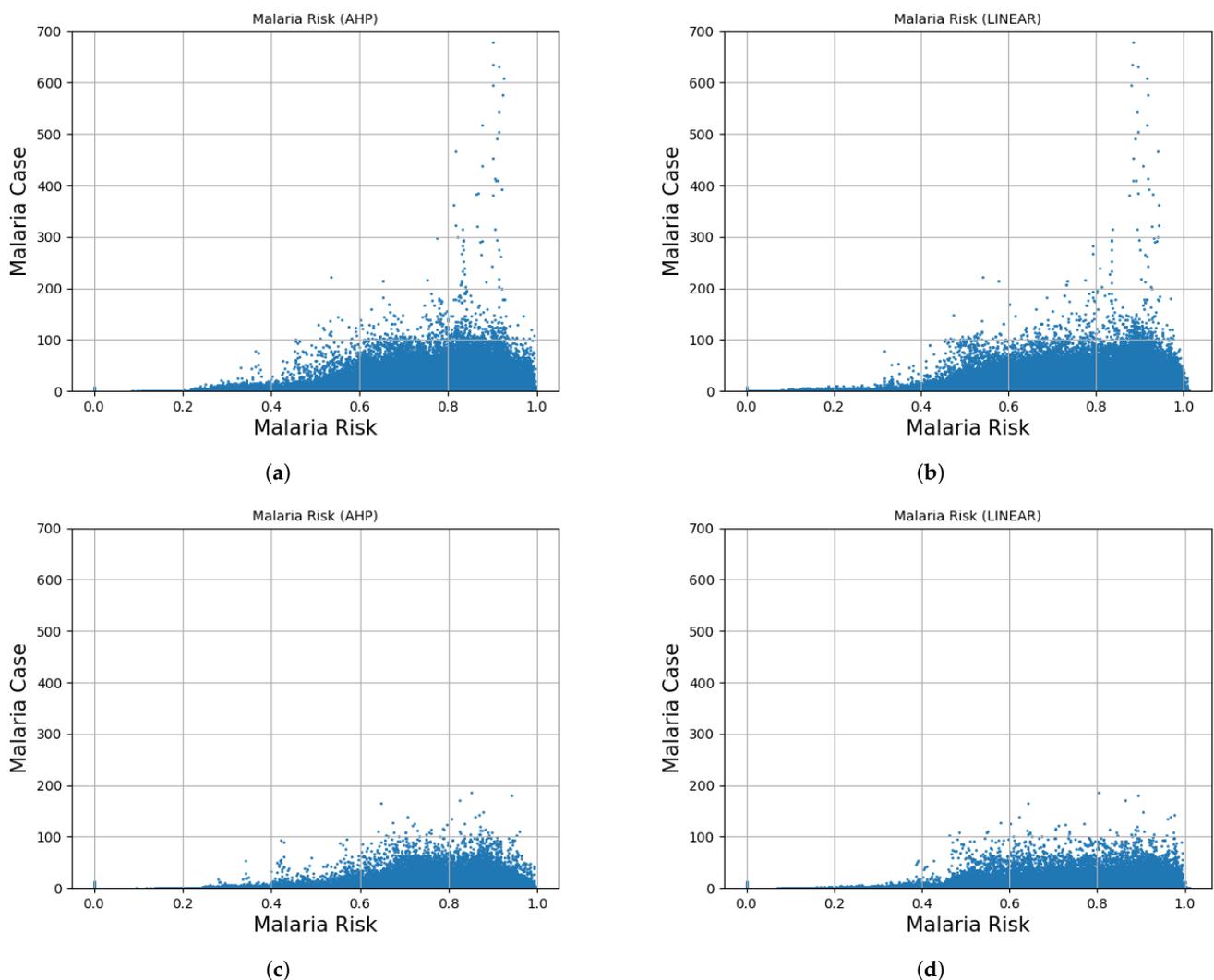


Figure 6. Relationship of malaria risk (S) with malaria cases for AHP method for the training period (a), LR method for the training period (b), AHP method for the validation period (c), and LR method for the validation period (d). The training period was from 1 January 2018 to 31 December 2020, and the validation period was from 1 January 2021 to 31 December 2021.

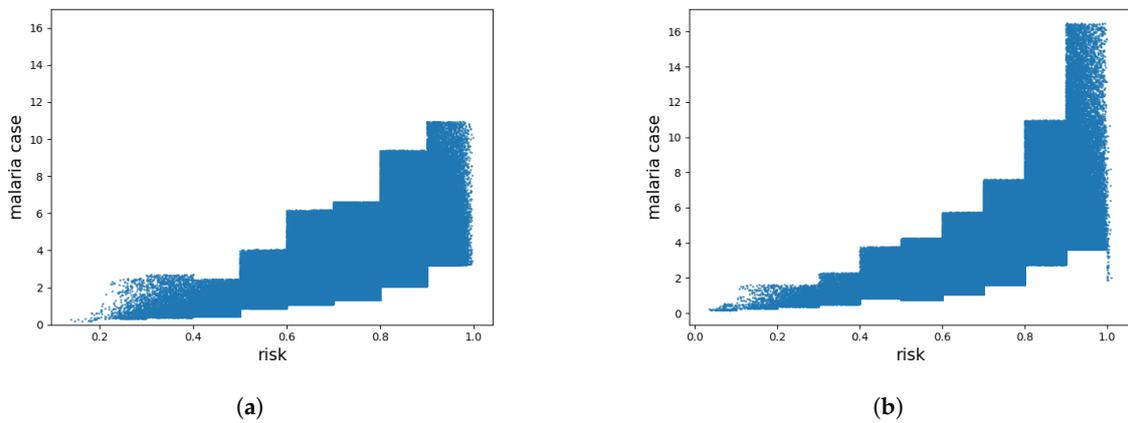


Figure 7. Relation of every 0.1 increment of malaria risk (S) and malaria cases of top 1–5% from 1 January 2018–31 December 2020, for AHP (a) and LR (b).

4.4.2. Comparison of Selected Case Percentiles

Figure 8 presents plots of the top 1%, 5%, 10%, and 20% of malaria cases, averaged over each 0.1 increment of malaria risk for each method. We analyzed this to assess whether the model exhibits similar risk patterns in terms of case number between the validation and training datasets. Both models demonstrate high accuracy, as their respective linear functions show similarities between periods.

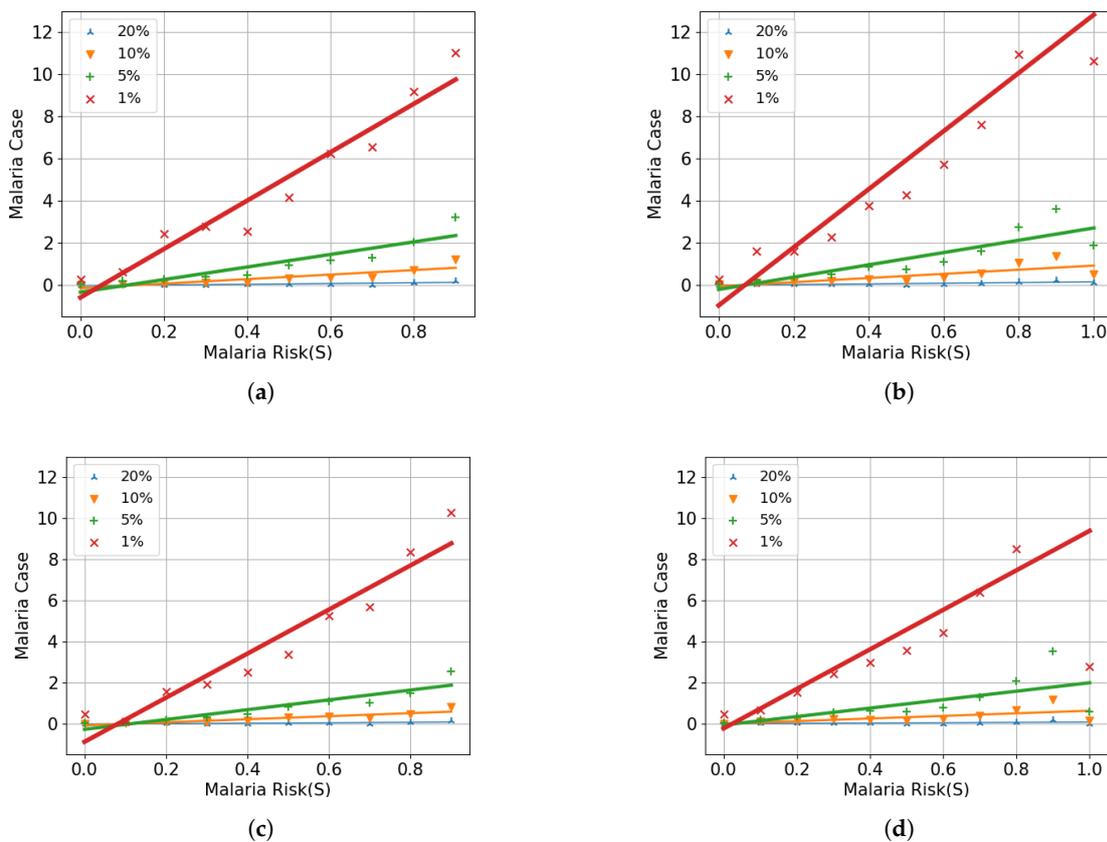


Figure 8. Relationships of the top 1%, 5%, 10%, and 20% of malaria cases with malaria risk, shown for the training period (1 January 2018–31 December 2020) in the top row and validation period (1 January 2021–31 December 2021) in the bottom row with the AHP method in the left column and LR method in the right column. (a) AHP method for the training period (2018–2020), (b) LR method for the training period (2018–2020) (c) AHP method for the validation period (2021) (d) LR method for the validation period (2021).

4.4.3. Comparing Plotted Risk with Case Numbers

Spatial data were visualized using heat maps, with risk levels indicated by varying color intensity, enabling the distinction of high-risk versus low-risk areas, as shown in Figure 9, which shows the calculated risk and the number of malaria cases. Figure 10 provides an enlarged map focusing on malaria risk in the range of 0.5–1 and case numbers in densely populated areas. This figure was prepared to facilitate a visual comparison between risk and malaria cases. The AHP method consistently indicates lower risk levels, whereas the LR method tends to show higher risks. Furthermore, the former exhibits a more gradual change in risk over time compared to the latter. This may be attributable to the fact that the LR weighting method places greater emphasis on elevation. Notably, the LR method produces a more pronounced contrast. However, in the eastern section, the AHP method demonstrates superior results in capturing the seasonal changes in cases. Conversely, the LR method performs well on the western side but tends to overestimate risk on the eastern side. This discrepancy is likely due to the AHP method assigning greater importance to precipitation than the LR method.

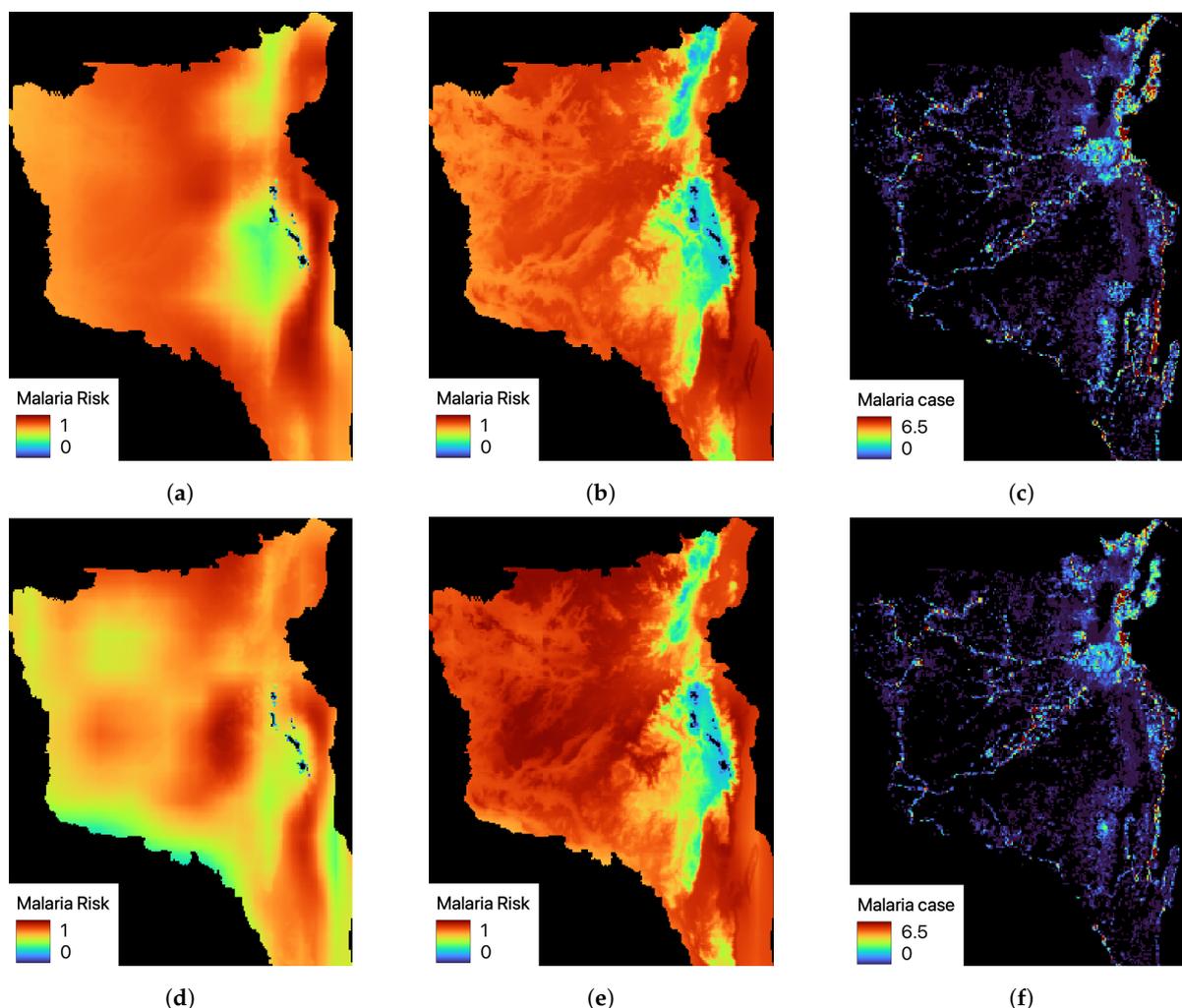


Figure 9. Comparison of malaria risk map and malaria case map. Observation dates are 1 April 2021 (rainy season) for the top panels and 21 June 2021 (dry season) for the bottom panels. To improve readability of the malaria maps, the maximum point is set to the 98th percentile of the 21 June 2021 (dry season) data. (a) Risk map obtained using AHP on 1 April 2021 (rainy season), (b) Risk map obtained using LR on 1 April 2021 (rainy season), (c) Malaria case map on 1 April 2021 (rainy season), (d) Risk map obtained using AHP on 21 June 2021 (dry season), (e) Risk map obtained using LR on 21 June 2021 (dry season), and (f) Malaria case map on 21 June 2021 (dry season).

Both methods identified a region of significantly elevated risk in the eastern part of South Kivu while the western side of the high-elevation area in central South Kivu exhibits low risk. The high-risk area coincides with a densely populated region, as depicted in Figure 5. Conversely, the area of low risk can be attributed to elevated terrain; as one of the highest regions in South Kivu, this area is estimated to have smaller populations of both mosquitoes and humans. Conversely, the eastern part of the region shows high risk, which may be attributable to the concentrated human population in that area.

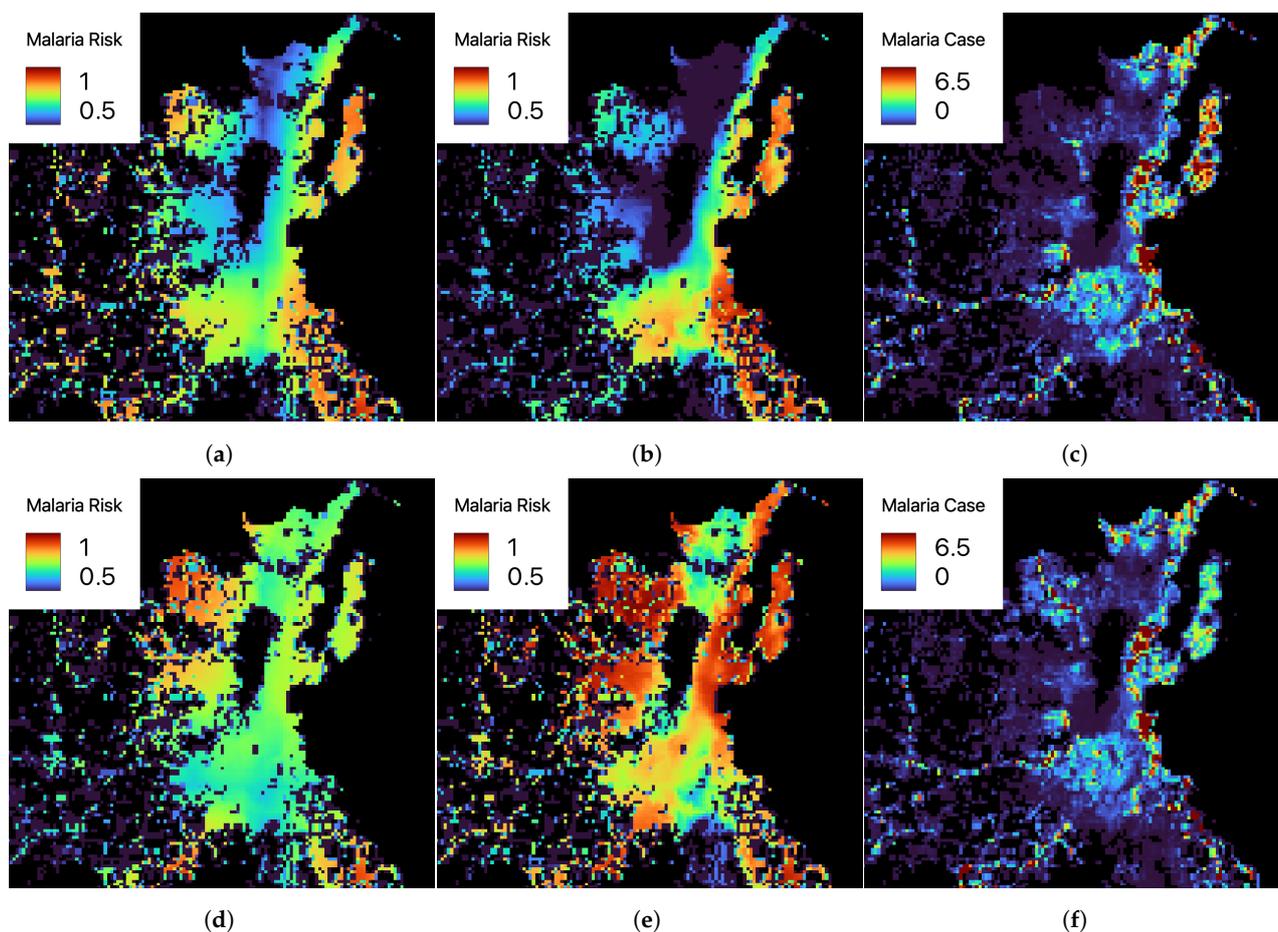


Figure 10. Zoomed map of malaria risk predicted and malaria cases observed at populated area on 1 April 2021 (rainy season) and 21 June 2021 (dry season). For the malaria map due to its visibility, the maximum point is set to the 98th percentile of the 21 June 2021 (dry season) data. (a) Risk map obtained using AHP on 1 April 2021 (rainy season), (b) Risk map obtained using LR on 1 April 2021 (rainy season) (c) Malaria case map on 1 April 2021 (rainy season) (d) Risk map obtained using AHP on 21 June 2021 (dry season) (e) Risk map obtained using LR on 21 June 2021 (dry season) (f) Malaria case map on 21 June 2021 (dry season).

4.4.4. Discussion of the Model

Our model focuses on forecasting the malaria risk using the top 1–5% of malaria cases to calculate factor weights, which can lead to overestimation of risk, even in areas with relatively low numbers of cases. In addition, the model parameters are adjusted only for malaria transmission risk in South Kivu, DRC. Therefore, to apply this method to other locations, further studies are needed to adjust the parameters. In other words, our model can be adapted to other locations if the parameters are adjusted. In summary, the AHP method showed a tendency to underestimate risk in areas with high case counts, whereas the LR method was more effective in areas heavily influenced by population density, which in turn was influenced by elevation, but tended to overestimate risk in areas with low case numbers.

Overall, the models weighted using both the AHP and LR methods produced favorable results, which indicate that areas with high malaria case counts also have elevated risk, whereas areas with low case numbers have lower risk. Our results indicate that the model produces predictions with high temporal and spatial resolution. This technique, including the AHP and LR methods, provides a structured and systematic approach to analyzing complex decisions, combining both subjective and objective elements. AHP captures intuitive judgments, ensuring consistency in comparisons among factors, while the LR method allows for the prediction of a dependent variable based on multiple independent variables. However, this method also has drawbacks. Namely, the AHP method involves subjective judgments, which, despite being quantified and standardized, might introduce biases or inconsistencies depending on the expertise and perspectives of the decision makers involved. Additionally, the LR method is generally not suitable for extrapolation beyond the range of observed data, as the relationships between variables may differ outside of this range.

5. Conclusions

Despite continuing demands and efforts to develop sustainable systems for predicting malaria risk, such systems have remained elusive [5]. To address this challenge, we developed a short-term prediction model for sustainable malaria risk forecasting. Malaria transmission is directly affected by mosquito lifespan and habitat, which are in turn influenced by climate and other factors [32,33]. Therefore, our model uses climate and evaluation data with a resolution of 2 km and considers the behavior of mosquitoes. Future research will address the model's sensitivity to various timeframes of climate data, such as data from a month or less prior to the time of analysis, as the impact of such data on prediction accuracy remains unknown. This study represents an important step toward the development of a sustainable malaria risk forecasting system, which has been a long-standing challenge. By providing a robust tool for predicting malaria outbreaks, this research can aid governmental and public health agencies in devising preemptive strategies, optimizing resource allocation, and enhancing community engagement and education efforts to mitigate the impact of the disease. More benefit will be obtained from targeted and effective disease prevention measures, and planners can incorporate the model's findings into broad health and urban development policies. For example, the predictive power of our model enables health agencies to identify areas at elevated risk of malaria outbreaks with unprecedented precision. This capability allows for more strategic deployment of limited resources, such as mosquito nets, insecticides, and antimalarial drugs. This model can support targeted campaigns for mosquito control and community health education, especially during high-risk periods identified by the model. In conclusion, the integration of our model into the public health policy framework can serve as a catalyst for more dynamic and evidence-based strategies for malaria prevention and treatment, ultimately decreasing malaria incidence and improving public health outcomes.

Overall, our findings provide insights into the development of a model for predicting malaria risk with high temporal and spatial resolution, thereby supporting malaria control and management efforts. Although this study focused on South Kivu, DRC, the model is designed to be versatile and not inherently region-specific; therefore, further studies conducted on a region-by-region basis are essential to fully ascertaining the applicability of the model to distinct climate patterns and environmental factors that influence malaria transmission, thereby supporting adaptation of the model to other regions. Moreover, as the LR method showed promising results but has not been widely applied, exploring the development of novel, user-friendly methodologies remains an important avenue of future research. Collaborative efforts between modelers, public health officials, and local governments can enhance the effectiveness of malaria interventions across diverse regions.

Author Contributions: Conceptualization, Ryunosuke Komura and Masayuki Matsuoka; methodology, Ryunosuke Komura and Masayuki Matsuoka; software, Ryunosuke Komura; validation, Ryunosuke Komura; formal analysis, Ryunosuke Komura; investigation, Ryunosuke Komura; resources, Ryunosuke Komura; data curation, Ryunosuke Komura; writing—original draft preparation, Ryunosuke Komura; writing—review and editing, Ryunosuke Komura; visualization, Ryunosuke Komura; supervision, Masayuki Matsuoka; project administration, Ryunosuke Komura; funding acquisition, Masayuki Matsuoka. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: For meteorological data (ERA5 data) and evaluation data (SRTM data), publicly available datasets were analyzed in this study. ERA5 data can be found at <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview> (accessed on 22 July 2022). SRTM data can be found at <https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1> (accessed on 22 July 2022). For malaria datasets, the database of the Directorate of Epidemiological Surveillance from South Kivu provincial health directorates of the DRC from 2018 to 2021 was exploited.

Acknowledgments: We are grateful to Eric Kalunda Panzi (Department of Community Health, Higher Institute of Medical Techniques of Kinshasa) and DRC’s Directorate of Epidemiological Surveillance who provided us with the malaria datasets. We would also like to thank the researchers who answered our questionnaire and kindly shared their knowledge.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Answers from 8 experts using Saaty’s continuous rating scale.

		Elevation	Wind Speed	Temperature	Precipitation	Humidity
Expert 1	Elevation	1	1	1	1/5	1
	Wind speed	1	1	5	1/5	1
	Temperature	1	1/5	1	1/9	1
	Precipitation	5	5	9	1	1
	Humidity	1	1	1	1	1
Expert 2	Elevation	1	1/9	1/9	1/7	1/3
	Wind speed	9	1	7	5	9
	Temperature	9	1/7	1	5	7
	Precipitation	7	1/5	1/5	1	7
	Humidity	3	1/9	1/7	1/7	1
Expert 3	Elevation	1	1/5	1/9	1/9	1/5
	Wind speed	5	1	1/5	1/9	1/5
	Temperature	9	5	1	1/5	1
	Precipitation	9	9	5	1	7
	Humidity	5	5	1	1/7	1
Expert 4	Elevation	1	1/7	1/5	5	1/5
	Wind speed	7	1	5	5	5
	Temperature	5	1/5	1	5	5
	Precipitation	1/5	1/5	1/5	1	3
	Humidity	5	1/5	1/5	1/3	1
Expert 5	Elevation	1	9	1/9	1	5
	Wind speed	1/9	1	1/9	1/9	1
	Temperature	9	9	1	1	7
	Precipitation	1	9	1	1	7
	Humidity	1/5	1	1/7	1/7	1
Expert 6	Elevation	1	1/5	1/9	1/5	1
	Wind speed	5	1	1/5	1/5	1/3
	Temperature	9	5	1	1	1
	Precipitation	5	5	1	1	1
	Humidity	1	3	1	1	1
Expert 7	Elevation	1	1/5	1/5	1/5	1/3
	Wind speed	5	1	3	1/3	1/3
	Temperature	5	1/3	1	1	1
	Precipitation	5	3	1	1	1/3
	Humidity	3	3	1	3	1
Expert 8	Elevation	1	1/9	1/9	1/9	1/9
	Wind speed	9	1	1/9	9	9
	Temperature	9	9	1	1	9
	Precipitation	9	1/9	1	1	9
	Humidity	9	1/9	1/9	1/9	1

References

1. Carter, R.; Mendis, K.N. Evolutionary and historical aspects of the burden of malaria. *Clin. Microbiol. Rev.* **2002**, *15*, 564–594. [CrossRef] [PubMed]
2. World Malaria Report 2021. Available online: <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021> (accessed on 10 January 2023).
3. Malaria's Impact Worldwide. 2021. Available online: https://www.cdc.gov/malaria/malaria_worldwide/impact.html (accessed on 12 November 2023).
4. Alonso, P.L. Malaria: A problem to be solved and a time to be bold. *Nat. Med.* **2021**, *27*, 1506–1509. [CrossRef] [PubMed]
5. Kim, Y.; Ratnam, J.; Morioka, Y.; Behera, S.; Tsuzuki, A.; Minakawa, N.; Sweijd, N.; Kruger, P.; Maharaj, R.; Imai, C.C.; et al. Malaria predictions based on seasonal climate forecasts in South Africa: A time series distributed lag nonlinear model. *Sci. Rep.* **2019**, *9*, 17882. [CrossRef]
6. Ferrao, J.L.; Niquisse, S.; Mendes, J.M.; Painho, M. Mapping and modelling malaria risk areas using climate, socio-demographic and clinical variables in Chimoió, Mozambique. *Int. J. Environ. Res. Public Health* **2018**, *15*, 795. [CrossRef] [PubMed]
7. Wieland, R.; Kuhls, K.; Lentz, H.H.; Conraths, F.; Kampen, H.; Werner, D. Combined climate and regional mosquito habitat model based on machine learning. *Ecol. Model.* **2021**, *452*, 109594. [CrossRef]
8. Rakotoarison, H.A.; Rasamimalala, M.; Rakotondramanga, J.M.; Ramiranirina, B.; Franchard, T.; Kapesa, L.; Razafindrakoto, J.; Guis, H.; Tantely, L.M.; Girod, R.; et al. Remote sensing and multi-criteria evaluation for malaria risk mapping to support indoor residual spraying prioritization in the central highlands of Madagascar. *Remote Sens.* **2020**, *12*, 1585. [CrossRef]
9. Adeola, A.M.; Olwoch, J.M.; Botai, J.O.; Rautenbach, C.d.; Kalumba, A.M.; Tsela, P.L.; Adisa, O.; Nsubuga, F.W.N. Landsat satellite derived environmental metric for mapping mosquitoes breeding habitats in the Nkomazi municipality, Mpumalanga Province, South Africa. *S. Afr. Geogr. J.* **2017**, *99*, 14–28. [CrossRef]
10. Hanafi-Bojd, A.A.; Vatandoost, H.; Oshaghi, M.; Charrahy, Z.; Haghdoost, A.; Zamani, G.; Abedi, F.; Sedaghat, M.; Soltani, M.; Shahi, M.; et al. Spatial analysis and mapping of malaria risk in an endemic area, South of Iran: A GIS based decision making for planning of control. *Acta Trop.* **2012**, *122*, 132–137. [CrossRef]
11. Minale, A.S.; Alemu, K. Mapping malaria risk using geographic information systems and remote sensing: The case of Bahir Dar City, Ethiopia. *Geospat. Health* **2018**, *13*, 660. [CrossRef]
12. Bhatt, B.; Joshi, J. Analytical hierarchy process modeling for malaria risk zones in Vadodara district, Gujarat. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *8*, 9–12. [CrossRef]
13. Asgarian, T.S.; Moosa-Kazemi, S.H.; Sedaghat, M.M. Impact of meteorological parameters on mosquito population abundance and distribution in a former malaria endemic area, central Iran. *Heliyon* **2021**, *7*, e08477. [CrossRef] [PubMed]
14. CDC—Parasites—Malaria. Available online: <https://www.cdc.gov/parasites/malaria/index.html> (accessed on 11 April 2023).
15. Life Cycle of Anopheles Species Mosquitoes | Mosquitoes | CDC. Available online: <https://www.cdc.gov/mosquitoes/about/life-cycles/anopheles.html> (accessed on 23 February 2023).
16. Panzi, E.K.; Kandala, N.I.; Kafinga, E.L.; Tampwo, B.M.; Kandala, N.B. Forecasting Malaria Morbidity to 2036 Based on Geo-Climatic Factors in the Democratic Republic of Congo. *Int. J. Environ. Res. Public Health* **2022**, *19*, 12271. [CrossRef]
17. Emina, J.B.; Doctor, H.V.; Yé, Y. Profiling malaria infection among under-five children in the Democratic Republic of Congo. *PLoS ONE* **2021**, *16*, e0250550. [CrossRef] [PubMed]
18. Nations, U. South Kivu Factsheet. 2015. Available online: https://monusco.unmissions.org/sites/default/files/south_kivu_factsheet_eng.pdf (accessed on 23 April 2023).
19. Climate: South Kivu in the Congo. Available online: <https://www.worlddata.info/africa/congo-kinshasa/climate-south-kivu.php#:~:text=South%20Kivu%20is%20one%20of,is%20warm%20or%20hot%20yearlong.&text=The%20number%20of%20hours%20of,the%20sun%20is%20actually%20visible> (accessed on 27 February 2023).
20. Bompangue, D.; Giraudoux, P.; Piarroux, M.; Mutombo, G.; Shamavu, R.; Sudre, B.; Mutombo, A.; Mondonge, V.; Piarroux, R. Cholera epidemics, war and disasters around Goma and Lake Kivu: An eight-year survey. *PLoS ONE Negl. Trop. Dis.* **2009**, *3*, e436. [CrossRef] [PubMed]
21. Google My Maps. Available online: <https://mymaps.google.com/> (accessed on 26 November 2023).
22. Shuttle Radar Topography Mission. Available online: <https://www2.jpl.nasa.gov/srtm/> (accessed on 18 December 2022).
23. ERA5 | ECMWF. Available online: <https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5> (accessed on 28 November 2023).
24. Global High Resolution Population Density Maps (Facebook Connectivity Lab, CIESIN) | UN-SPIDER Knowledge Portal. Available online: <https://www.un-spider.org/links-and-resources/data-sources/global-high-resolution-population-density-maps-facebook> (accessed on 18 December 2022).
25. Saaty, R.W. The analytic hierarchy process—What it is and how it is used. *Math. Model.* **1987**, *9*, 161–176. [CrossRef]
26. Al-Harbi, K.M.S. Application of the AHP in project management. *Int. J. Proj. Manag.* **2001**, *19*, 19–27. [CrossRef]
27. Tranmer, M.; Elliot, M. *Multiple Linear Regression*; The Cathie Marsh Centre for Census and Survey Research (CCSR): Manchester, UK, 2008; Volume 5, pp. 1–5.
28. gdalwarp—GDAL Documentation. Available online: <https://gdal.org/programs/gdalwarp.html#cmdoption-gdalwarp-r> (accessed on 23 February 2023).

29. Lawrence, M.G. The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bull. Am. Meteorol. Soc.* **2005**, *86*, 225–234. [[CrossRef](#)]
30. Eastman, J.R. Multi-criteria evaluation and GIS. *Geogr. Inf. Syst.* **1999**, *1*, 493–502.
31. Bayoh, M.N. Studies on the Development and Survival of *Anopheles gambiae* Sensu Stricto at Various Temperatures and Relative Humidities. Ph.D. Thesis, Durham University, Durham, UK, 2001.
32. Fuller, D.O.; Ahumada, M.L.; Quiñones, M.L.; Herrera, S.; Beier, J.C. Near-present and future distribution of *Anopheles albimanus* in Mesoamerica and the Caribbean Basin modeled with climate and topographic data. *Int. J. Health Geogr.* **2012**, *11*, 13. [[CrossRef](#)] [[PubMed](#)]
33. Stoops, C.A.; Gionar, Y.R.; Sismadi, P.; Elyazar, I.R.; Bangs, M.J.; Sukowati, S. Environmental factors associated with spatial and temporal distribution of *Anopheles* (Diptera: Culicidae) larvae in Sukabumi, West Java, Indonesia. *J. Med. Entomol.* **2007**, *44*, 543–553. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.