

RESEARCH

Open Access



# Understanding work-related travel and its relation to malaria occurrence in Thailand using geospatial maximum entropy modelling

Natalie Memarsadeghi<sup>1</sup>, Kathleen Stewart<sup>1\*</sup>, Yao Li<sup>1\*</sup>, Siriporn Sornsakrin<sup>2</sup>, Nichaphat Uthaimongkol<sup>2</sup>, Worachet Kuntawunginn<sup>2</sup>, Kingkan Pidtana<sup>2</sup>, Chatree Raseebut<sup>3</sup>, Mariusz Wojnarski<sup>2</sup>, Krisada Jongsakul<sup>2</sup>, Danai Jearakul<sup>3</sup>, Norman Waters<sup>2</sup>, Michele Spring<sup>2</sup> and Shannon Takala-Harrison<sup>4</sup>

## Abstract

**Background** Estimating malaria risk associated with work locations and travel across a region provides local health officials with information useful to mitigate possible transmission paths of malaria as well as understand the risk of exposure for local populations. This study investigates malaria exposure risk by analysing the spatial pattern of malaria cases (primarily *Plasmodium vivax*) in Ubon Ratchathani and Sisaket provinces of Thailand, using an ecological niche model and machine learning to estimate the species distribution of *P. vivax* malaria and compare the resulting niche areas with occupation type, work locations, and work-related travel routes.

**Methods** A maximum entropy model was trained to estimate the distribution of *P. vivax* malaria for a period between January 2019 and April 2020, capturing estimated malaria occurrence for these provinces. A random simulation workflow was developed to make region-based case data usable for the machine learning approach. This workflow was used to generate a probability surface for the ecological niche regions. The resulting niche regions were analysed by occupation type, home and work locations, and work-related travel routes to determine the relationship between these variables and malaria occurrence. A one-way analysis of variance (ANOVA) test was used to understand the relationship between predicted malaria occurrence and occupation type.

**Results** The MaxEnt (full name) model indicated a higher occurrence of *P. vivax* malaria in forested areas especially along the Thailand–Cambodia border. The ANOVA results showed a statistically significant difference between average malaria risk values predicted from the ecological niche model for rubber plantation workers and farmers, the two main occupation groups in the study. The rubber plantation workers were found to be at higher risk of exposure to malaria than farmers in Ubon Ratchathani and Sisaket provinces of Thailand.

**Conclusion** The results from this study point to occupation-related factors such as work location and the routes travelled to work, being risk factors in malaria occurrence and possible contributors to transmission among local populations.

\*Correspondence:

Kathleen Stewart  
stewartk@umd.edu

Yao Li  
liyao@umd.edu

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Malaria has been a public health priority of the Greater Mekong Subregion for years, although in recent years, high treatment failure rates in *Plasmodium falciparum* in eastern Thailand have necessitated rapid policy responses. To address the concerning rates of antimalarial drug resistance, In 2020, Thailand's Ministry of Public Health changed the first-line therapy for *P. falciparum* to artesunate-pyronaridine (AS-PY) in Sisaket, to address high treatment failure rates [1]. This resulted in a steep decline of *P. falciparum* cases to the point where *Plasmodium vivax* has become the predominant species causing malaria cases in Thailand, with a much smaller proportion of *P. falciparum* cases. Many of these cases have been clustered along international borders, including the borders with Myanmar and Cambodia [2]. However overall, malaria prevalence has been decreasing in recent years, and Thailand has a goal to eliminate malaria by 2024. From 2012 to 2017, there was a 67% decline in the number of cases including an accelerated decline of 39% from 2016 to 2017 [3]. While Thailand is moving forward with its elimination goals, there are still significant pockets of malaria in Thailand as recent reports [4] indicate that malaria cases have doubled in fiscal year 2022 compared with the previous fiscal year.

Estimating malaria risk across a region provides local health officials with information useful for the mitigation of possible local transmission paths, and for understanding the risk of exposure for local populations. In this study, a machine learning approach was utilized to estimate the spatial distribution of malaria in two provinces in eastern Thailand, Ubon Ratchathani and Sisaket, where both *P. vivax* and *P. falciparum* had been found in clinical studies and was being monitored by a team from the United States Armed Forces Research Institute of Medical Sciences (AFRIMS). The research objective was to estimate the association of occupation-related factors, including occupation type and work locations, with estimated malaria occurrence in these two provinces to understand how local populations may be vulnerable to malaria based on where they reside, where their workplace is located, and the routes they take to reach their workplace. Malaria occurrence here refers to the probability of having malaria at a location.

To generate an estimate of the spatial variability of malaria occurrence within these provinces, a maximum entropy-based modelling tool (MaxEnt) was used to model the malaria species distribution. MaxEnt is an ecological niche modelling tool that predicts species distribution across a defined geographic area using case data and relevant environmental variables [5]. The true distribution of a species is represented as a probability distribution that considers the constraints of incidence data in

relation to the empirical average of environmental data at sites where cases are detected [6–8]. MaxEnt produces a probability map showing whether a species existed in a location or not, and this tool was used to predict the species distribution of *P. vivax* malaria for the period, January 2019 to April 2020 in Ubon Ratchathani and Sisaket. Environmental and sociodemographic variables were used to train the ecological niche model, including precipitation, temperature, elevation, land cover and population density as well as the locations of known local positive *P. vivax* and *P. falciparum* cases in these two provinces. These variables were used to determine the suitable environmental conditions for local malaria transmission and to create a probability surface that captures the ecological niche, reflecting the likelihood of malaria occurrence across the study area. The relationship between the probability of malaria occurrence and occupation type, work-related locations, as well as estimated travel routes between home and work locations (at village-level granularity) has been analysed to better understand the impact that residential and work locations, as well as work-related travel have on risk of exposure in outdoor settings and local transmission of malaria within this study area.

Several studies have used an ecological niche modelling approach to model the species distribution of *Anopheles* mosquitoes. Studies where the maximum entropy tool was used to determine the spatial distribution of the local *Anopheles* mosquito species have been undertaken in regions all over the world, including Asia and the Middle East [9], South America [10, 11], and Africa [12, 13]. MaxEnt has also been used to model the effects of climate change on malaria transmission and *Anopheles* species density in China [14], Iran [15], and parts of Africa [16]. In the model, local *P. vivax* malaria cases were used to train the machine learning model and estimate the percent probability of malaria occurrence given certain geographical and population factors.

The correlation between demographic variables and malaria risk has been previously studied, for example, an investigation of malaria risk factors in northern Namibia used survey data of malaria positive and negative individuals to identify significant risk factors for malaria in Namibia. In this study, cattle herders and police were the occupation types at the highest risk, and cross-border travel, open sleeping quarters, and working in agriculture overnight had a significant relationship with malaria infection [17, 18]. Significant risk factors for *P. vivax* malaria infections in Western Thailand included previous history of clinical malaria, occupation in agriculture, and travel to Myanmar [19]. In further studies, malaria prevalence has been very high in forest and forest fringes along rural stretches of border areas in Thailand, and

forest workers, and local migrants were the individuals at highest risk [20–23]. Researchers found that challenges for malaria control existed especially in border regions with highly mobile migrant workers [24]. Recent research findings showed that limited vector protection for forest goers was one of the main reasons underlying malaria transmission hotspots located in the border areas [25, 26]. As occupation and work-related travel were expected to play a key role in exposure to malaria infection [25], this study focused on how different environmental factors impact the spatial distribution and extent of the ecological niche in Ubon Ratchathani and Sisaket, a *P. vivax* transmission prone area of eastern Thailand.

**Methods**

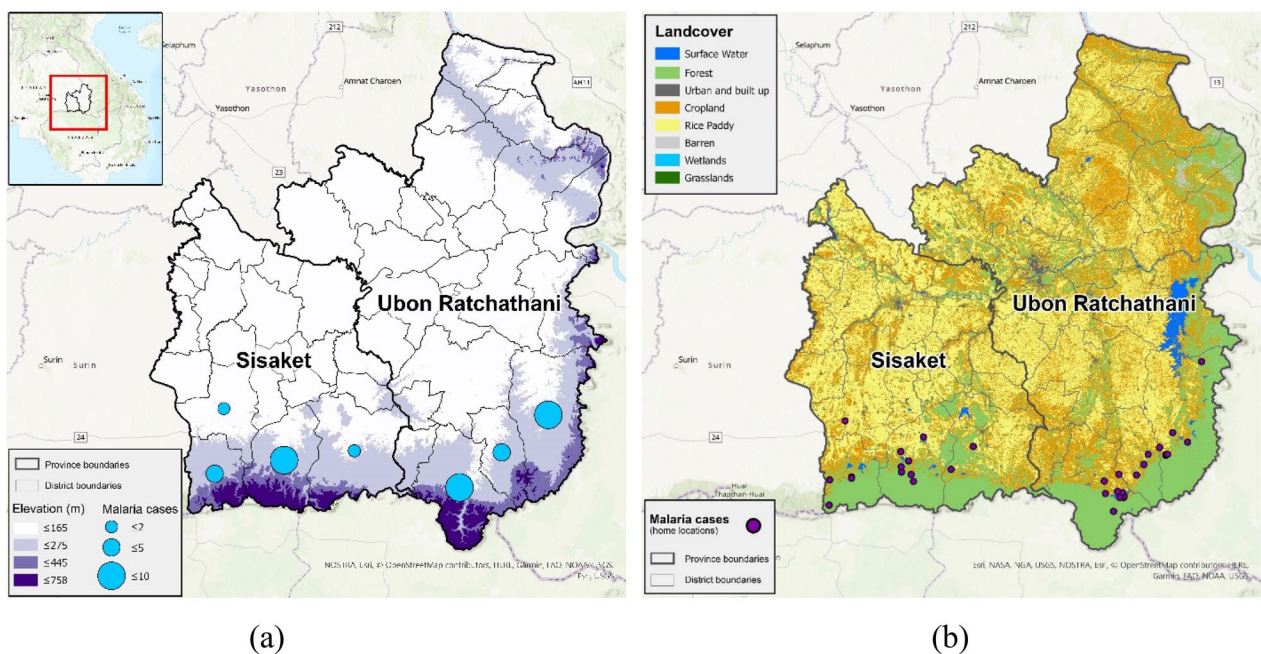
**Data and study area**

Ubon Ratchathani and Sisaket were selected based on an ongoing collaboration with researchers at the United States Armed Forces Research Institute of Medical Sciences (AFRIMS) based in Bangkok, Thailand who were conducting clinical studies of malaria in these two provinces (Fig. 1a).

Travel histories and other demographic data were collected by the AFRIMS research team as part of an ongoing study from March 2019 to April 2020 from 40 individuals reporting to malaria clinics and posts in Ubon Ratchathani and Sisaket. Polymerase chain reaction tests were applied to all samples to confirm speciation. Nearly all cases (39 cases) were diagnosed as *P. vivax*. (Fig. 1a).

Due to the single case of *P. falciparum* that was reported, the study sample data is referred collectively as "malaria case data," with the majority of cases being *P. vivax*. The survey participants met several requirements to be selected as participants in the study by AFRIMS, namely they were males or non-pregnant females, were older than 18 years and were civilian or military. They understood spoken Thai or a local language. In the survey, participants were asked to name their (permanent) home village to confirm that the data represent a local population. Other data used in this study included variables such as malaria diagnosis, age, sex, occupation type, and GPS coordinates for work locations and travel distance to work. For the AFRIMS survey data, the GPS coordinates were obtained for an approximate location near to their permanent residence in the village where they reside, and for work locations the approximate geographic coordinates capturing a close location were obtained for where they worked. This information was used in the analysis of occupation and work-related travel in relation to the estimated probability of malaria occurrence.

A landcover product for Myanmar, SERVIR KEKONG Land Cover Portal, was used for ecological niche modeling [27]. Based on the dataset, the primary land cover classes for Ubon Ratchathani and Sisaket provinces were forested areas, crop lands, and rice paddy fields. There is an escarpment with higher elevation (peak of 761 m) that is highly forested along the border of both provinces with Cambodia (Fig. 1b). This escarpment area has very low

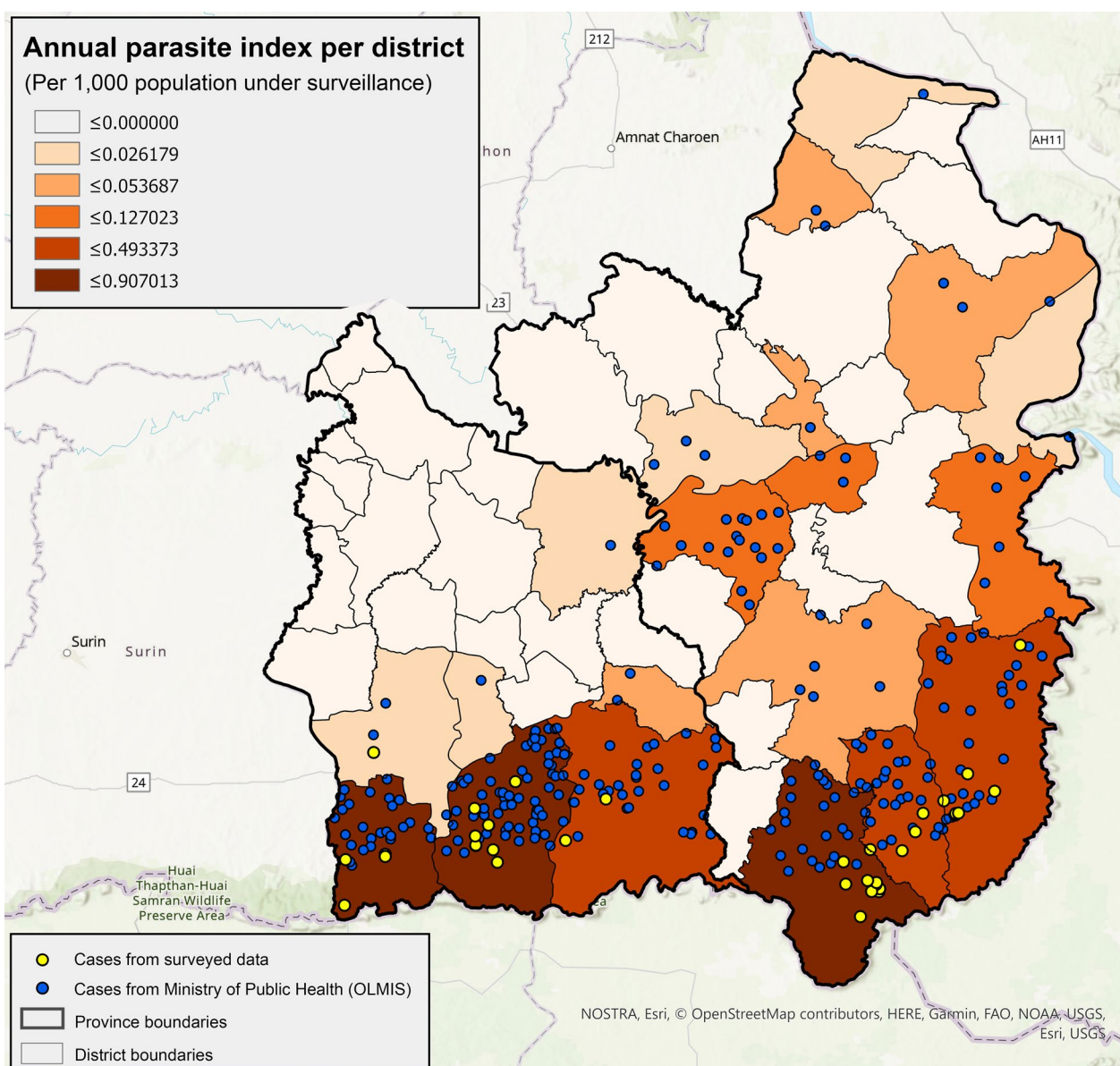


**Fig. 1** a Study area of Ubon Ratchathani and Sisaket, Thailand and the malaria cases per district in study sample enrolled by AFRIMS, and b landcover from SERVIR KEKONG Land Cover Portal and home village locations for malaria cases from data provided by AFRIMS

population density, however many of the surveyed individuals regularly travelled into this region for their work.

In addition to the research data collected by AFRIMS, the malaria case data from the online malaria information system (OLMIS) managed by the Ministry of Public Health for Thailand (MOPH) was also obtained (Fig. 2). These data included the number of *P. vivax* and *P. falciparum* malaria cases in each district for Ubon Ratchathani and Sisaket from January 2019 to April 2020. The MOPH recorded 125 *P. vivax* malaria cases in Ubon Ratchathani and 201 cases in Sisaket, for a

total of 326 *P. vivax* cases within the two provinces during this period. It is possible that the 40 cases identified by the AFRIMS study are included in the 326 cases reported by the OLMIS dashboard. After simulating locations for these cases, these data were used to train the machine learning model for malaria niche estimation. Lastly, the road network data were obtained from Open Street Map (OSM) for Ubon Ratchathani and Sisaket, and used these data to estimate travel routes from home villages to estimated work locations of the survey participants.



**Fig. 2** Annual parasite index by district recorded by the OLMIS between January 2019 and April 2020 classified using natural breaks overlaid with both simulated home village locations from OLMIS, and recorded home village locations from AFRIMS survey

Eight different remotely-sensed environmental data layers were used for malaria ecological niche estimation including total precipitation, landcover, population, elevation, slope, and average, minimum and maximum temperatures for both provinces. Precipitation estimates were obtained from WorldClim and included the monthly total amount of precipitation at a 21 km resolution, which was then averaged to create an annual precipitation layer for 2018 [28]. Landcover data are categorical and sourced from the Mekong Landcover Viewer [29]. From the original data, there are 20 different land cover classifications, however these classes were consolidated into nine classes by generalizing multiple forest classes to a general classification of forested areas. Shuttle Radar Topography Mission (SRTM) 2014 elevation data was used in this research, including for calculating slope using ArcGIS Pro 2.6.2 [30]. Temperature variables were annual averages for 2020 and were obtained from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) [31]. Lastly, population density data were sourced from Oak Ridge National Laboratory's 2019 Landsat resource [32].

## Methodology

Using the malaria case data and the environmental data, the malaria parasite distribution was determined using an ecological niche model for the study area. The relationship between the ecological niche, and work locations and occupation types were analysed using a leave-one-out validation process to determine which variables have the strongest relationship with areas with a higher probability of malaria occurrence. To implement this process, a number of models were created. Each variable was excluded in turn, and a model created with the remaining variables. Then a model was created using each variable in isolation. In addition, a model was created using all variables, as before. The area under the ROC curve (AUC) values from all the results were compared to determine the relative strength of each variable.

As the cases provided through the OLMIS for Thailand were geolocated only to the district level, a random simulation workflow was conducted using R to generate estimates of home village locations for these malaria-positive individuals (Fig. 2).

The random simulation process was executed based on the following rules:

- Points were randomly generated within the district boundary;
- The number of points generated corresponded to the number of cases within each district;
- Locations with low residential population numbers based on Landsat 2019 population estimates were

excluded, including the forested escarpment area along the Cambodia border;

- Uniform random sampling was used as the population distribution for the area is widely spread, with only a few urban areas with higher population density such as the City of Ubon Ratchathani, the former capital of the province, with a population of 79,000 (Department of Provincial Administration, 2014). However, the risk of contracting malaria in urban areas is likely to be less than in rural areas as health care providers tend to cluster in urban areas, and the lack of habitats for the vector and a lack of travel to rural areas in urban areas. Public health measures are more likely to be in place and to have sufficient funding to be maintained in urban areas [33, 34].

The generated points were used as occurrence data in the ecological niche model. One thousand different sets of random points by district were simulated in order to generate the highest level of randomness and to avoid bias in the results [35, 36]. All environmental variables were pre-processed using R to be the same projection, extent, and spatial resolution.

The MaxEnt modelling tool was run over 1000 iterations for each set of randomly generated home village locations and the eight pre-processed environmental layers in R. Three different training–testing splits were analysed to test the sensitivity of the models and evaluate the quality of the results including 70-30, 80-20 and 100% of the data as training data. The area under the curve (AUC) values for the 1000 iterations were averaged to determine the training–testing split that was used for the analyses.

A random sampling approach was used to estimate 100 work locations for rubber plantation workers and farmers. Farmers were assumed to have worked in both “Rice paddy” and “Cropland” regions while rubber plantation workers were assumed to work in “Forest” regions (Fig. 2). This assumption was tested using the demographic data collected by AFRIMS. The work locations were then generated using a uniform random sampling package from R package ‘sf’ applying these restrictions for the working regions for farmers and rubber plantation workers, respectively.

Logistic regression was applied to estimate the association between the estimated probability of malaria occurrence and occupation types. This study focused only on two occupations, farmers and rubber plantation workers, classified as categorical variables when building the model. These two occupation groups were tested because they comprised 65% of the collected data and represent different land cover classes (forested areas, croplands, and rice paddies), allowing us to analyse one group that works in forested areas and one that does not. A one-way

Analysis of Variance (ANOVA) test was applied to determine the difference in malaria risk between different occupations. In this study, the probability of malaria was measured being present at a location as determined from the ecological niche model. This probability served as the dependent variable while occupation type was the independent variable. The ANOVA test made it possible to understand if there were significant differences in the probability of malaria at the work locations for either rubber plantation workers or farmers. The ecological niche model values were extracted at the estimated work locations obtained based on the Landcover data and analysed for the corresponding occupation type.

To further analyse the relationship between occupation-related factors and malaria occurrence, a road network and buffer analysis was conducted to determine the risk of exposure that could be associated with the travel routes to work. The shortest paths from home villages to the estimated work locations for the cases identified by AFRIMS were calculated using ArcGIS Pro 2.6.2 and the OSM street network. Using a 1 km buffer around work locations and a 500-m buffer that extended from the work location along the shortest path routes, these areas were also analysed with respect to their overlap with the niche areas.

## Results

Base on the AFRIMS survey data, the primary occupation types of those surveyed included farmers (35%), rubber plantation workers (30%), service workers (12.5%), forestry officers (10%) and monks (7.5%). When comparing the results of the three different training and testing split model runs for estimating ecological niche, there was little deviation in the average AUC values used to assess predictive accuracy, and the contributions of environmental variables. The average AUC differed less than 0.015 between the model runs, and the top contributing environmental variables were consistently elevation, average and maximum temperature, and landcover (Table 1). As the 70-30 training testing split results returned the highest accuracy of 79%, this split was used to further analyse occupation and travel in the context of ecological niche.

The estimated ecological niche results showed that the highest probability locations of malaria occurrence were located along the Thailand–Cambodia border, particularly in the forested escarpment area (Fig. 3). The highest probability locations were also detected for northern Ubon Ratchathani province near the Thailand-Lao border. Using a natural breaks classification, the model results were classified into four classes reflecting the probability of malaria occurrence and assigned work locations from AFRIMS survey data into a class based on

**Table 1** Percent contributions of environmental variables used to train the ecological niche model

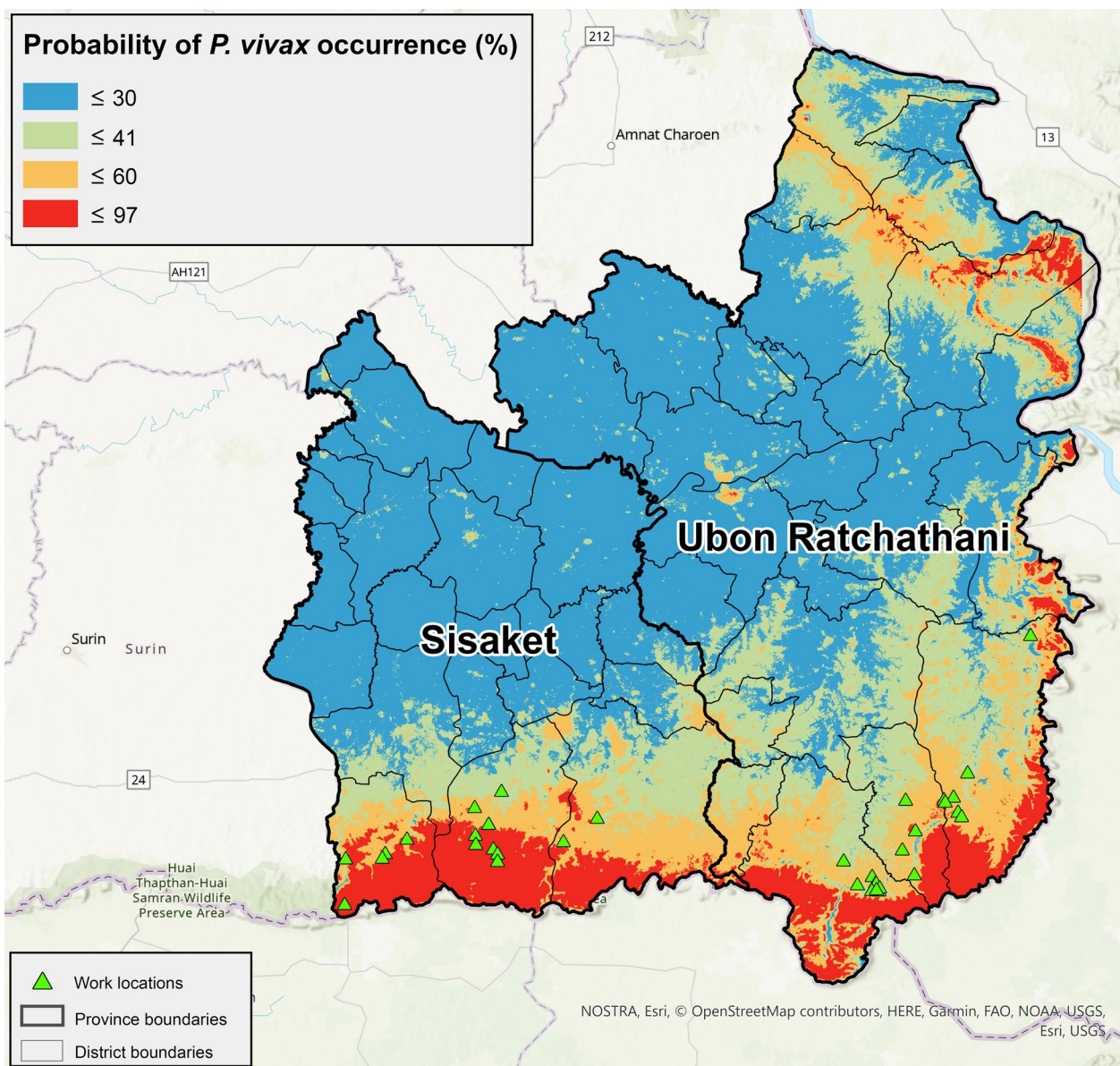
Environmental Variable	70-30 (%)	80-20 (%)	100 (%)
Elevation	49.61	49.82	49.57
Landcover	8.88	8.75	8.58
Population	0.99	0.99	0.97
Precipitation	3.6	3.57	3.23
Slope	3.45	3.35	3.27
Average temperature	16.76	16.45	17.36
Maximum temperature	10.84	11	11.02
Minimum temperature	5.85	6.01	6

the ecological niche model value at these locations. The results showed that the majority of the study sample population fell into class 3 (79%), where there was a 41–60% chance of malaria occurrence, with 95% of work locations falling in class 3 or 4 (Fig. 3). None of the work locations were found in a class 1 area, i.e., areas with the lowest probability of malaria occurrence.

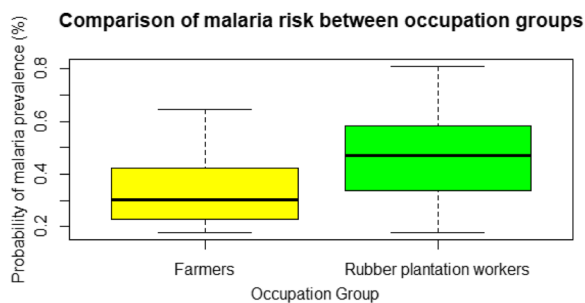
To understand the potential malaria risk for rubber plantation workers and farmers, the simulated work locations for the two occupation groups were analysed using the landcover data and the study sample data was used to validate the simulation points. The results showed that 72% of simulated work locations for both farmers and rubber plantation workers were consistent with the reported work locations from the survey.

A logistic regression analysis was applied to occupation type and predicted MaxEnt values. The outputs showed the coefficient for predicted MaxEnt value was positive with a coefficient of 9.03 and a p-value of  $9.3e-10$ . The one-way ANOVA test returned a statistically significant difference between the average MaxEnt predicted values for the two occupation types of rubber plantation workers and farmers, returning a p-value of  $1.73e-13$ . Based on the box-plot results, the average MaxEnt predicted values for rubber plantation workers were higher than those for farmers (Fig. 4). Rubber plantation workers were estimated to have significantly higher risk of exposure to malaria compared to farmers based on the predicted values from the machine learning analysis.

Analysing the estimated travel routes between home village and work locations in the context of the niche extent showed that travel for the two occupation groups occurred in areas with higher probability of malaria occurrence (Fig. 5). The travel routes were found to go through areas corresponding to higher niche classes. This is important to understand because regular daily travel increases the chance of exposure to infection by travelling through higher probability areas of local malaria



**Fig. 3** Estimated ecological niche and work locations based on the GPS coordinates from AFRIMS data

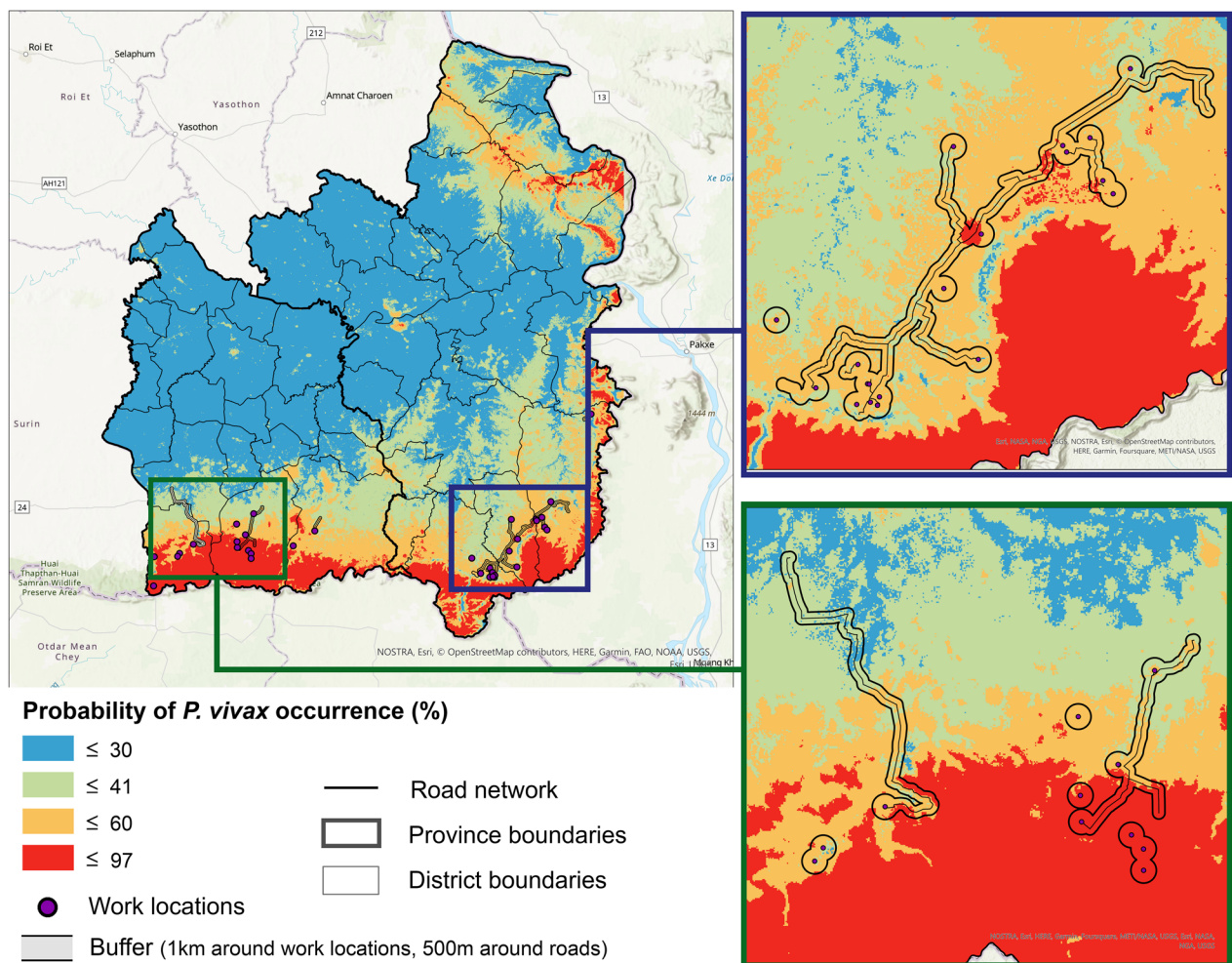


**Fig. 4** Comparison of average MaxEnt values by occupation type

occurrence. When averaging model values within buffer areas, the results showed that rubber plantation workers were travelling in areas with higher probabilities of malaria occurrence (average ecological niche value of 0.515) slightly more than other occupations (average ecological niche value within buffers for all occupations of 0.503).

**Discussion**

Computing the ecological niche showed higher probabilities of *P. vivax* occurrence was along the Thailand–Cambodia border. Additional high probability locations



**Fig. 5** Analysing occupation-related travel routes and malaria occurrence probability using buffer analysis

were detected in northern Ubon Ratchathani province near the Thailand-Lao border. The species distribution was modelled primarily for *P. vivax* malaria as there was a higher probability of *P. vivax* occurrence reflected by both the MOPH and AFRIMS data. This region of the world is also commonly referred to as having “border malaria”, i.e., a higher prevalence of malaria in South-east Asia exists along international borders [2, 37, 38] and the niche analysis also returned this pattern of species distribution. Similar to the findings in this research, a study focused on the Buriram and Surin provinces of Thailand determined through multiple linear regression analysis that “the high-risk areas of malaria cases were on the Thai-Cambodian border” and that malaria morbidity rates were strongly associated with forested areas [39]. Further, in an overview of the epidemiological patterns of malaria in the Greater Mekong Subregion from 1998 to 2007, researchers found that malaria was most prevalent in forested and forest-fringe areas, primarily along

international borders [40]. The niche model results also reflected these findings, as it similarly predicted higher probabilities of *P. vivax* occurrence in a similar ecological area. The red and orange regions of the border areas which are representing regions with high risk of malaria of the mapped niche shown in Figs. 4, 5, are also consistent with the active malaria transmission foci map for 2020 for Ubon Ratchathani and Sisaket as produced by the Sudathip et al. [41].

The ANOVA test results showed significant differences between forest and non-forest-based occupations, indicating a significant relationship between occupation type and working in an area of higher malaria occurrence. Further analysis showed that compared to farmers, rubber plantation workers were working in areas with a higher risk of malaria occurrence. This higher risk may be due to the fact that the working hours and outdoor practices of rubber plantation workers often overlap with peak mosquito biting times [42, 43].



Previous studies have found a relationship between occupation type and malaria risk. For example, out of over 4000 malaria cases reviewed in seven provinces of southern Thailand (Chumphon, Nakhon Si Thammarat, Krabi, Phangnga, Phuket, Ranong and Surat Thani provinces), nearly 62% malaria cases occurred in rubber plantation workers [44]. One study, based on qualitative ethnographic research of the Oddar Meanchey Province of Cambodia, found that those who worked in forests experienced multiple episodes of malaria more frequently than other occupations. Local residents understood the increased malaria risk, but for economic purposes, continued to pursue forest-based livelihoods. This geographic pattern is further supported by the analysis in this study, which estimated the highest probability of malaria occurrence in the forested areas along the Thailand–Cambodia border, and observed low but positive correlations between occupation types that worked in forests [45]. Researchers also found that forest-related occupations and economic drivers, such as logging of rosewood trees and other luxury timber, were the main reason for forest going which increase their exposure to malaria [46]. Similarly, another study conducted in Prachuap Khiri Khan Province of Thailand found that occupation as a rubber farmer/tapper was highly correlated with malaria affected households, and this correlation was statistically significant [47]. The same conclusion was drawn in this research that these same occupations worked in and travelled through areas with higher niche values and probabilities of malaria occurrence. The findings showed that likely daily travel routes take individuals from homes that may not be in a risk area to a work location that is in a risk area and that travel can be a possible driver for infection. Travel has been shown to play a key role in transmission [48] and, therefore, working in an area corresponding to a higher niche category, and travelling in and out, may increase the chance of transmission and infection. A recent study in Myanmar also found that forest workers have a higher malaria exposure compared with farmers as they were more likely to travel to regions with higher malaria prevalence for work [49].

One possible limitation of this research is that the analysis did not consider the presence of existing malaria control measures such as personal protection behaviours, or preventive measures for the simulation analysis. Likewise, the cases that might be due to staying outside at work locations were not considered. For the analysis of travel routes, although a 1 km buffer around work locations and a 500-m buffer along the travel routes were used to include more potential travel areas and to mitigate the uncertainty of the travel routes, these are estimates only and require more data to refine these values. For the AFRIMS survey data, the GPS coordinates were

obtained using an approximate location near to the permanent residence in the village where they reside, and for work locations approximate geographic coordinates capturing a close location for where they worked were used. Future research could include ground truth data for the work locations and travel routes to validate the simulated routes used in this study. For example, GPS trajectory data can be used to provide more details about travel routes [50, 51]. For this study, district-level data was used when simulating the home locations for *P. vivax* cases from the OLMIS dataset. Future research could be undertaken at sub-district or village level data. More detailed and precise landcover data can also be used in the future to differentiate the working regions for rubber plantation workers and that of other forest workers. For this study, the possibility of recrudescence or relapse of *P. vivax* was not taken into account and for the AFRIMS survey data, it is not known if these were newly infected or recurrent cases. The data from the OLMIS dashboard also did not specify whether cases were first episodes or likely relapses. Distinguishing new infections from relapses is a continuing challenge in studies of *P. vivax*, and could have impacted the inferences.

## Conclusions

In this study, the probability of malaria occurrence across Ubon Ratchathani and Sisaket provinces of Thailand was examined. A maximum entropy model was developed that used the known locations of positive malaria cases sourced from OLMIS, to produce a mapped probability surface that represented the percent probability of malaria occurrence in these provinces. Other inputs to the model included environmental variables sourced from remotely sensed datasets including temperature, landcover, precipitation, and elevation, and population density, and these were used to train the machine learning model. The highest probability of malaria occurrence in Sisaket was found along the Thailand–Cambodia border. For Ubon Ratchathani, other than the Thailand–Cambodia border area, high malaria risk areas were also detected in the north of Ubon Ratchathani near the Thailand-Lao border.

An analysis of the relationship between occupation types and the model results returned a significant difference between the average probability of malaria occurrence for the work locations simulated using landcover data of different outdoor occupation groups. As part of this research, travel routes were generated and applied buffer analysis and then analysed these areas in the context of the probability of malaria occurrence to understand the relationship between daily work-related travel and the potential risk of exposure to malaria. The results showed that individuals working in and travelling to areas

with higher probabilities of malaria occurrence were at higher risk, and that rubber plantation workers were travelling in areas with higher probabilities of malaria occurrence more frequently than other occupation types. Daily travel to work that takes individuals into areas with higher probabilities of malaria occurrence raises the risk of exposure relating to these travel behaviours.

#### Abbreviations

ANOVA	Analysis of variance
AUC	Area under the curve
MOPH	The Ministry of Public Health for Thailand
OSM	Open street map
OLMIS	Online malaria information system
MaxEnt	Maximum entropy-based modelling

#### Acknowledgements

We would like to thank Montri Arsanok, Ladaporn Bodhidatta, Parat Boonyarangka, COL Nithinart Chaitaveep, COL (Ret.) Mark Fukuda, CAPT John S. Griesenbeck, CDR Nicholas Martin, Chaiyawat Mathavarat, Kittijarankon Phontham, Phimphan Pisutsan, Kamonporn Poramathikul, COL (Ret.) Philip Smith, Sabaithip Sriwichai, Umaporn Suksawad, Chadin Thongpiam, COL Darunee Utennam, MAJ Brian Vesely, and Saowaluk Wongarunkochakorn for their help in collecting data.

#### Disclaimer

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the author, and are not to be construed as official, or as reflecting true views of the Department of the Army or the Department of Defense. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70-25.

#### Author contributions

Designed the experiments: NM, YL and KS. Collected data: SS, NU, WK, KP, CR, MW, KJ, DJ, NW, and MS. Analysed data: NM, and YL. Interpreted the data: NM, YL, ST-H and KS. Wrote and revised the manuscript: NM, KS, ST-H and YL. All authors read and approved of the final manuscript.

#### Funding

Research reported in this publication was supported by a National Institute of Allergy and Infectious Diseases R01 award to ST-H (NIH 1R01AI145852) and by a National Science Foundation award to KS (BCS2049805). The malaria surveillance study is supported by funding obtained through US Department of Defense Global Emerging Infections Surveillance and Defense Malaria Assistance Programs. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or National Science Foundation.

#### Availability of data and materials

Landcover data is publicly available through <http://servir-rlcms.appspot.com/static/html/home.html>.

Malaria cases data from MOPH is available through: [http://malaria.ddc.moph.go.th/malariar10/index\\_newversion.php](http://malaria.ddc.moph.go.th/malariar10/index_newversion.php)

#### Declarations

##### Ethics approval and consent to participate

Walter Reed Army Institute of Research (No. IRB00000794), Institute for the Development of Human Research Protections of Thailand (No. IRB00006539) and Royal Thai Army Medical Department (No. IRB00001439) approved the studies that collected the data. All methods were carried out in accordance with relevant guidelines and regulations.

##### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA. <sup>2</sup>Armed Forces Research Institute of Medical Sciences (AFRIMS), Bangkok, Thailand. <sup>3</sup>Division of Vector Borne Diseases, Ministry of Public Health, Ubon Ratchathani, Thailand. <sup>4</sup>Center for Vaccine Development and Global Health, University of Maryland School of Medicine, Baltimore, MD, USA.

Received: 19 October 2022 Accepted: 1 February 2023

Published online: 13 February 2023

#### References

- WHO. World malaria report 2021. Geneva: World Health Organization; 2021.
- Ministry of Public Health. National malaria elimination strategy 2017–2026. Asia Pacific Malaria Elimination Network, 2020.
- WHO. The Mekong malaria elimination programme: countries of the Greater Mekong making inroads in the race towards elimination. Geneva: World Health Organization; 2017.
- Ministry of Public Health, Thailand. Thailand malaria elimination program. 2022 [http://malaria.ddc.moph.go.th/malariar10/index\\_newversion.php](http://malaria.ddc.moph.go.th/malariar10/index_newversion.php). Accessed 8 Sep 2022.
- Phillips SJ. A brief tutorial on Maxent. *ATT Res.* 2005;190:231–59.
- Elith J, Leathwick JR. Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst.* 2009;40:677–97.
- Phillips SJ, Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography.* 2008;31:161–75.
- Guisan A, Thuiller W. Predicting species distribution: offering more than simple habitat models. *Ecol Lett.* 2005;8:993–1009.
- Ma A, Wang JF, Wang DQ, Ren ZP. Prediction of potential distribution of *Anopheles sinensis* in China based on MaxEnt. *Chin J Vector Biol Control.* 2014;25:393–8.
- Valderrama L, Ayala S, Reyes C, González CR. Modeling the potential distribution of the malaria vector *Anopheles (Ano.) pseudopunctipennis* Theobald (Diptera: Culicidae) in Arid Regions of Northern Chile. *Front Public Health.* 2021;9:611152.
- Padilla O, Rosas P, Moreno W, Toulkeridis T. Modeling of the ecological niches of the anopheles spp in Ecuador by the use of geo-informatic tools. *Spat Spatio-Temporal Epidemiol.* 2017;21:1–11.
- Akpan GE, Adepoju KA, Oladosu OR, Adelabu SA. Dominant malaria vector species in Nigeria: modelling potential distribution of *Anopheles gambiae* sensu lato and its siblings with MaxEnt. *PLoS ONE.* 2018;13:e0204233.
- Gwitira I, Murwira A, Zengeya FM, Shekede MD. Application of GIS to predict malaria hotspots based on *Anopheles arabiensis* habitat suitability in Southern Africa. *Int J Appl Earth Obs Geoinformation.* 2018;64:12–21.
- Abrha H, Hagos H, Brhane E, Hadgu M, Mamo G. Spatio-temporal dynamics of malaria expansion under climate change in semi-arid areas of Ethiopia. *Environ Hazards.* 2019;18:400–13.
- Hanafi-Bojdi AA, Sedaghat MM, Vatandoost H, Azari-Hamidian S, Pakdad K. Predicting environmentally suitable areas for *Anopheles superpictus* Grassi (s.l.), *Anopheles maculipennis* Meigen (s.l.) and *Anopheles sacharovi* Favre (Diptera: Culicidae) in Iran. *Parasit Vectors.* 2018;11:382.
- Hundessa S, Li S, Li Liu D, Guo J, Guo Y, Zhang W, et al. Projecting environmental suitable areas for malaria transmission in China under climate change scenarios. *Environ Res.* 2018;162:203–10.
- Smith JL, Mumbengegwi D, Haindongo E, Cueto C, Roberts KW, Gosling R, et al. Malaria risk factors in northern Namibia: the importance of occupation, age and mobility in characterizing high-risk populations. *PLoS ONE.* 2021;16:e0252690.
- Jongdeepsai M, Ean M, Heng C, Buntan T, Tripura R, Callery JJ, et al. Acceptability and feasibility of malaria prophylaxis for forest goers: findings from a qualitative study in Cambodia. *Malar J.* 2021;20:446.
- Nguitragool W, Karl S, White M, Koepfli C, Felger I, Singhasivanon P, et al. Highly heterogeneous residual malaria risk in western Thailand. *Int J Parasitol.* 2019;49:455–62.

20. Corbel V, Nosten F, Thanispong K, Luxemburger C, Kongmee M, Chareonviriyaphap T. Challenges and prospects for dengue and malaria control in Thailand. *Southeast Asia Trends Parasitol.* 2013;29:623–33.
21. Suwonkerd W, Ritthison W, Ngo CT, Tainchum K, Bangs MJ, Chareonviriyaphap T. Vector biology and malaria transmission in Southeast Asia. In: Manguin S, editor. *Anopheles mosquitoes—new insights malar vectors.* London: IntechOpen; 2013.
22. Sukkanon C, Masangkay FR, Mala W, Kotepui KU, Wilairatana P, Chareonviriyaphap T, et al. Prevalence of *Plasmodium* spp. in *Anopheles* mosquitoes in Thailand: a systematic review and meta-analysis. *Parasit Vectors.* 2022;15:285.
23. Angrisano F, Robinson LJ. *Plasmodium vivax*—How hidden reservoirs hinder global malaria elimination. *Parasitol Int.* 2022;87:102526.
24. Khamsiriwatchara A, Sudathip P, Sawang S, Vijakadge S, Potithavoranan T, Sangvichean A, et al. Artemisinin resistance containment project in Thailand. (I): Implementation of electronic-based malaria information system for early case detection and individual case management in provinces along the Thai-Cambodian border. *Malar J.* 2012;11:247.
25. Roh ME, Lausatianragit K, Chaitaveep N, Jongsakul K, Sudathip P, Raseebut C, et al. Civilian-military malaria outbreak response in Thailand: an example of multi-stakeholder engagement for malaria elimination. *Malar J.* 2021;20:458.
26. Nofal SD, Peto TJ, Adhikari B, Tripura R, Callery J, Bui TM, et al. How can interventions that target forest-goers be tailored to accelerate malaria elimination in the Greater Mekong Subregion? A systematic review of the qualitative literature. *Malar J.* 2019;18:32.
27. SERVIR KEKONG Land Cover Portal. <http://servir-rlcms.appspot.com/static/html/home.html>. Accessed 12 Feb 2023.
28. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol.* 2017;37:4302–15.
29. Saah D, Johnson G, Ashmall B, Tondapu G, Tenneson K, Patterson M, et al. Collect earth: an online tool for systematic reference data collection in land cover and use applications. *Environ Model Softw.* 2019;118:166–71.
30. NASA Shuttle Radar Topography Mission (SRTM). Shuttle Radar Topography Mission (SRTM) Global. Distributed by OpenTopography. 2013. <https://doi.org/10.5069/G9445JDF>.
31. Wangdi K, Gatton ML, Kelly GC, Clements AC. Cross-border malaria: a major obstacle for malaria elimination. *Adv Parasitol.* 2015;89:79–107.
32. Rose A, McKee J, Sims K, Bright E, Reith A, Urban M. LandScan Global 2020 [Internet]. 2020th edn. Oak Ridge, TN: Oak Ridge National Laboratory; 2021. [landscan.ornl.gov](http://landscan.ornl.gov)
33. Parker DM, Carrara VI, Pukrittayakamee S, McGready R, Nosten FH. Malaria ecology along the Thailand–Myanmar border. *Malar J.* 2015;14:388.
34. Sudathip P, Kongkasuriyachai D, Stelmach R, Bisanzio D, Sine J, Sawang S, et al. The investment case for malaria elimination in Thailand: a cost-benefit analysis. *Am J Trop Med Hyg.* 2019;100:1445–53.
35. Syfert MM, Smith MJ, Coomes DA. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS ONE.* 2013;8:e55158.
36. Royle JA, Chandler RB, Yackulic C, Nichols JD. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods Ecol Evol.* 2012;3:545–54.
37. Cui L, Yan G, Sattabongkot J, Cao Y, Chen B, Chen X, et al. Malaria in the greater Mekong subregion: heterogeneity and complexity. *Acta Trop.* 2012;121:227–39.
38. Zhou G, Sirichaisinthop J, Sattabongkot J, Jones J, Bjørnstad ON, Yan G, et al. Spatio-temporal distribution of *Plasmodium falciparum* and *P. vivax* malaria in Thailand. *Am J Trop Med Hyg.* 2005;72:256–62.
39. Loyd RA, Kaewpitoon SJ. Malaria risk areas in Thailand border. *J Med Assoc Thai.* 2015;98:51.
40. Delacollette C, D'Souza C, Christophel E, Thimasarn K, Abdur R, Bell D, et al. Malaria trends and challenges in the Greater Mekong Subregion. *Southeast Asian J Trop Med Public Health.* 2009;40:674–91.
41. Sudathip P, Dharmarak P, Rossi S, Promda N, Newby G, Larson E, et al. Leveraging geo-referenced malaria information to increase domestic financial support for malaria elimination in Thailand. *Malar J.* 2022;21:213.
42. Tangena J-AA, Thammavong P, Wilson AL, Brey PT, Lindsay SW. Risk and control of mosquito-borne diseases in Southeast Asian rubber plantations. *Trends Parasitol.* 2016;32:402–15.
43. Bhumiratana A, Sorosjinda-Nunthawarasilp P, Kaewwaen W, Maneekan P, Pimnon S. Malaria-associated rubber plantations in Thailand. *Travel Med Infect Dis.* 2013;11:37–50.
44. Chaivisit P, Chujun S, Chutinantakul A. Modelling malaria incidence in the upper part of southern Thailand. *J Public Health Dev.* 2020;18:16–27.
45. Bannister-Tyrrell M, Gryseels C, Sokha S, Dara L, Sereiboth N, James N, et al. Forest goers and multidrug-resistant malaria in Cambodia: an ethnographic study. *Am J Trop Med Hyg.* 2019;100:1170.
46. Jongdeepaisal M, Khonputsa P, Prasert O, Maneenet S, Pongsoipetch K, Jatapai A, et al. Forest malaria and prospects for anti-malarial chemoprophylaxis among forest goers: findings from a qualitative study in Thailand. *Malar J.* 2022;21:8.
47. Satitvipawee P, Wongkhang W, Pattanasin S, Hoithong P, Bhumiratana A. Predictors of malaria-association with rubber plantations in Thailand. *BMC Public Health.* 2012;12:1115.
48. Smith DL, Perkins TA, Reiner RC Jr, Barker CM, Niu T, Chaves LF, et al. Recasting the theory of mosquito-borne pathogen transmission dynamics and control. *Trans R Soc Trop Med Hyg.* 2014;108:185–97.
49. Li Y, Stewart K, Han KT, Han ZY, Aung PP, Thein ZW, et al. Understanding spatio-temporal human mobility patterns for malaria control using a multi-agent mobility simulation model. *Clin Infect Dis.* 2023;76:e867–74.
50. Edwards HM, Sriwichai P, Kirabittir K, Prachumsri J, Chavez IF, Hii J. Transmission risk beyond the village: entomological and human factors contributing to residual malaria transmission in an area approaching malaria elimination on the Thailand–Myanmar border. *Malar J.* 2019;18:221.
51. Shirayama Y, Phompida S, Shibuya K. Geographic information system (GIS) maps and malaria control monitoring: intervention coverage and health outcome in distal villages of Khammouane province, Laos. *Malar J.* 2009;8:217.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

