

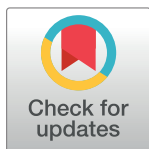
RESEARCH ARTICLE

Climatic, land-use and socio-economic factors can predict malaria dynamics at fine spatial scales relevant to local health actors: Evidence from rural Madagascar

Julie D. Pourtois^{1,2*}, Krti Tallam¹, Isabel Jones^{1,2}, Elizabeth Hyde³, Andrew J. Chamberlin², Michelle V. Evans⁴, Felana A. Ihtamamala^{5,6}, Laura F. Cordier⁶, Bénédicte R. Razafinjato⁶, Rado J. L. Rakotonanahary^{5,6}, Andritiana Tsirinomen'ny Aina⁶, Patrick Soloniaina⁶, Sahondraritera H. Raholiarimanana⁷, Celestin Razafinjato⁷, Matthew H. Bonds^{5,6}, Giulio A. De Leo^{1,2}, Susanne H. Sokolow^{8,9}, Andres Garchitorea^{4,6}

1 Biology Department, Stanford University, Stanford, CA, United States of America, **2** Hopkins Marine Station, Stanford University, Pacific Grove, CA, United States of America, **3** School of Medicine, Stanford University, Stanford, CA, United States of America, **4** MIVEGEC, Université de Montpellier, CNRS, IRD, Montpellier, France, **5** Department of Global Health and Social Medicine, Harvard Medical School, Boston, MA, United States of America, **6** NGO Pivot, Ifanadiana, Madagascar, **7** Programme National de Lutte contre le Paludisme, Ministère de la Santé Publique, Antananarivo, Madagascar, **8** Woods Institute for the Environment, Stanford University, Stanford, CA, United States of America, **9** Marine Science Institute and Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA, United States of America

* jp22@stanford.edu, juliepourtois22@gmail.com



OPEN ACCESS

Citation: Pourtois JD, Tallam K, Jones I, Hyde E, Chamberlin AJ, Evans MV, et al. (2023) Climatic, land-use and socio-economic factors can predict malaria dynamics at fine spatial scales relevant to local health actors: Evidence from rural Madagascar. *PLOS Glob Public Health* 3(2): e0001607. <https://doi.org/10.1371/journal.pgph.0001607>

Editor: Muhammad Asaduzzaman, University of Oslo Faculty of Medicine: Universitetet i Oslo Det medisinske fakultet, NORWAY

Received: September 20, 2022

Accepted: January 23, 2023

Published: February 22, 2023

Copyright: © 2023 Pourtois et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All code was written in R 4.0.3 and is available in the GitHub repository https://github.com/jpourtois/malaria_ifanadiana. Use of incidence and survey data was approved by the Ministry of Health for this study. The dataset is not available publicly but may be shared upon request to the address research@pivotworks.org.

Abstract

While much progress has been achieved over the last decades, malaria surveillance and control remain a challenge in countries with limited health care access and resources. High-resolution predictions of malaria incidence using routine surveillance data could represent a powerful tool to health practitioners by targeting malaria control activities where and when they are most needed. Here, we investigate the predictors of spatio-temporal malaria dynamics in rural Madagascar, estimated from facility-based passive surveillance data. Specifically, this study integrates climate, land-use, and representative household survey data to explain and predict malaria dynamics at a high spatial resolution (i.e., by Fokontany, a cluster of villages) relevant to health care practitioners. Combining generalized linear mixed models (GLMM) and path analyses, we found that socio-economic, land use and climatic variables are all important predictors of monthly malaria incidence at fine spatial scales, via both direct and indirect effects. In addition, out-of-sample predictions from our model were able to identify 58% of the Fokontany in the top quintile for malaria incidence and account for 77% of the variation in the Fokontany incidence rank. These results suggest that it is possible to build a predictive framework using environmental and social predictors that can be complementary to standard surveillance systems and help inform control strategies by field actors at local scales.

Funding: This work was supported by internal funding from PIVOT and grants from the Agence Nationale de la Recherche (Project ANR-19-CE36-0001-01), from Institut de Recherche pour le Développement (Project IRD Coup de Pouce "MAGIE"), and from the Herrnstein Family Foundation to AG. GADL was partially supported by the National Science Foundation (NSF DEB – 2011179) and the Belmont Collaborative Forum on Climate, Environment and Health (NSF ICER-2024383). JDP was supported by the Stanford Graduate Fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Interventions for malaria control have been highly successful in many countries over the last few decades, leading to a 47% decrease in mortality rates globally between 2001 and 2013 [1]. However, access to preventative and curative health care remains limited in many rural areas of low-income countries, where the burden of malaria is concentrated [2]. This problem is expected to intensify as population growth and climate change are projected to increase the number of people at risk for malaria in these areas [1]. Ensuring universal access to malaria diagnostics and treatment thus remains a key goal in the global malaria strategic plan for 2016–2030, which recognizes that improved surveillance is essential to inform those efforts [1]. However, limited human and financial resources are significant obstacles to reaching this goal in low-income countries [3]. Most malaria surveillance occurs passively, only capturing those cases that reach health facilities, missing the vast majority of infections where access to health care is low [4–6]. To help overcome these obstacles, improved methods to predict malaria incidence at local scales could help health care practitioners optimize the distribution of limited resources when and where they are most needed.

There has been substantial progress on predicting malaria at regional and national scales: aggregations of data from national Demographic and Health Surveys or from national surveillance systems, in combination with large-scale remotely sensed environmental data available in public repositories, are able to explain and predict spatial patterns in malaria at the global and regional scale with relatively good resolution. A leading example is the Malaria Atlas Project, which maps malaria incidence and relevant variables across the world [7], but many other studies are done at these scales [8–12]. While global and regional patterns may be useful for international organizations or national governments to estimate total malaria burdens and medical treatment needs [12], they cannot necessarily inform malaria control at the local level (e.g. within a government district), where most control activities are actually implemented [13]. Malaria incidence can be highly heterogeneous at small spatial scales [14], so a key need to improve local strategies to curb malaria transmission is to better understand how disease risk varies at fine spatial and temporal scales, and whether there are predictable factors (environmental, socio-economic, climatic) driving this variability. This could help medical practitioners and local health programs anticipate resource needs or inform the implementation of targeted control activities. There remain, however, substantial challenges to downscaling malaria predictions at these scales, especially the resolution and quality of current routine surveillance systems and the lower heterogeneity in environmental and socio-demographic predictors at small scales.

High quality data on malaria incidence is limited in many countries. Studies attempting to predict malaria patterns often rely on surveys or active surveillance methods, which have limited geographic and temporal scope [14–17]. Routine health system data provide a more sustainable data source and are available in most countries, but the spatial resolution of such data is low when aggregate numbers of cases over a catchment area are reported, as is often the case, rather than the precise locations of cases [18]. In addition, the quality of health facility-based data can be highly variable and difficult to assess [19], and there are many geographic and financial barriers to care, resulting in surveillance data biased towards populations living close to higher-quality health centers [18]. The health system burden may thus not be an accurate representation of the community burden [19]. As a result, new ways of reporting data and adjusting it to remove known biases are necessary before one can even attempt to understand local drivers of malaria, and predict the spatio-temporal variability of incidence.

Many social and environmental factors have been shown to influence exposure and vulnerability to malaria infection. Household wealth is generally associated with preventive

behaviours, while the coverage and use of insecticide-treated bed nets are effective in reducing malaria transmission [20], both of which can be derived from national surveys. Climatic and environmental factors such as temperature, precipitation and land use have been shown to affect mosquito habitat and life history traits [21,22]. The resolution of many of these predictors is constantly improving with the help of remote sensing, as is the resolution of predictions [23]. At national or regional scales, the heterogeneity (range of variation) of these predictors is high, making it feasible to detect associations of climatic variables with malaria incidence or prevalence data [9,17,24]. However, measuring heterogeneity presents a challenge at finer scales, even with environmental data remotely sensed at higher spatial resolution: a smaller range of variation in predictors, an increase in spatial auto-correlation, and a greater role of additional unobservable (stochastic) processes and behavioural factors can make fine scale projections of spatial variability incredibly challenging. As a result, it is unclear whether more localized studies, even with the best available data, can find consistent predictors of local disease dynamics at the finer scales at which health care interventions occur [25].

The aim of this work is to explore whether data gathered from routine surveillance systems can be used along with ecological, environmental and socioeconomic information to improve our understanding of spatio-temporal malaria dynamics at fine scales. We built on previous work where passive surveillance data on malaria incidence from the rural district of Ifanadiana, in south-eastern Madagascar, was adjusted to correct for reporting biases derived from financial and geographic barriers to health care [18]. The resulting data provide adjusted monthly malaria incidence per Fokontany (i.e. cluster of villages), improving nearly ten-fold the spatial resolution of malaria incidence rates as compared with aggregate health facility reports. We coupled this with satellite information on climate and land-use dynamics, as well as socio-economic information from a longitudinal cohort study representative of the district population. We then identified correlates of malaria incidence across time and space with a mixed-effects generalized linear model and explored causal relationships among these variables using structural equation modelling. Ifanadiana was an ideal region to explore these questions, as heterogeneity in climate, land-use, and household wealth is high relative to its spatial size [25]. Improving predictions of malaria dynamics at a high spatial resolution could pave the way towards local forecasting and early warning systems with operational applications for on-the-ground distribution of health care resources.

Results

A total of 326,334 malaria cases were estimated in the adjusted dataset for Ifanadiana from January 2014 to December 2017, equivalent to an average of 1.65 infections per capita over that four-year period. We observed large variations in infections across space and time. Total infections per capita over four years ranged from 0 to 6.9 across Fokontany. This corresponds to an average incidence of 42 cases per thousand people (‰) per month, ranging from 0 to 177 ‰ across Fokontany (Fig 1A). The discrepancy between the expected total number of infections based on these averages and the observed total number of infections presented above can be explained by missing data (each Fokontany was missing 4.7 months of data on average). The adjusted malaria incidence remained clustered along the main roads and in Fokontany close to health centers (Fig 1A). Malaria incidence was seasonal, with lower incidence from June to October (7–22 ‰ per month, averaged over all Fokontany and over 4 years) and higher incidence from November to May (36–92 ‰ per month, averaged over all Fokontany and over 4 years) (S1 Fig). November to May were also associated with high precipitation and high temperatures (S1 Fig). We focus our statistical analyses of malaria predictors on the high season to reduce the number of excess zeroes in our dataset and to increase our ability to capture spatial variation.

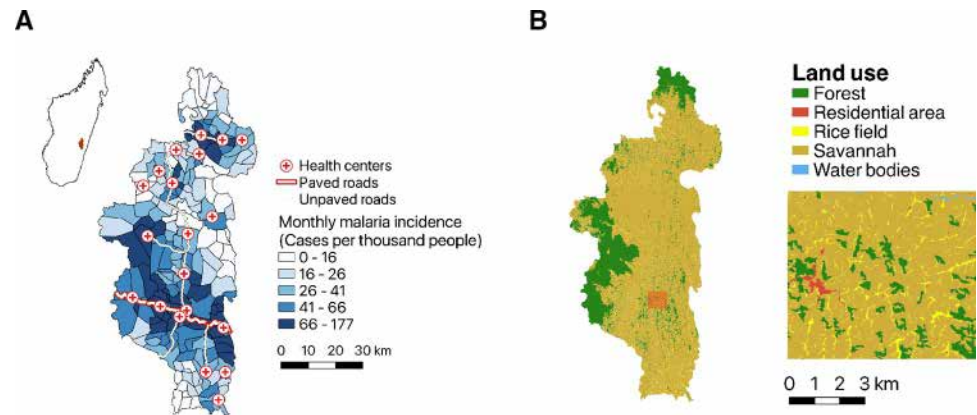


Fig 1. (A) Map of roads (paved and unpaved), health centers and adjusted monthly malaria incidence averaged over four years in Ifanadiana, and its location in Madagascar (inset map). (B) Broad and detailed view of the land use map and its five land classes. Both maps were made with QGIS, with boundary data from OCHA (<https://data.humdata.org/dataset/cod-ab-mdg>) under a CC BY 4.0 License and land use data available through OpenStreetMap.

<https://doi.org/10.1371/journal.pgph.0001607.g001>

The hottest month of the year was November, with a mean land surface temperature (LST) across 4 years of 33°C (S1 Fig). The coldest month of the high season (November to May) was May, with a mean LST across 4 years of 23°C. May was also the driest month of the high season, with an average total precipitation of 65 mm. The wettest months were January and March, with a mean precipitation over four years of 330 and 384 mm respectively. Land use varied spatially, with forested areas concentrated in the West and North, and an open landscape with residential areas and rice fields in the rest of the district (Fig 1B). Spatial distributions for all predictor variables are available in S2 Fig.

Socio-economic, land use and climatic variables accurately predicted spatial and temporal hotspots in GLMMs

The generalized linear mixed model (GLMM) included a zero-inflated negative binomial structure, fixed effects for socio-economic, land use and climatic variables (Fig 2A), and an Ornstein-Uhlenbeck and Matern covariance structure to account for temporal and spatial autocorrelation respectively (see Methods section). Overall, the conditional part of the model had a larger influence on the relationship between our predictors and malaria incidence than the zero-inflated part of the model (S1 Table). The socio-economic variables included in the GLMM were distance to health centers, wealth score and bed net use. An increase of one standard deviation (SD) in the log-transformed distance to health centers increased the odds of seeing a structural zero by 120% and decreased the expected malaria incidence by 19% (95% CI: 0.73–0.91, Fig 2A). An increase in log-transformed wealth score was associated with a 23% increase in the expected malaria incidence (95% CI: 1.08–1.40). Bed net use was not associated with a significant change in malaria incidence (95% CI: 0.83–1.13). Among land use variables, the log-transformed proportion of residential land use was negatively associated with malaria incidence, with a 14% decrease in malaria incidence with each SD change (95% CI: 0.78–0.94). The log-transformed proportion of rice fields was positively associated with malaria incidence, with a 12% increase in malaria incidence with each SD increase (95% CI: 1.00–1.26). Forest loss over the previous 3 years did not significantly affect malaria incidence (95% CI: 0.98–1.13). With regard to climate, a 1 SD increase in log-transformed total precipitation, lagged by one month, was associated with a 33% increase in malaria incidence (95% CI: 1.26–1.41). Mean

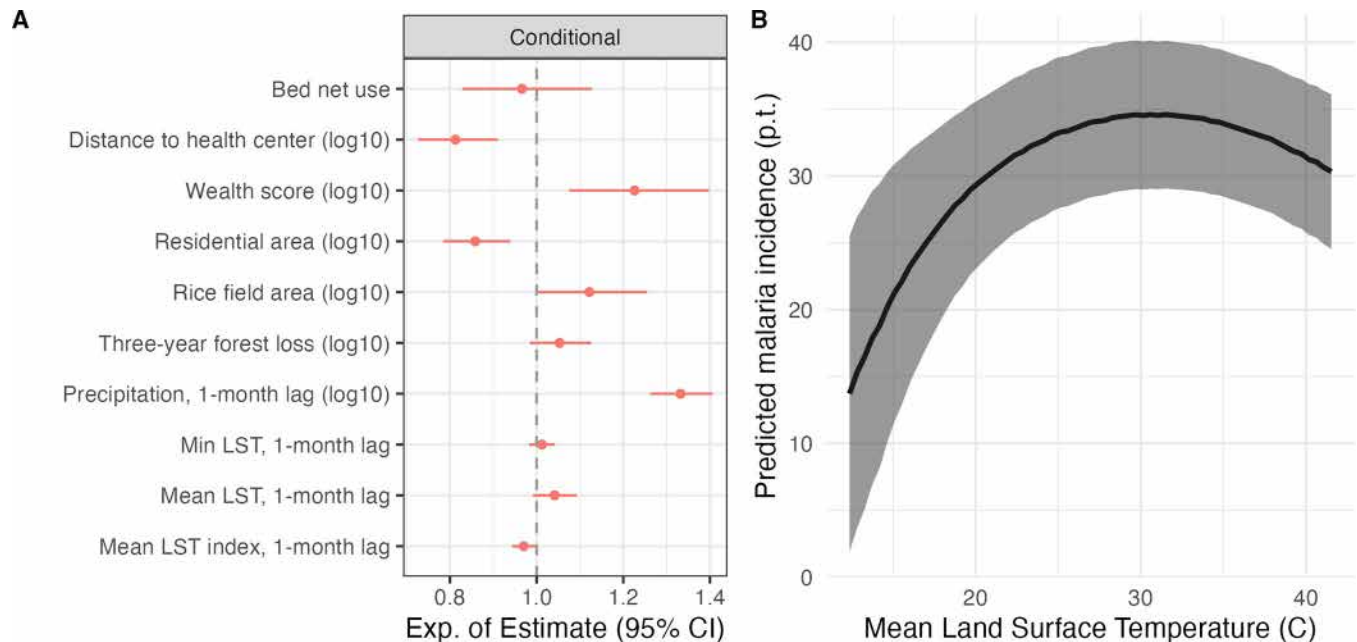


Fig 2. GLMM analysis. (A) Exponential of the coefficient estimates and 95% confidence interval for the conditional part of the GLMM with zero-inflated negative binomial distribution and spatial and temporal covariance structure. The exponential of the coefficient represents the expected multiplicative change in malaria incidence for a one-unit change of the predictor variable. Variables were scaled so that a one-unit change represents a change of 1 standard deviation for all variables. (B) Marginal effect of mean LST and mean LST index (one-month lag, standardized) on malaria incidence, conditioned on both conditional and zero-inflated parts of the model for a random Fokontany and random time point, with standard error interval.

<https://doi.org/10.1371/journal.pgph.0001607.g002>

land surface temperature (LST) was associated with a 46% decrease in the odds of a structural zero (95% CI: 0.34–0.86). Finally, mean temperatures closer to the optimal temperature of 25°C for malaria (i.e., mean LST index) were associated with higher malaria incidence (cond. factor: 0.97, z.i. factor: 1.31). This resulted in a unimodal relationship between mean LST and malaria incidence (Fig 2B).

We used a mixed-effect GLM with month and Fokontany ID as random effects to make predictions that are not dependent on malaria incidence in the prior month (unlike the model using a temporal covariance structure). Our in-sample model predictions across space and time had a root mean square error of 52 cases per thousand people, and a high correlation with observed malaria incidence in log-space (Pearson's $R = 0.73$). When averaged over time, our predictions and observed malaria incidence were strongly correlated (Pearson's $R = 0.997$, Spearman's $\rho = 0.997$, Fig 3A and 3B). More specifically, our predictions correctly identified 39 out of 40 (top 20%) fokontany with the highest average malaria incidence from 2014 to 2017. Our model was also able to capture the seasonality of malaria incidence during the high transmission season, capturing the incidence peak from January–March for all three years (S3 Fig). When trained on the 2014–2015 and 2015–2016 malaria seasons (November–May) and tested on the 2016–2017 season, we obtained a RMSE of 56 cases per thousand people. For comparison, the mean malaria incidence per month during the high season was 60.5, ranging from 0 to 844 cases per thousand people per month. In addition, the model identified 9 out of 20 (top 10%) and 23 out of 40 (top 20%) Fokontany with the highest average malaria incidence for the 2017–2018 season, compared to an expected 2 out of 20 and 8 out of 40 if identified at random. We found a high correlation between our predictions and observed malaria incidence averaged over time (Pearson's $R = 0.69$, Spearman's $\rho = 0.77$, Fig 3C). Finally, we evaluated the performance of our model with fixed effects only. Without random effects, our model

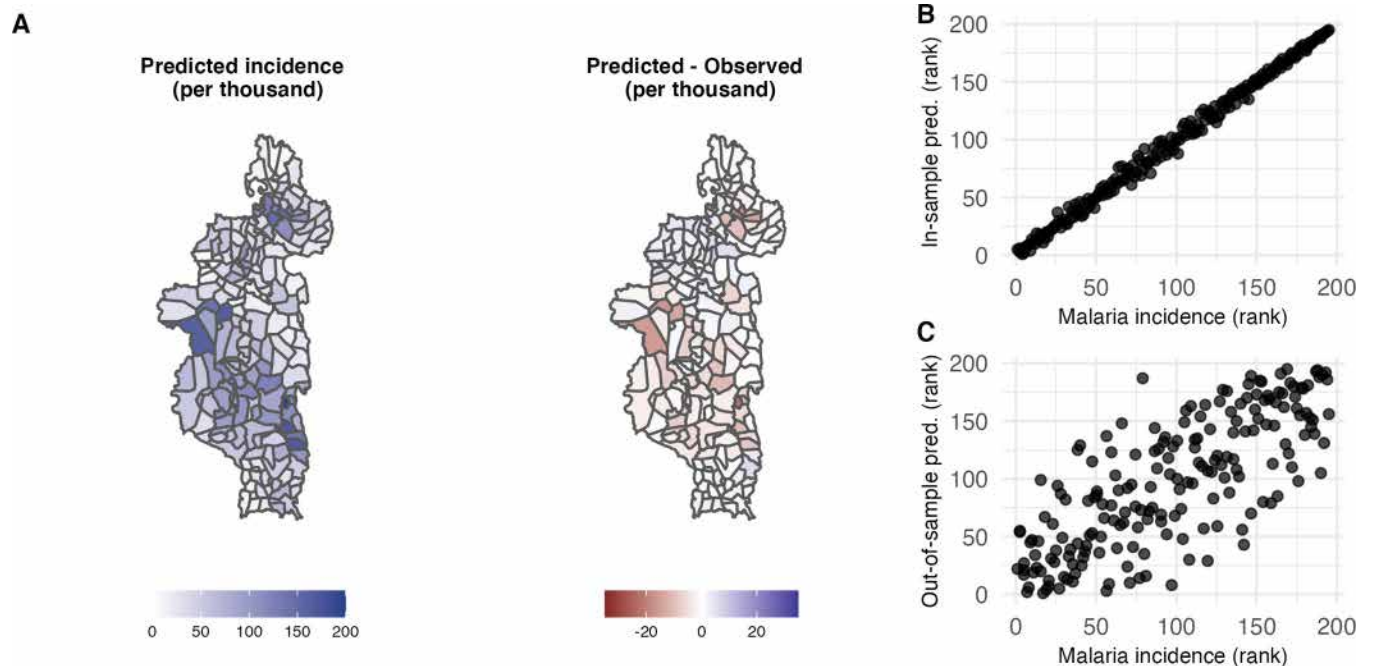


Fig 3. Predictions from GLMM with month and Fokontany ID as random effects. (A) Left: Within-sample predictions for monthly malaria incidence. Right: Difference between observed and predicted malaria incidence across space. Overpredictions are shown in blue and underpredictions in red. Fokontany boundary data is available from OCHA (<https://data.humdata.org/dataset/cod-ab-mdg>) under a CC BY 4.0 License. (B) Scatterplot of ranked within-sample malaria predictions across ranked observed malaria incidence, for data averaged across space. One point represents one Fokontany. (C) Scatterplot of ranked out-of-sample malaria predictions across ranked observed malaria incidence, for data averaged across space. One point represents one Fokontany. The model was trained on the malaria seasons from 2014 to 2016 and tested on the malaria season from November 2016 to May 2017.

<https://doi.org/10.1371/journal.pgph.0001607.g003>

evaluated across time and space had a RMSE of 67, and we found a Pearson's correlation coefficient of 0.50 and a Spearman rank correlation coefficient of 0.60 between our predictions and observed malaria incidence averaged across time. We obtained a RMSE of 59 for out-of-sample predictions for the 2016–2017 season.

Path analysis provides more insight at the spatial level

We used a path analysis to further investigate the spatial relationship between our predictors and their effect on malaria incidence (Fig 4). For this, data on both malaria and its predictors were averaged by Fokontany over the study period.

Many of the relationships that were found to be significant in the path analysis confirmed our hypotheses (see [Methods](#) section). More specifically, we found that distance to the road was negatively correlated with wealth. Wealth was also associated with a decrease in forest loss, which was itself a predictor of forest edge. Altitude was associated with lower precipitation and land surface temperature. Both land surface temperature and precipitation were positively associated with bed net use. Precipitation also had a positive effect on rice field area. Direct effects between predictors and malaria incidence were mostly consistent with our GLMM results, but fewer predictors had a significant relationship with malaria: the effect of bed nets, temperature, and forest loss on malaria incidence were not significant in the path analysis.

Discussion

Socio-ecological drivers of human malaria have been identified at global, national and subnational spatial scales [26–28], but it is currently unclear whether associations at large spatial

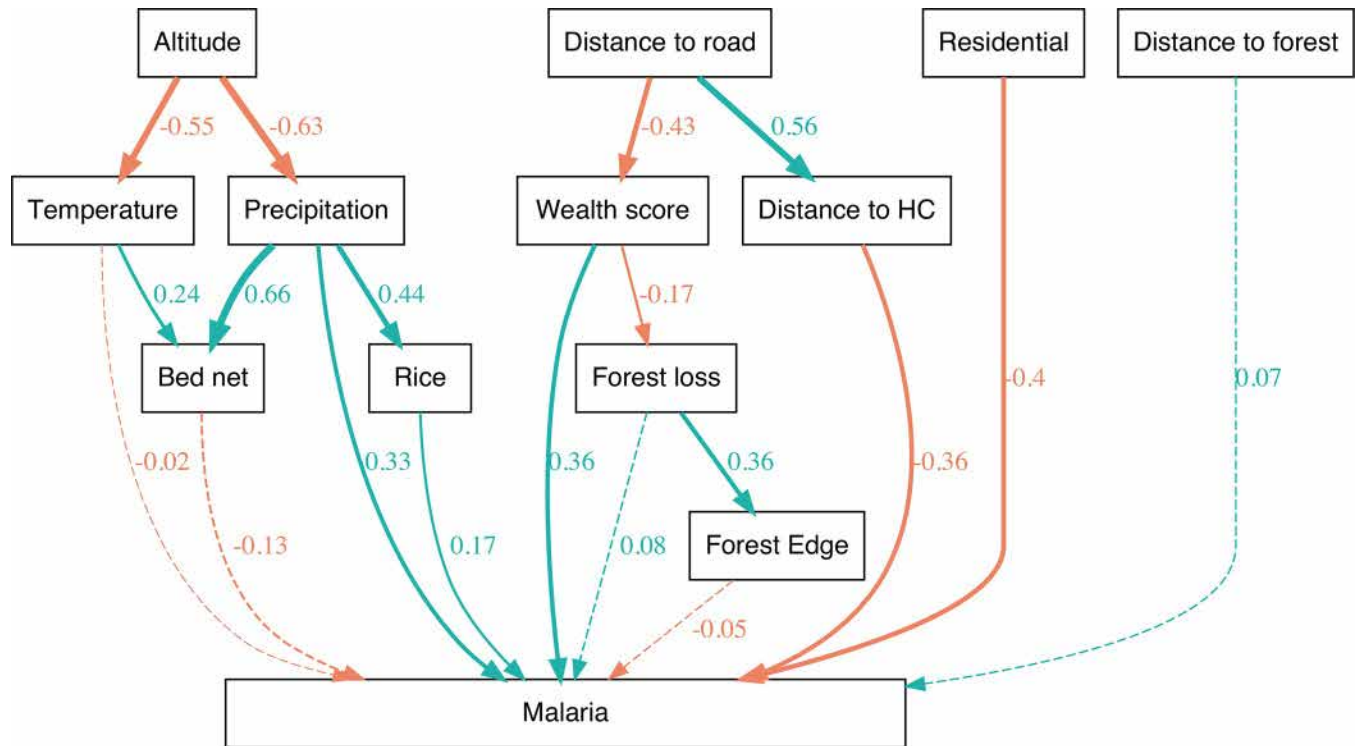


Fig 4. Model structure and standardized coefficient estimates for the path analysis. Positive and negative effects are shown with green and red arrows respectively. Arrow width is also proportional to coefficient estimates. Dashed arrows show non-significant effects.

<https://doi.org/10.1371/journal.pgph.0001607.g004>

scales hold at smaller spatial scales relevant for malaria control efforts by local authorities [14]. Here, we used the rural district of Ifanadiana, in south-eastern Madagascar, to investigate the relationship between hypothesized social and environmental determinants of malaria and malaria incidence at a very fine spatial scale. For this, we used a geographically-explicit dataset of malaria incidence at the Fokontany-level (a village or groups of villages) previously adjusted for known underreporting drivers in passive surveillance systems, and combined it with field surveys and remotely-sensed data via two separate statistical frameworks to account for direct and indirect pathways. We found that widely accepted drivers of malaria, including precipitation and temperature, were still relevant to predict incidence at fine spatial scales in our setting, in addition to identifying other land cover and socio-economic variables specific to the study region. Together, these factors allowed us to identify more than half of communities in the top quintile for malaria transmission for the 2016–2017 malaria season (compared to a 20% accuracy by chance), and explain over three quarters of the variation in malaria incidence rank.

Despite the very local scale of our study, large variations in temperature, both across space and time, occurred in Ifanadiana district. However, we only found an association between malaria incidence and temperature using a spatio-temporal GLM, and not in a spatial path analysis where data was averaged over time. This could be explained by the fact that, once averaged over time, mean land surface temperature only ranges between 22 and 28°C across Fokontany. In addition, malaria transmission is not only affected by mean temperature, but also by temperature variation, which could not be captured in the path analysis. Malaria dynamics have been shown to be strongly associated with temperature, which affects different stages of the mosquito larval development and mosquito lifespan [29,30]. It has been suggested that the effect of temperature on malaria is unimodal, with a predicted optimal temperature

for malaria transmission of 25°C for *Anopheles* mosquito species and *Plasmodium falciparum* [22]. Previously, the association between temperature and malaria outcomes has been studied at relatively large scales, where heterogeneity can be larger [8,31–33]. Our analysis supported a unimodal relationship between temperature and malaria incidence. In addition, we found that malaria incidence was more likely to be non-zero when mean temperature was higher, suggesting that malaria transmission may be cold-limited in Ifanadiana district even during the warm, high transmission season.

The relationship between malaria and precipitation is complex and more context-specific than with temperature. Precipitation creates breeding grounds for mosquitoes but can also wash away larvae [34]. Therefore, depending on the local climate and mosquito life cycle stage, there may exist positive [35], negative, or unimodal relationships [28] between precipitation and malaria. We found evidence of a positive relationship between precipitation during the previous month and malaria incidence in our spatiotemporal model, supporting the ‘breeding grounds’ hypothesis. Unlike temperature, this relationship was also supported by our spatial SEM, where it had both a direct effect on malaria and an indirect effect via its association with larger areas of rice fields. This suggests that spatial variation in precipitation was large enough to affect malaria transmission patterns, by affecting the distribution of mosquito habitats such as rice fields and smaller pools.

Overall, land use emerged as an important ecological predictor of local malaria transmission in Ifanadiana. In this analysis, we considered three factors representing land cover at the Fokontany level: proportion of residential areas, proportion of rice fields and forest loss. We did not find any relationship between forest loss in the last three years and malaria incidence. Some studies have found a positive relationship between malaria and deforestation [36,37], with intact forests supporting non-vector species of mosquitoes while deforested areas provide vector habitat through agriculture, timber extraction and mining [38]. In south-eastern Madagascar, *Anopheles* mosquito species have been found to be rare in forested areas, while being very common in agricultural areas and near livestock pens [21]. Irrigated agriculture, in particular, has been identified as a strong predictor of *Anopheles* mosquito habitat across Madagascar [39], and the proportion of rice fields was positively associated with malaria in our spatiotemporal model as well. However, an increase in malaria with deforestation is characteristic of the frontier stage, when extensive changes in land use are occurring [38,40]. Instead, the majority of Fokontany in Ifanadiana are characterized by an open fragmented landscape and limited forested areas (Fig 1B), and would not be considered a frontier environment. Finally, Fokontany with a larger proportion of residential areas had lower malaria incidence. These results suggest that malaria transmission is associated with rural areas in the district, which have higher agriculture and irrigation relative to more urbanized areas (larger towns). This is consistent with previous findings, with lower malaria prevalence in children in urban areas than in rural areas, across the sub-Saharan region [41] and Madagascar [39,42].

Ifanadiana is heterogeneous with respect to wealth and access to health care. Despite using incidence data already adjusted for underreporting due to biases in geographical and financial access to care, we still found an association with both average household wealth and distance to primary health centers. Specifically, we found that wealthier fokontany had higher incidence rates and that fokontany more distant from a primary health center had lower incidence rates. In Ifanadiana, household wealth is concentrated along major transit axes, and health centers are located in larger towns but many populations live 2–6h away. Rather than reflecting social drivers of malaria incidence, these findings are likely the result of differential access to health-care, a known phenomenon in Ifanadiana [43,44]. This reflects the fact that people who live further away from a health center and who have less economic resources are less likely to access health centers and be diagnosed with malaria [45,46].

Overall, the SEM allowed for a more nuanced understanding of local malaria transmission, revealing how socio-environmental factors affected malaria via direct and indirect pathways. For example, we found a negative relationship between the wealth index and forest loss. The relationship between poverty and deforestation is complex and may vary between regions and countries [47,48]. Our results suggest that poverty in Ifanadiana may drive people to depend on logging for income. The link between deforestation and poverty has been previously explored in other areas. For example, Jones et al. (2020) [49] found that clinic discounts were associated with a 70% reduction in deforestation in Borneo, as well as increased clinic usage. We also found higher reported bed net use was not associated with a change in malaria incidence in the spatiotemporal GLM, even though bed net use has generally been found to be an effective malaria prevention tool, including in Madagascar [45]. Results from the SEM suggest that bed net use was high in areas where climate was mediating a higher transmission risk (areas with higher temperature and precipitation), and when that was accounted for, it had a protective, although non-significant, effect on malaria spatial patterns.

While the primary purpose of this study was to investigate predictors of malaria at small spatial scales, we also explored the predictive potential of our GLM model. Despite a relatively high error (RMSE of 50 per thousand per month, compared to an average incidence of 61 per thousand per month during the high season), our model performed well when predicting the rank of each fokontany, both for within- and out-of-sample predictions. This difference in model performance can be explained by the large observed range in malaria incidence from 0 to 844 per thousand per month. The error associated with our model made it difficult to distinguish between Fokontany with small or intermediate levels of malaria incidence but allowed us to identify Fokontany with the highest malaria incidence. More specifically, we correctly identified 80% of the Fokontany above the median malaria prevalence, based on data from the two previous seasons. This would be of help to inform local intervention efforts from routine surveillance data, using relative rather than absolute metrics to identify Fokontany of concern. It is important to note that the time series used here are too short to allow modelling of longer-term temporal trends, so that predictions built from our 4-year period could be inadequate for future years due to anomalies, extreme weather events, or other factors not observed within our study period. The reliability of forward predictions could be increased as additional malaria and predictor data become available and longer time series are used. Longer time series would also allow for more explicit time series analysis. Finally, predictive power could also be increased in future work by increasing the number of predictors or by using specialized models that optimize predictive power at the expense of explanatory power, such as Support Vector Regression, Random Forest regression and other machine learning and deep learning models [50–52].

While we had access to health system, household survey and land cover data at high spatial resolution within one rural district, many of the limitations of our models were still due to either limited resolution in space or time or missing data. For example, the resolution of rainfall data was limited spatially while the resolution of detailed land use data was limited temporally. As a result, any directional relationship between malaria incidence and changes in land cover (other than deforestation) could not be explored. In addition, collinearity between some of these land cover variables decreased our ability to discern their respective impact. For this project, we used a dataset that adjusted from known biases in passive surveillance, which are difficult to correct completely [18,19]. If there were remaining biases due to factors other than those we identified and accounted for in our zero-inflated model (wealth, distance to health-care), this could affect the spatiotemporal structure of malaria incidence and could impact our conclusions. We corrected for spatial autocorrelation using a Matern covariance structure but some spatial structure remained in the residuals, suggesting that other unknown variables

correlated in space might contribute to malaria in this area. Finally, our study was undertaken in only one district of Madagascar, which limits its generalizability to other settings. Thus, methods developed here should be replicated elsewhere or expanded to larger areas in order to explore their wider applicability.

Conclusion

This paper represents one of the first attempts to explain and predict malaria over fine spatial scales from routine passive surveillance data. We found that climatic, land use and socio-economic variables can be significant predictors of malaria incidence even at such fine scales, providing key insights for a context-specific understanding of local malaria transmission. In addition, our explanatory model performed well when predicting malaria dynamics, paving the way for more predictive frameworks that could help inform local health actors with the implementation of malaria control activities. Future work should focus on improving the spatiotemporal resolution of socio-environmental drivers of malaria and developing automated data pipelines and predictive frameworks to facilitate the integration of nowcasting or forecasting systems within local program implementation in areas of high malaria transmission.

Materials and methods

Study site

Ifanadiana is a rural district located in southeastern Madagascar, and is home to approximately 200,000 people across 195 Fokontany, the smallest administrative unit comprising a village or group of villages (average size of 20 km² in Ifanadiana). The district is characterized by heterogeneous socio-economic, land use and climatic conditions, as well as malaria incidence (Fig 1A), which make it an ideal setting for this study. The dominant land class across the district is open or degraded forest (i.e. 'savannah'), with the exception of Ranomafana National Park, a heavily forested protected area to the west of the district (Fig 1B). Rice fields and residential areas are interspersed throughout the district, among the savannah. The district is also characterized by a strong altitude gradient from east (100m) to west (1100m), which influences precipitation and temperature patterns (S1 Fig).

The passive malaria surveillance system in Ifanadiana relies on reported cases from primary health care facilities, each serving a population of about 10,000 people and an average catchment area of about 200 km². The populations' wealth and health care access are influenced by their proximity to the main paved road crossing the district and to health facilities, which are located in the larger towns (Figs 1A and S1) [53]. More than 75% of the population in Ifanadiana lives more than an hour away from a primary health care facility, more than one third lives over two hours away, and over one in ten people live further than three hours away [54]. Primary care utilization decreases exponentially with increasing distance and travel time to these health facilities [44]. In 2014, the NGO Pivot partnered with the Ministry of Public Health with the goal of strengthening the health system at all levels of care and ultimately ensuring universal access to health care, including malaria diagnosis and treatment. The intervention combines improving health system readiness (infrastructure, staffing, equipment, and medicine), integrated clinical programs, and robust data collection systems to inform program implementation. This work was done in the context of current efforts to improve malaria surveillance and control activities in the district.

Data acquisition and processing

Malaria incidence data. We used a dataset of monthly malaria cases at the Fokontany level from January 2014 to December 2017, previously adjusted from passive surveillance data

in Ifanadiana [18]. Methods used to produce this dataset are described in detail by Hyde et al. [18]. Briefly, de-identified patient-level data were collected from the 19 public health centers across Ifanadiana (covering 195 Fokontany), including age, Fokontany of residence, and malaria status, which was determined using rapid diagnostic tests. Use of Ministry of Health (MoH) data for this study was authorized by the Secretary General of the MoH, by the Medical Inspector of Ifanadiana district, and by Harvard's Institutional Review board (IRB). The IHOPE cohort study was approved by the Madagascar National Ethics Committee and Harvard Medical School IRB. Population data was obtained from the Ministry of Public Health to calculate per-capita incidence estimates. Individual-level data was then aggregated to obtain monthly malaria incidence per capita at the level of the Fokontany for individuals of all ages, and for children under five years old. This dataset was finally adjusted for a variety of factors known to bias passive surveillance data, including distance to healthcare, existence of fee reimbursement systems and number of staff. For this, a benchmark multiplier was combined with a health care utilization index obtained from statistical models of non-malaria patients. Variations to the multiplier and several strategies for pooling neighboring communities together were explored to allow for fine-tuning of the final estimates. The resulting dataset was validated based on the reduction in said biases and through comparisons to areas with optimal health care access [18]. Here, we focus our analysis on the adjusted malaria incidence rates during the malaria high transmission season, which occurred from November to May (S1 Fig), to reduce the effect of missing data and false zeroes.

Climatic variables. We included two climatic variables known to influence malaria dynamics: temperature and precipitation [8,32,55]. Precipitation is an important driver of seasonal mosquito population dynamics and temperature determines larval mosquito and parasite development rates [30]. We obtained monthly land surface temperature (LST) by averaging biweekly data points from January 1st 2014 to December 31st 2017 from the thermal infrared band (Band 10) of Landsat 8 satellites [56], which has a spatial resolution of 30m (Table 1). We summed daily precipitation over each month to obtain monthly precipitation during the same period from the CHIRPS dataset, which provides rainfall estimates at a ~5km resolution from rain gauges and satellite data [57]. From these raster data and using Fokontany administrative limits, we calculated for each Fokontany total monthly precipitation, as well as spatial minimum, maximum and mean monthly LST. We excluded from the analysis anomalous LST (mean and minimum temperatures below 0°C and -5°C respectively) and capped maximum LST at 50°C. Temperature has been shown to have non-linear effects on malaria: transmission potential is a unimodal, symmetrical function of temperature with an optimum around 25°C and critical thermal thresholds below 17°C and above 34°C [22]. We thus defined a temperature suitability index by taking the squared difference between observed mean temperatures and this optimal temperature of 25°C.

Land cover variables. We included predictors related to vegetation cover, land use and land use change which influence the distribution of mosquito breeding habitat, and may influence human behavior. We used qGIS [58] and ArcGIS [59] to calculate the proportion of each land cover class in each Fokontany from a land use map (Fig 1B) developed by Ihantamalala et al. [54]. This dataset was created via semi-supervised classification of Sentinel-2 imagery with identification of training data via OpenStreetMap, as described in Ihantamalala et al. [54]. Even though these land use features were obtained in 2018, they tend to be relatively stable over the short term in a rural area such as ours, so they were assumed to be a good representation of land cover for the period 2014–2017. Water bodies mostly included rivers and were not included in this analysis as moving water is not suitable mosquito breeding habitat. In addition, forest habitat was also excluded and replaced with other land cover metrics relevant to malaria transmission, such as distance from residential areas to forest, and forest edge [60].

Table 1. Predictor variables.

Predictor		Model	Spatial scale	Temporal scale	Source
Detection	Wealth score	GLM, SEM	Fokontany	Every two years (collected in 2014 2016, 2018)	[62]
	Distance to health center	GLM, SEM	Fokontany	Fixed (collected in 2018)	[54]
	Distance to road	SEM	Fokontany	Fixed	[54]
Behavioural	Bed net use (high season)	GLM, SEM	Fokontany	Every two years (collected in 2014, 2016, 2018)	[62]
Land use	Altitude	SEM	10x10m	Fixed	
	Residential area	GLM, SEM	10x10m	Fixed (collected in 2018)	[54]
	Rice field area	GLM, SEM			
	Distance from residential areas to forest	SEM			
	Forest edge	SEM			
Forest loss	Mean loss over last 3 years	GLM	30x30m	Yearly	Global Forest Change
	Mean loss over last 10 years	SEM			
Temperature (1-month and 2-month lag)	Min	GLM	30x30m (resampled)	Monthly (mean monthly value)	USGS Landsat 8, Band 10
	Max	GLM			
	Mean	GLM, SEM			
	Mean-suitability index	GLM			
Precipitation	Total	GLM, SEM	5566x5566 m	Monthly	CHIRPS

<https://doi.org/10.1371/journal.pgph.0001607.t001>

This is because large forested areas are arbitrarily divided between different Fokontany, possibly biasing the analysis. These metrics were calculated using the original map in ArcGis.

We used data from the Global Forest Change dataset to determine the mean forest loss over the previous 3 and 10 years for each Fokontany and year during our study period [61]. The Global Forest Change dataset provides a global estimation of forest loss between 2000 and 2020 based on Landsat satellite data, including the year during which a pixel of 30m by 30m was deforested, if it was. We first used qGIS to visualize this information per year for Ifanadiana district and to calculate the proportion of deforested pixels for each year and for each Fokontany between 2004 and 2017. Finally, for each year during our study period (2014 to 2017), we estimated the average of that proportion over the last 3 and 10 years.

Socio-demographic variables. We included in our models important social and behavioral attributes of communities that can influence health-seeking behavior. Some of these (e.g., distance to health center) were already used in the initial adjustments of the malaria incidence dataset, but we included them here to account for residual uncorrected biases in malaria detection.

The distance of each Fokontany population to the nearest health center was obtained from Ihantamalala et al. [54]. Briefly, the shortest distance between each building in a Fokontany and the closest health center was calculated using the OSRM software, based on a full mapping of over 100,000 buildings and 23,000 km of footpaths previously conducted on OpenStreetMap [54]. Average distance to the nearest road was calculated by taking the average Euclidian distance to a road for every 100x100m pixel in the Fokontany.

We obtained household wealth and bed net use rates from a district-representative longitudinal cohort, which sampled 1600 households from 80 clusters across Ifanadiana every two

years (2014, 2016 and 2018), selected through a two-stage cluster sampling scheme [62]. Data was collected by the Madagascar National Institute of Statistics via household and individual surveys based on the internationally validated Demographic and Health Surveys (DHS), which include information on different indicators of health and financial well-being. From this, we estimated a wealth index using standard DHS methods [63]. Briefly, a principal component analysis of household durable assets was estimated, which included access to electricity, water and toilets, material of roofing for houses, number of residents per bedroom and cooking fuel among others. Averages per cluster were estimated for each of the cohort variables each year, and these were extrapolated from clusters using inverse distance weights, obtaining a raster for the whole district. From this interpolated raster, we then averaged them over each Fokontany and year. Finally, wealth scores were interpolated to obtain monthly estimates from survey data obtained in 2014, 2016 and 2018 (Table 1).

Data processing. Data processing was performed in R. Residential area, rice field area, wealth score, distance between residential areas and forest, forest edge, distance to health center, forest loss and precipitation were log-transformed to correct for large deviations from normality. All variables were then scaled. Maps of all raw predictor variables are available in S2 Fig.

Mixed-effect GLM

We fit a mixed-effect model with a zero-inflated negative binomial distribution to monthly malaria incidence at the Fokontany level using the glmmTMB package [64]. We used the AIC criterion to compare the fit of different distributions. In addition, we used the DHARMA package to perform residual diagnostics and confirm good model fit of the zero-inflated negative binomial model (KS test, $p = 0.5$) [65]. A zero-inflated negative binomial is used to represent count data that is over dispersed and has excess zeroes, which are assumed to be generated by two different processes. For example, zeroes in our dataset may be due to true absence of malaria incidence or to very low access to health care leading to underreporting. A more complete interpretation of the model is provided in the Results and Discussion section. We included Fokontany and month of the year as crossed random effects to account for repeated observations across time and space.

Fixed effects were determined using a two-step variable selection process, using AICc as our measure of fit. In both steps, five variables were always included based on the existing literature and exploratory analysis: bed net use, wealth score, distance to health center, residential area and rice field area. During the first step, we evaluated the model for all possible combinations of variables, listed in Table 1. In this step, we constrained models to always include all temperature and precipitation variables (either 1-month lag or 2-month lag), based on previous literature [8,22,32,55]. During the second step, we explored the influence of the three temperature variables by fixing all variables from the model with the lowest AICc in the first step and comparing models that included all possible combinations of minimum, mean, and maximum temperature. We present the model from the second step with the lowest AICc in the main manuscript. The steps described above are summarized in S2 Table. The model we present here was broadly consistent to the conditional average of the top 10% of all models explored (as determined by AICc) (S3 Table).

After variable selection, we assessed the spatial and temporal autocorrelation in the model's residuals. We tested for temporal autocorrelation using a Durbin-Watson test on residuals aggregated over space. Spatial autocorrelation was assessed with a Moran's I test on residuals aggregated over time and using the latitude and longitude of each Fokontany to make a distance matrix. We found both temporal ($DW = 0.54$, $p < 0.0001$) and spatial autocorrelation ($I = 0.10$, $p < 0.0001$) in the residuals of the model using month and Fokontany as random

effects. We then updated our model structure to account for temporal autocorrelation by including a Ornstein–Uhlenbeck (OU) covariance structure, which allows for uneven time steps (only the high malaria season is used in this model). We fit a OU covariance matrix for each commune, which was sufficient to remove temporal autocorrelation in the residuals ($DW = 1.8$, $p = 0.30$). We also updated our model structure to use a Matern covariance structure using latitude and longitude to address spatial autocorrelation. The residuals of the resulting model showed reduced but still significant spatial autocorrelation ($I = 0.23$, $p = 0.0002$). We use the full model with OU and Matern covariance structure to obtain the model coefficients presented in this work and use the model with month and Fokontany as random effects for model predictions as to not rely on past malaria incidence for predictions and because the Matern covariance structure does not improve predictions. Coefficients for the model used for predictions are presented in [S4 Table](#) and equations for both model are available in [S1 Text](#).

In our analysis, we first trained the model on the whole set of observations and then obtained in-sample predictions. For out-of-sample predictions, we trained the model on the 2014–2015 and 2015–2016 malaria seasons and tested it on the 2016–2017 season. Finally, we explored the contributions from different parts of our model (fixed vs. random, conditional vs. zero-inflated). We used the package ‘ggeffects’ to explore the marginal effects of our predictors [66]. We also fit the model with fixed effects only and compared its predictive performance to the full model.

Structural equation modeling (SEM)

We used the piecewiseSEM package in R to explore putative relationships between our predictor variables in mediating malaria incidence, by accounting for indirect pathways and relationships between the explanatory variables [67]. We conducted the SEM analysis using data aggregated across time by taking the temporal mean for each Fokontany. The set of variables with a direct path to malaria incidence was chosen *a priori* based on the results of our GLMM exercise. Further variables were included to test specific hypotheses, namely:

- Wealth and average distance to health centers are both correlated with average distance to roads (Figs [1A](#) and [S1](#)).
- Forest loss is higher in areas with lower wealth score [47].
- Forest loss is associated with an increase in forest edge [68]. We decided not to include forest area directly as forest area and forest loss may be related in complex ways (forest loss decreases forest area, but also can only occur if forest is still present).
- Areas with more rice fields have higher precipitation [69].
- Precipitation and temperature are influenced by altitude [70,71].
- Bed net use is correlated with known malaria risk factors, including higher local precipitation and temperature. We hypothesize that bed net use increases in high-risk areas, because of intervention policies and behaviour change following high malaria incidence [72].

We chose these hypotheses based on the variables available in our dataset and the causal relationship we could test considering their temporal resolution. For example, we could not test causal relationships between land cover classes and deforestation as land cover data were collected once in time. It is important to note that, although the SEM provides insights that are complementary to those from the GLM, it was not within the scope of these study to link these two models explicitly in a unified statistical framework.

Supporting information

S1 Fig. (A) Distribution of adjusted malaria incidence per Fokontany per month, averaged over four years. (B) Average precipitation and mean land surface temperature (LST) per month, averaged across Fokontany. Precipitation is shown with grey bars and mean LST is shown with a red line.

(PDF)

S2 Fig. Maps of all unscaled predictor variables considered in the GLMM. The Fokontany boundary shapefile is available from OCHA (<https://data.humdata.org/dataset/cod-ab-mdg>) under a CC BY 4.0 License.

(PDF)

S3 Fig. Distribution of observed and predicted malaria incidence across Fokontany and across time for the high malaria season (November-May). Outliers are omitted.

(PDF)

S1 Table. Exponential of GLM coefficient estimates for model with spatial and temporal covariance structures.

(DOCX)

S2 Table. Model selection summary. Note that wealth score, distance to health center, bed net use, residential area and rice field area are always included. Variables that were included as a set are shown in boxes. Variables always included in the second step are indicated with a F.

(DOCX)

S3 Table. Conditional average of the top 10% models (as determined by AICc).

(DOCX)

S4 Table. Exponential of GLMM coefficient estimates for model with spatial and temporal random effects.

(DOCX)

S1 Text. Model equations.

(DOCX)

Acknowledgments

We are grateful to everyone who contributed to the participatory mapping of Ifanadiana, especially Vincent Herbreteau, Christophe Révillion, and Jérémy Commins. We thank the staff of the local Ministry of Health team in Ifanadiana district as well as PIVOT's monitoring and community teams for their support during health system data collection. Thanks are due to Ann Miller and the teams from INSTAT for their contribution to cohort data collection. Finally, we are grateful to the reviewers for their time and comments, which greatly improved the manuscript.

Author Contributions

Conceptualization: Isabel Jones, Elizabeth Hyde, Giulio A. De Leo, Susanne H. Sokolow, Andres Garchitorena.

Data curation: Julie D. Pourtois, Elizabeth Hyde, Andrew J. Chamberlin, Felana A. Ihantamalala, Andres Garchitorena.

Formal analysis: Julie D. Pourtois, Krti Tallam, Isabel Jones, Michelle V. Evans, Susanne H. Sokolow.

Funding acquisition: Andres Garchitorena.

Investigation: Krti Tallam, Andrew J. Chamberlin.

Methodology: Julie D. Pourtois, Susanne H. Sokolow.

Resources: Laura F. Cordier, Bénédicte R. Razafinjato, Rado J. L. Rakotonanahary, Andritiana Tsirinomen'ny Aina, Patrick Soloniaina, Sahondraritera H. Raholiarimanana, Celestin Razafinjato, Matthew H. Bonds.

Supervision: Giulio A. De Leo, Susanne H. Sokolow, Andres Garchitorena.

Validation: Julie D. Pourtois.

Visualization: Julie D. Pourtois.

Writing – original draft: Julie D. Pourtois, Andres Garchitorena.

Writing – review & editing: Julie D. Pourtois, Krti Tallam, Isabel Jones, Elizabeth Hyde, Andrew J. Chamberlin, Michelle V. Evans, Felana A. Ihantamalala, Laura F. Cordier, Bénédicte R. Razafinjato, Rado J. L. Rakotonanahary, Andritiana Tsirinomen'ny Aina, Patrick Soloniaina, Sahondraritera H. Raholiarimanana, Celestin Razafinjato, Matthew H. Bonds, Giulio A. De Leo, Susanne H. Sokolow, Andres Garchitorena.

References

1. WHO. Global technical strategy for malaria 2016–2030. World Health Organization; 2015.
2. Ouma PO, Maina J, Thurairira PN, Macharia PM, Alegana VA, English M, et al. Access to emergency hospital care provided by the public sector in sub-Saharan Africa in 2015: a geocoded inventory and spatial analysis. *Lancet Glob Heal* [Internet]. 2018 Mar 1 [cited 2021 Nov 22]; 6(3):e342–50. Available from: <http://www.worldpop.org.uk>. [https://doi.org/10.1016/S2214-109X\(17\)30488-6](https://doi.org/10.1016/S2214-109X(17)30488-6) PMID: 29396220
3. Alonso PL, Tanner M. Public health challenges and prospects for malaria control and elimination [Internet]. Vol. 19, *Nature Medicine*. Nature Publishing Group; 2013 [cited 2022 Jan 3]. p. 150–5. Available from: <http://www.nature.com/explore/app>. <https://doi.org/10.1038/nm.3077> PMID: 23389615
4. Olotu A, Fegan G, Williams TN, Sasi P, Ogada E, Bauni E, et al. Defining Clinical Malaria: The Specificity and Incidence of Endpoints from Active and Passive Surveillance of Children in Rural Kenya. Snounou G, editor. *PLoS One* [Internet]. 2010 Dec 16 [cited 2022 Apr 27]; 5(12):e15569. Available from: <https://doi.org/10.1371/journal.pone.0015569> PMID: 21179571
5. Singh N, Bharti PK, Kumre NS. Active v. passive surveillance for malaria in remote tribal belt of Central India: Implications for malaria elimination. *Pathog Glob Health* [Internet]. 2016 Jul 3 [cited 2022 Apr 27]; 110(4–5):178–84. Available from: <https://pubmed.ncbi.nlm.nih.gov/27585412/>. <https://doi.org/10.1080/20477724.2016.1223920> PMID: 27585412
6. WHO. World malaria report 2013 [Internet]. 2013 [cited 2022 Apr 27]. Available from: <https://www.who.int/publications/i/item/9789241564694>.
7. Hay SI, Snow RW. The Malaria Atlas Project: Developing Global Maps of Malaria Risk. *PLoS Med* [Internet]. 2006 Dec 5 [cited 2022 Jan 3]; 3(12):e473. Available from: <https://doi.org/10.1371/journal.pmed.0030473> PMID: 17147467
8. Mabaso MLH, Craig M, Ross A, Smith T. ENVIRONMENTAL PREDICTORS OF THE SEASONALITY OF MALARIA TRANSMISSION IN AFRICA: THE CHALLENGE. *Am J Trop Med Hyg*. 2007; 76(1):33–8. PMID: 17255225
9. Adimi F, Soebiyanto RP, Safi N, Kiang R. Towards malaria risk prediction in Afghanistan using remote sensing. *Malar J*. 2010; 9(1). <https://doi.org/10.1186/1475-2875-9-125> PMID: 20465824
10. Ryan SJ, McNally A, Johnson LR, Mordecai EA, Ben-Horin T, Paaijmans K, et al. Mapping physiological suitability limits for malaria in Africa under climate change. *Vector-Borne Zoonotic Dis* [Internet]. 2015 Dec 1 [cited 2022 Jan 3]; 15(12):718–25. Available from: <https://pubmed.ncbi.nlm.nih.gov/26579951/>. <https://doi.org/10.1089/vbz.2015.1822> PMID: 26579951

11. Yamana TK, Bomblies A, Eltahir EAB. Climate change unlikely to increase malaria burden in West Africa. *Nat Clim Chang* [Internet]. 2016 Nov 1 [cited 2022 Jan 3]; 6(11):1009–13. Available from: www.nature.com/natureclimatechange.
12. Snow RW, Guerra CA, Noor AM, Myint HY, Hay SI. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* [Internet]. 2005 Mar 10 [cited 2022 Jan 3]; 434(7030):214–7. Available from: www.ncbi.nlm.nih.gov/entrez/query.fcgi. <https://doi.org/10.1038/nature03342> PMID: 15759000
13. Shiffman J, Shawar YR. Strengthening accountability of the global health metrics enterprise [Internet]. Vol. 395, *The Lancet*. Lancet Publishing Group; 2020 [cited 2021 Nov 22]. p. 1452–6. Available from: [/pmc/articles/PMC7162633/](https://pubmed.ncbi.nlm.nih.gov/32305072/). [https://doi.org/10.1016/S0140-6736\(20\)30416-5](https://doi.org/10.1016/S0140-6736(20)30416-5) PMID: 32305072
14. Rice BL, Golden CD, Randriamady HJ, Rakotomalala AANA, Vonona MA, Anjaranirina EJJ, et al. Fine-scale variation in malaria prevalence across ecological regions in Madagascar: a cross-sectional study. *BMC Public Health* [Internet]. 2021 May 29 [cited 2021 Nov 17]; 21(1):1–13. Available from: <https://link.springer.com/articles/10.1186/s12889-021-11090-3>.
15. Riedel N, Vounatsou P, Miller JM, Gosoniu L, Chizema-kawesha E, Mukonka V, et al. Geographical patterns and predictors of malaria risk in Zambia: Bayesian geostatistical modelling of the 2006 Zambia national malaria indicator survey (ZMIS). 2010;1–13. <https://doi.org/10.1186/1475-2875-9-37> PMID: 20122148
16. Baragatti M, Fournet F, Henry MC, Assi S, Ouedraogo H, Rogier C, et al. Social and environmental malaria risk factors in urban areas of Ouagadougou, Burkina Faso. *Malar J* [Internet]. 2009 Jan 13 [cited 2022 Jan 3]; 8(1):1–14. Available from: <https://link.springer.com/articles/10.1186/1475-2875-8-13>. <https://doi.org/10.1186/1475-2875-8-13> PMID: 19144144
17. Kazembe LN, Kleinschmidt I, Holtz TH, Sharp BL. Spatial analysis and mapping of malaria risk in Malawi using point-referenced prevalence of infection data. *Int J Health Geogr* [Internet]. 2006 Sep 20 [cited 2022 Jan 3]; 5(1):1–9. Available from: <https://link.springer.com/articles/10.1186/1476-072X-5-41>. <https://doi.org/10.1186/1476-072X-5-41> PMID: 16987415
18. Hyde E, Bonds MH, Ihantamalala FA, Miller AC, Cordier LF, Razafinjato B, et al. Estimating the local spatio-temporal distribution of malaria from routine health information systems in areas of low health care access and reporting. *Int J Health Geogr*. 2021 Dec 1; 20(1). <https://doi.org/10.1186/s12942-021-00262-4> PMID: 33579294
19. Rowe AK, Kachur SP, Yoon SS, Lynch M, Slutsker L, Steketee RW. Caution is required when using health facility-based data to evaluate the health impact of malaria control efforts in Africa. *Malar J* [Internet]. 2009 Dec 3 [cited 2022 Jan 4]; 8(1):209. Available from: <https://malariajournal.biomedcentral.com/articles/10.1186/1475-2875-8-209>. <https://doi.org/10.1186/1475-2875-8-209> PMID: 19728880
20. Yang GG, Kim D, Pham A, Paul CJ. A meta-regression analysis of the effectiveness of mosquito nets for malaria control: The value of long-lasting insecticide nets. *Int J Environ Res Public Health* [Internet]. 2018 Mar 19 [cited 2022 Jan 5]; 15(3). Available from: [/pmc/articles/PMC5877091/](https://pubmed.ncbi.nlm.nih.gov/29562673/). <https://doi.org/10.3390/ijerph15030546> PMID: 29562673
21. Zohdy S, Derfus K, Headrick EG, Andrianjafy MT, Wright PC, Gillespie TR. Small-scale land-use variability affects *Anopheles* spp. distribution and concomitant *Plasmodium* infection in humans and mosquito vectors in southeastern Madagascar. *Malar J* [Internet]. 2016 Dec 24 [cited 2020 Jul 29]; 15(1):114. Available from: <http://www.malariajournal.com/content/15/1/114>. <https://doi.org/10.1186/s12936-016-1164-2> PMID: 26944051
22. Mordecai EA, Paaijmans KP, Johnson LR, Balzer C, Ben-Horin T, de Moor E, et al. Optimal temperature for malaria transmission is dramatically lower than previously predicted. *Ecol Lett* [Internet]. 2013 Jan 1 [cited 2020 Nov 17]; 16(1):22–30. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1111/ele.12015>. PMID: 23050931
23. Masek JG, Wulder MA, Markham B, McCorkel J, Crawford CJ, Storey J, et al. Landsat 9: Empowering open science and applications through continuity. *Remote Sens Environ*. 2020 Oct 1; 248:111968.
24. Hay SI, Snow RW, Rogers DJ. Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. *Trans R Soc Trop Med Hyg*. 1998 Jan 1; 92(1):12–20. [https://doi.org/10.1016/S0035-9203\(98\)90936-1](https://doi.org/10.1016/S0035-9203(98)90936-1) PMID: 9692138
25. Evans M V., Bonds MH, Cordier LF, Drake JM, Ihantamalala F, Haruna J, et al. Socio-demographic, not environmental, risk factors explain fine-scale spatial patterns of diarrhoeal disease in Ifanadiana, rural Madagascar. *Proc R Soc B Biol Sci* [Internet]. 2021 Mar 10 [cited 2021 Nov 10]; 288(1946). Available from: <https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2020.2501>.
26. Weiss DJ, Lucas TCD, Nguyen M, Nandi AK, Bisanzio D, Battle KE, et al. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: a spatial and temporal modelling study. *Lancet* [Internet]. 2019 Jul 27 [cited 2022 Feb 15]; 394(10195):322–31. Available from: [https://doi.org/10.1016/S0140-6736\(19\)31097-9](https://doi.org/10.1016/S0140-6736(19)31097-9) PMID: 31229234

27. Noor AM, Kinyoki DK, Mundia CW, Kabaria CW, Mutua JW, Alegana VA, et al. The changing risk of *Plasmodium falciparum* malaria infection in Africa: 2000–10: A spatial and temporal analysis of transmission intensity. *Lancet* [Internet]. 2014 May 17 [cited 2022 Feb 15]; 383(9930):1739–47. Available from: [https://doi.org/10.1016/S0140-6736\(13\)62566-0](https://doi.org/10.1016/S0140-6736(13)62566-0) PMID: 24559537
28. Lowe R, Chirombo J, Tompkins AM. Relative importance of climatic, geographic and socio-economic determinants of malaria in Malawi. *Malar J* [Internet]. 2013 Nov 14 [cited 2022 Feb 15]; 12(1):416. Available from: <https://malariajournal.biomedcentral.com/articles/10.1186/1475-2875-12-416>. <https://doi.org/10.1186/1475-2875-12-416> PMID: 24228784
29. Mordecai EA, Caldwell JM, Grossman MK, Lippi CA, Johnson LR, Neira M, et al. Thermal biology of mosquito-borne disease. Byers J (Jeb), editor. *Ecol Lett* [Internet]. 2019 Oct 8 [cited 2022 Feb 8]; 22(10):1690–708. Available from: <https://onlinelibrary.wiley.com/doi/> <https://doi.org/10.1111/ele.13335> PMID: 31286630
30. Shapiro LLM, Whitehead SA, Thomas MB. Quantifying the effects of temperature on mosquito and parasite traits that determine the transmission potential of human malaria. *PLoS Biol* [Internet]. 2017 Oct 16 [cited 2021 Nov 10]; 15(10):e2003489. Available from: <https://doi.org/10.1371/journal.pbio.2003489> PMID: 29036170
31. Craig M, Le Sueur D, Snow B. A climate-based distribution model of malaria transmission in sub-Saharan Africa. Vol. 15, *Parasitology Today*. Elsevier Current Trends; 1999. p. 105–11. [https://doi.org/10.1016/s0169-4758\(99\)01396-4](https://doi.org/10.1016/s0169-4758(99)01396-4) PMID: 10322323
32. Okunlola OA, Oyeyemi OT. Spatio-temporal analysis of association between incidence of malaria and environmental predictors of malaria transmission in Nigeria. *Sci Rep* [Internet]. 2019;1–11. Available from: <http://dx.doi.org/10.1038/s41598-019-53814-x>.
33. Kigozi R, Zinszer K, Mpimbaza A, Sserwanga A, Kigozi SP, Kanya M. Assessing temporal associations between environmental factors and malaria morbidity at varying transmission settings in Uganda. *Malar J* [Internet]. 2016 Oct 19 [cited 2022 Feb 9]; 15(1):1–6. Available from: <https://link.springer.com/articles/10.1186/s12936-016-1549-2>. <https://doi.org/10.1186/s12936-016-1549-2> PMID: 27756304
34. Paaijmans KP, Wandago MO, Githeko AK, Takken W. Unexpected high losses of *Anopheles gambiae* larvae due to rainfall. *PLoS One* [Internet]. 2007 Nov 7 [cited 2022 Feb 15]; 2(11):1146. Available from: </pmc/articles/PMC2063461/>. <https://doi.org/10.1371/journal.pone.0001146> PMID: 17987125
35. Pascual M, Cazelles B, Bouma MJ, Chaves LF, Koelle K. Shifting patterns: Malaria dynamics and rainfall variability in an African highland. *Proc R Soc B Biol Sci* [Internet]. 2008 Jan 22 [cited 2022 Feb 15]; 275(1631):123–32. Available from: <https://royalsocietypublishing.org/>.
36. Brock PM, Fornace KM, Grigg MJ, Anstey NM, William T, Cox J, et al. Predictive analysis across spatial scales links zoonotic malaria to deforestation. *Proc R Soc B Biol Sci*. 2019; 286(1894). <https://doi.org/10.1098/rspb.2018.2351> PMID: 30963872
37. Hahn MB, Gangnon RE, Barcellos C, Asner GP, Patz JA. Influence of deforestation, logging, and fire on malaria in the Brazilian Amazon. *PLoS One*. 2014; 9(1). <https://doi.org/10.1371/journal.pone.0085725> PMID: 24404206
38. De Castro MC, Monte-Mór RL, Sawyer DO, Singer BH. Malaria risk on the Amazon frontier. *Proc Natl Acad Sci U S A*. 2006; 103(7). <https://doi.org/10.1073/pnas.0510576103> PMID: 16461902
39. Arisco NJ, Rice BL, Tantely LM, Girod R, Emile GN, Randriamady HJ, et al. Variation in *Anopheles* distribution and predictors of malaria infection risk across regions of Madagascar. *Malar J* [Internet]. 2020 Sep 29 [cited 2022 Feb 15]; 19(1):348. Available from: <https://malariajournal.biomedcentral.com/articles/10.1186/s12936-020-03423-1> PMID: 32993669
40. MacDonald AJ, Mordecai EA. Amazon deforestation drives malaria transmission, and malaria burden reduces forest clearing. *Proc Natl Acad Sci U S A*. 2019; 116(44). <https://doi.org/10.1073/pnas.1905315116> PMID: 31611369
41. Brousse O, Georganos S, Demuzere M, Dujardin S, Lennert M, Linard C, et al. Can we use local climate zones for predicting malaria prevalence across sub-Saharan African cities? *Environ Res Lett* [Internet]. 2020 Dec 1 [cited 2022 Jan 3]; 15(12):124051. Available from: <https://doi.org/10.1088/1748-9326/abc996> PMID: 35211191
42. Ihantamalala FA, Herbreteau V, Rakotoarimanana FMJ, Rakotondramanga JM, Cauchemez S, Rahoili-jaona B, et al. Estimating sources and sinks of malaria parasites in Madagascar. *Nat Commun* [Internet]. 2018 Dec 1 [cited 2022 Feb 15]; 9(1):1–8. Available from: www.nature.com/naturecommunications.
43. Garchitorena A, Miller AC, Cordier LF, Rabeza VR, Randriamanambintsoa M, Razanadrakato HTR, et al. Early changes in intervention coverage and mortality rates following the implementation of an integrated health system intervention in Madagascar. *BMJ Glob Heal*. 2018; 3(3):1–12. <https://doi.org/10.1136/bmjgh-2018-000762> PMID: 29915670

44. Garchitorena A, Ihantamalala FA, Révillion C, Cordier LF, Randriamihaja M, Razafinjato B, et al. Geographic barriers to achieving universal health coverage: evidence from rural Madagascar. *Health Policy Plan* [Internet]. 2021 Dec 1 [cited 2022 Feb 15]; 36(10):1659–70. Available from: <https://pubmed.ncbi.nlm.nih.gov/34331066/>. <https://doi.org/10.1093/heapol/czab087> PMID: 34331066
45. Kesteman T, Rafalimanantsoa SA, Razafimandimby H, Rasamimanana HH, Raharimanga V, Ramarosandratana B, et al. Multiple causes of an unexpected malaria outbreak in a high-transmission area in Madagascar. *Malar J* [Internet]. 2016 Feb 2 [cited 2022 Feb 15]; 15(1):57. Available from: <http://www.malariajournal.com/content/15/1/57>.
46. Storey JD, Babalola SO, Ricotta EE, Fox KA, Toso M, Lewicky N, et al. Associations between ideational variables and bed net use in Madagascar, Mali, and Nigeria. *BMC Public Health* [Internet]. 2018 Apr 11 [cited 2022 Feb 15]; 18(1):484. Available from: <https://bmcpubhealth.biomedcentral.com/articles/10.1186/s12889-018-5372-2>. <https://doi.org/10.1186/s12889-018-5372-2> PMID: 29642883
47. Miyamoto M. Poverty reduction saves forests sustainably: Lessons for deforestation policies. *World Dev* [Internet]. 2020 Mar 1 [cited 2021 Nov 17]; 127. Available from: <http://creativecommons.org/licenses/by/4.0/>.
48. Zwane AP. Does poverty constrain deforestation? Econometric evidence from Peru. *J Dev Econ*. 2007 Sep 1; 84(1):330–49.
49. Jones IJ, MacDonald AJ, Hopkins SR, Lund AJ, Liu ZYC, Fawzi NI, et al. Improving rural health care reduces illegal logging and conserves carbon in a tropical forest. *Proc Natl Acad Sci U S A*. 2020; 117(45).
50. Chekol BE, Hagraas H. Employing Machine Learning Techniques for the Malaria Epidemic Prediction in Ethiopia. In: 2018 10th Computer Science and Electronic Engineering Conference, CEEC 2018—Proceedings. Institute of Electrical and Electronics Engineers Inc.; 2019. p. 89–94.
51. Wang M, Wang H, Wang J, Liu H, Lu R, Duan T, et al. A novel model for malaria prediction based on ensemble algorithms. Plawiak P, editor. *PLoS One* [Internet]. 2019 Dec 26 [cited 2022 May 6]; 14(12): e0226910. Available from: <https://doi.org/10.1371/journal.pone.0226910> PMID: 31877185
52. Harvey D, Valkenburg W, Amara A. Predicting malaria epidemics in Burkina Faso with machine learning. Gadekallu TR, editor. *PLoS One* [Internet]. 2021 Jun 18 [cited 2022 May 6]; 16(6):e0253302. Available from: <https://doi.org/10.1371/journal.pone.0253302> PMID: 34143829
53. Garchitorena A, Miller AC, Cordier LF, Ramananjato R, Rabeza VR, Murray M, et al. In Madagascar, use of health care services increased when fees were removed: Lessons for universal health coverage. *Health Aff*. 2017; 36(8):1443–51. <https://doi.org/10.1377/hlthaff.2016.1419> PMID: 28784737
54. Ihantamalala F, Herbreteau V, Révillion C, Randriamihaja M, Commins J, Andreambeloston T, et al. Improving geographical accessibility modeling for operational use by local health actors. *Int J Health Geogr* [Internet]. 2020; 19(27):2020.03.09.20033100. Available from: <http://medrxiv.org/content/early/2020/03/13/2020.03.09.20033100.abstract>. <https://doi.org/10.1186/s12942-020-00220-6> PMID: 32631348
55. Beck-Johnson LM, Nelson WA, Paaijmans KP, Read A, Thomas MB, Bjørnstad ON. The importance of temperature fluctuations in understanding mosquito population dynamics and malaria risk. *R Soc Open Sci* [Internet]. 2017 Mar 8 [cited 2021 Nov 10]; 4(3). Available from: <http://dx.doi.org/10.1098/rsos.160969> *Electronics supplementary material is available online at* <https://dx.doi.org/> PMID: 28405386
56. Roy DP, Wulder MA, Loveland TR, C.E. W, Allen RG, Anderson MC, et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens Environ*. 2014 Apr 5; 145:154–72.
57. Funk CC, Peterson PJ, Landsfeld MF, Pedreros DH, Verdin JP, Rowland JD, et al. A quasi-global precipitation time series for drought monitoring [Internet]. *Data Series*. Reston, VA; 2014. Available from: <http://pubs.er.usgs.gov/publication/ds832>.
58. QGIS Association. QGIS Geographic Information System. 2020.
59. ESRI. ArcGIS Release 10.1. Redlands, CA; 2012.
60. Gao D, van den Driessche P, Cosner C. Habitat fragmentation promotes malaria persistence. *J Math Biol* [Internet]. 2019 Dec 1 [cited 2021 Nov 22]; 79(6–7):2255–80. Available from: <https://doi.org/10.1007/s00285-019-01428-2> PMID: 31520106
61. Hansen MC, Potapov P V., Moore R, Hancher M, Turubanova SA, Tyukavina A, et al. High-resolution global maps of 21st-century forest cover change. *Science* (80-) [Internet]. 2013 Nov 15 [cited 2021 Nov 22]; 342(6160):850–3. Available from: <http://earthenginepartners>. <https://doi.org/10.1126/science.1244693> PMID: 24233722
62. Miller AC, Ramananjato RH, Garchitorena A, Rabeza VR, Gikic D, Cripps A, et al. Baseline population health conditions ahead of a health system strengthening program in rural Madagascar. *Glob Health Action* [Internet]. 2017 Jan 16 [cited 2021 Nov 1]; 10(1):1329961. Available from: <https://www.tandfonline.com/doi/full/10.1080/16549716.2017.1329961>. PMID: 28621206

63. Rutstein SO, Johnson K. The DHS wealth index. DHS comparative reports no. 6. Calverton, Md ORC Macro. 2004.
64. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. {glmmTMB} Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *R J* [Internet]. 2017; 9(2):378–400. Available from: <https://journal.r-project.org/archive/2017/RJ-2017-066/index.html>.
65. Hartig F. Package ‘DHARMa’: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models [Internet]. 2021. Available from: <http://florianhartig.github.io/DHARMa/>.
66. Lüdtke D. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *J Open Source Softw*. 2018; 3(26):772.
67. Lefcheck JS. piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. *Methods Ecol Evol* [Internet]. 2016; 7(5):573–9. Available from: <http://dx.doi.org/10.1111/2041-210X.12512>.
68. Fischer R, Taubert F, Müller MS, Groeneveld J, Lehmann S, Wiegand T, et al. Accelerated forest fragmentation leads to critical increase in tropical forest edge area. *Sci Adv* [Internet]. 2021 Sep 1 [cited 2021 Nov 17]; 7(37):1DUMMMY. Available from: <https://www.science.org>. <https://doi.org/10.1126/sciadv.abg7012> PMID: 34516875
69. Sawano S, Hasegawa T, Goto S, Konghakote P, Polthanee A, Ishigooka Y, et al. Modeling the dependence of the crop calendar for rain-fed rice on precipitation in Northeast Thailand. *Paddy Water Environ* [Internet]. 2008 Jan 19 [cited 2021 Nov 17]; 6(1):83–90. Available from: <http://www.irri.org/science/ricestat/index>.
70. Barry R. Mountain weather and climate [Internet]. 1992 [cited 2022 Aug 8]. Available from: https://books.google.com/books?hl=en&lr=&id=-EcTaULIWxcC&oi=fnd&pg=PP1&ots=qjRGMAGqqZ&sig=F6h4_rpfb-AubW-CpMmQqFi8okM.
71. Körner C. The use of “altitude” in ecological research. *Trends Ecol Evol*. 2007 Nov 1; 22(11):569–74. <https://doi.org/10.1016/j.tree.2007.09.006> PMID: 17988759
72. Pinchoff J, Hamapumbu H, Kobayashi T, Simubali L, Stevenson JC, Norris DE, et al. Factors associated with sustained use of long-lasting insecticide-treated nets following a reduction in malaria transmission in southern Zambia. *Am J Trop Med Hyg* [Internet]. 2015 Nov 4 [cited 2022 Aug 8]; 93(5):954–60. Available from: <https://www.ajtmh.org/view/journals/tpmd/93/5/article-p954.xml>. <https://doi.org/10.4269/ajtmh.15-0093> PMID: 26324729