

UCSF

UC San Francisco Previously Published Works

Title

Heterogeneous exposure and hotspots for malaria vectors at three study sites in Uganda

Permalink

<https://escholarship.org/uc/item/5hg2t62q>

Authors

Kang, Su Yun
Battle, Katherine E
Gibson, Harry S
et al.

Publication Date

2018

DOI

10.12688/gatesopenres.12838.2

Peer reviewed



RESEARCH ARTICLE

REVISED **Heterogeneous exposure and hotspots for malaria vectors at three study sites in Uganda [version 2; referees: 2 approved]**

Su Yun Kang ¹, Katherine E. Battle ¹, Harry S. Gibson¹, Laura V. Cooper², Kilama Maxwell³, Moses Kanya³, Steven W. Lindsay⁴, Grant Dorsey⁵, Bryan Greenhouse⁵, Isabel Rodriguez-Barraquer⁵, Robert C. Jr. Reiner⁶, David L. Smith ^{6*}, Donal Bisanzio^{7,8*}

¹Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK

²Department of Veterinary Medicine, Cambridge University, Cambridge, UK

³Infectious Diseases Research Collaboration, Kampala, Uganda

⁴Department of Biosciences, Durham University, Durham, UK

⁵Department of Medicine, University of California, San Francisco, San Francisco, CA, USA

⁶Institute for Health Metrics & Evaluation, University of Washington, Seattle, WA, USA

⁷RTI International, Washington DC, USA

⁸Centre for Tropical Diseases, Sacro Cuore-Don Calabria Hospital, Negrar, Italy

* Equal contributors

v2 **First published:** 18 Jul 2018, 2:32 (<https://doi.org/10.12688/gatesopenres.12838.1>)

Latest published: 13 Nov 2018, 2:32 (<https://doi.org/10.12688/gatesopenres.12838.2>)

Abstract

Background: Heterogeneity in malaria transmission has household, temporal, and spatial components. These factors are relevant for improving the efficiency of malaria control by targeting heterogeneity. To quantify variation, we analyzed mosquito counts from entomological surveillance conducted at three study sites in Uganda that varied in malaria transmission intensity. Mosquito biting or exposure is a risk factor for malaria transmission.

Methods: Using a Bayesian zero-inflated negative binomial model, validated via a comprehensive simulation study, we quantified household differences in malaria vector density and examined its spatial distribution. We introduced a novel approach for identifying changes in vector abundance hotspots over time by computing the Getis-Ord statistic on ratios of household biting propensities for different scenarios. We also explored the association of household biting propensities with housing and environmental covariates.

Results: In each site, there was evidence for hot and cold spots of vector abundance, and spatial patterns associated with urbanicity, elevation, or other environmental covariates. We found some differences in the hotspots in rainy vs. dry seasons or before vs. after the application of control interventions. Housing quality explained a portion of the variation among households in mosquito counts.

Conclusion: This work provided an improved understanding of heterogeneity in malaria vector density at the three study sites in Uganda and offered a valuable opportunity for assessing whether interventions could be spatially targeted to be aimed at abundance hotspots which may increase malaria risk.

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
version 2 published 13 Nov 2018	 report	
version 1 published 18 Jul 2018	 report	 report

1 **Gillian Stresman** , London School of Hygiene and Tropical Medicine, UK

2 **Mark Janko** , Duke University, USA

Discuss this article

[Comments](#) (0)

Indoor residual spraying was shown to be a successful measure of vector control interventions in Tororo, Uganda. Cement walls, brick floors, closed eaves, screened airbricks, and tiled roofs were features of a house that had shown reduction of household biting propensity. Improvements in house quality should be recommended as a supplementary measure for malaria control reducing risk of infection.

Keywords

Heterogeneity, hotspots, housing, malaria vectors, spatial, zero-inflated negative binomial

Corresponding author: Su Yun Kang (su.kang@bdi.ox.ac.uk)

Author roles: **Kang SY:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Battle KE:** Data Curation, Writing – Review & Editing; **Gibson HS:** Data Curation, Writing – Review & Editing; **Cooper LV:** Conceptualization, Writing – Review & Editing; **Maxwell K:** Data Curation, Writing – Review & Editing; **Kamya M:** Data Curation, Writing – Review & Editing; **Lindsay SW:** Data Curation, Methodology, Writing – Review & Editing; **Dorsey G:** Data Curation, Writing – Review & Editing; **Greenhouse B:** Data Curation, Writing – Review & Editing; **Rodriguez-Barraquer I:** Data Curation, Writing – Review & Editing; **Reiner RCJ:** Writing – Review & Editing; **Smith DL:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Bisanzio D:** Conceptualization, Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Bill and Melinda Gates Foundation [OPP1159934; Su Yun Kang and David L. Smith], [OPP1152978; Katherine E. Battle], [OPP1106023; Harry S. Gibson]; Clinton Health Access Initiative [Katherine E. Battle]; National Institutes of Health & International Centers of Excellence in Malaria Research (ICMER) program [U19AI089674; David L. Smith, Kilama Maxwell, Moses Kamya, Steven W. Lindsay, Grant Dorsey, Bryan Greenhouse, Isabelle Rodriguez-Barraquer].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Kang SY *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Kang SY, Battle KE, Gibson HS *et al.* **Heterogeneous exposure and hotspots for malaria vectors at three study sites in Uganda [version 2; referees: 2 approved]** Gates Open Research 2018, 2:32 (<https://doi.org/10.12688/gatesopenres.12838.2>)

First published: 18 Jul 2018, 2:32 (<https://doi.org/10.12688/gatesopenres.12838.1>)

REVISED Amendments from Version 1

In the revised manuscript, we addressed all the comments and suggestions raised by the two reviewers. We paid specific attention to the nuances in language, a source of concern for the reviewers, to ensure that our findings were not overstated. We made clear that our results showed spatio-temporal heterogeneity of mosquito abundance and biting propensity and not malaria risk by replacing the phrase 'malaria risk' with phrases such as 'vector abundance' or 'mosquito abundance' throughout the manuscript. The Introduction section has been reframed in the context of the current literature with some changes made to improve the clarity of the background of the study. The specific changes can be found in the responses to reviewers' comments. Finally, a couple of sentences were also added to the end of the second paragraph of the Discussion to further clarify our findings.

See referee reports

Introduction

Despite recent progress in controlling *Plasmodium falciparum* transmission¹, malaria remains a significant cause of preventable death². Human malaria is transmitted by more than 70 species of mosquitoes in the genus *Anopheles*³. Because of differences in the ecology and the competence of these vectors, malaria transmission intensity is highly heterogeneous over geographies and seasons, and among villages and households⁴. Heterogeneous transmission presents both a challenge and an opportunity. Heterogeneous biting rates among people driven partly by household mosquito abundance (*i.e.*, superspreading) tend to stabilize endemic transmission, but provide an opportunity to increase efficiency of malaria control by targeting interventions at those who are bitten most^{5,6}, such as those with the poorest housing quality⁷. The role of heterogeneity has been investigated in several recent studies^{4,8,9}, including the potential for spatially targeting interventions in foci, or hotspots within foci¹⁰, but there is no current consensus about what exactly is meant by a hotspot or a focus or how to identify one.

Geographical patterns, spatial uncertainty, and seasonality in malaria endemicity have been quantified rigorously in recent studies using large aggregated databases describing malaria metrics and environmental covariates^{1,11,12}, as well as high-quality research data^{13,14}. Spatial heterogeneity, spatial dynamics, and seasonality are of great interest for spatial and seasonal targeting, which could enable tailoring interventions and coverage targets to the local context and identifying hotspots^{15,16}. While these studies capture large-scale spatial and temporal patterns, transmission is a local phenomenon, and many questions about the micro-epidemiology of malaria remain poorly quantified. Although several studies have also described fine-grain spatial patterns in malaria risk, for example 4 and 8, it has proven difficult to quantify heterogeneity in individual or household biting. This is largely due to the fact that the complex endogenous dynamics of vector populations and malaria transmission are often defined by exogenous factors such as local topography, rainfall, hydrology, humidity, temperature, house construction, and malaria control. Mosquito biting or exposure is a risk factor for malaria transmission. The prospects for targeting malaria rely on quantifying spatial and temporal heterogeneity and identifying individuals or households with greatest exposure (*i.e.*, transmission hotspots). The scan statistic is commonly

used for hotspot identification¹⁷, but the anomalies detected may be stochastic fluctuations that are neither stable nor of any importance for transmission or control. Further, scanning for hotspots does not provide any insight in terms of the drivers of heterogeneity. Accurate quantification of heterogeneity and hotspots at a fine-grain resolution therefore requires taking a different approach.

Understanding malaria transmission and quantifying the possible efficiency gains from targeting interventions require methods to accurately measure heterogeneous distribution of bites and its underlying causes. A large longitudinal study of malaria recently conducted at three study sites in Uganda provides a unique opportunity to do so¹⁸, so that it was possible to measure and study the household component of biting over seasons and over time. This involves using mosquito count data from monthly entomological surveillance conducted at 330 households between October 2011 and March 2015 for Walukuba subcounty, Jinja District and Kihhi subcounty, Kanungu District; and between October 2011 and September 2016 for Nagongera subcounty, Tororo District.

With the aid of a comprehensive simulation study, we extend the existing methodology in the literature for understanding among-household heterogeneity in malaria vector density and potentially for other mosquito-borne diseases. We evaluated heterogeneous biting propensities among households during different seasons, before and after the application of vector control interventions, and among these three sites which differed substantially in their average malaria transmission intensity. We estimated household biting propensities, which measure the average ratio of mosquitoes caught in a household compared with the population expectation, which is attributable to household characteristics such as housing structure and human hosts within the household. Studies have shown that households within a settlement exhibit spatial heterogeneity in vector distribution, where mosquito densities vary from household to household¹⁹. It has also been shown that individuals living in households located close to aquatic habitats tend to receive more mosquito bites due to the relatively higher mosquito abundance in the surroundings. This is known to be affected by wind direction²⁰ and the type of materials used to build the house²¹. Also, differences in biting attractiveness of human hosts contribute to the variability in mosquito biting among people living in the same household²². Individuals who receive the most bites are most likely to be infected and can, when infected, intensify transmission by transmitting malaria parasites to a large number of mosquitoes¹³, which will tend to stabilize transmission at low intensity but reduce overall exposure at high intensity as many of the bites are absorbed on fewer individuals⁶.

Here, we describe the household biting propensities or attractiveness by first estimating and removing seasonality effects from the household mosquito counts, leaving only the remaining household-level heterogeneity for further inference and study. Inferences drawn from the household biting propensities instead of household mosquito counts provide a clearer picture of household differences in terms of attracting mosquitoes, which are unaffected by seasonal and environmental factors. Next, using these biting propensities, we introduce a novel approach for

identifying changes in malaria hotspots over time in which we compute the Getis-Ord statistic (G_i^*) on ratios of household biting propensities for different scenarios, which reflects the relative attractiveness of households in attracting mosquitoes under different circumstances.

Methods

Data sources

Entomological surveillance was conducted in three areas of Uganda: between October 2011 and March 2015 for Walukuba subcounty, Jinja District (0°26'33.2" N, 33°13'32.3" E) and Kihhihi subcounty, Kanungu District (0°45'3.1" S, 29°42'3.6" E); and between October 2011 and September 2016 for Nagongera subcounty, Tororo District (0°46'10.6" N, 34°1'34.1" E). Details of this study, including the overall study design, study sites, a detailed description of the entomological methods, and ethical approval have been described elsewhere^{18,23}. See [Figure S1](#) to [Figure S3](#) for the maps of the three study sites ([Supplementary File 1](#)).

Jinja District is located in the east central region bordering Lake Victoria and has a population of roughly 470,000 with 63% of the population living in rural area²⁴. Jinja town was characterized historically by moderate malaria transmission intensity, but it was the site with lowest transmission in our study. Kanungu District is a rural area in south-western Uganda bordering the Democratic Republic of Congo and has a population of approximately 250,000 with 80% of the population living in villages²⁴. Malaria in Kanungu has been characterized by relatively low transmission intensity and low endemicity, but it had moderate transmission in our study. Tororo District is located in south-eastern Uganda on the Kenyan border and has a population of about 520,000 with 86% of the population living in villages²⁴. Tororo is characterized by very high malaria transmission.

CDC light traps (Model 512; John W. Hock Company, Gainesville, FL, USA) were placed in 330 randomly-selected households; 116 in Jinja, 107 in Kanungu, and 107 in Tororo. The traps were positioned one meter above the floor at the foot of the bed where a study participant slept. Traps were set on one day each month at 19:00 h and collected the following morning at 07:00h. All anophelines were identified taxonomically based on morphological criteria according to established taxonomic keys²⁵. Up to 50 anophelines per household were tested for presence of *Plasmodium falciparum* sporozoites using ELISA²⁶. The total number of anophelines trapped in Jinja per household each day ranged from 0 to 121 with a median of 0; ranged from 0 to 306 with a median of 0 in Kanungu; and ranged from 0 to 1011 with a median of 9 in Tororo.

There are typically two rainy seasons in Uganda (March to May and August to October) with annual rainfall of 1,000–1,500 mm. Residents of all study households were provided with long-lasting insecticidal bednets (LLINs) at enrollment. Over the course of the study, compliance (defined as sleeping under an LLIN the previous night at the time of each clinic visit) was over 98% for all three sites. Thus LLIN coverage within the study households (where mosquitoes were collected) was very high and consistent over time and across the three sites. In

terms of LLIN coverage in the surrounding communities, coverage levels varied across the sites and over time based on repeated cross-sectional surveys. Changes in coverage over time were primarily due to a national LLIN distribution campaign conducted in November 2013 at two sites (Jinja and Tororo) and June 2014 at the third site (Kanungu). However, in a separate time series analysis²⁷, there was no significant difference in human biting rates (mosquito abundance) before and after community LLIN distribution for all three sites. Indoor residual spraying of insecticide (IRS) was only used in Tororo, where 3 rounds of the carbamate insecticide Bendiocarb were initiated in December 2014, June 2015, and December 2015 in Nagongera.

Statistical analyses

Here we partitioned the observed variance of mosquito counts at the three Ugandan study sites among the sources of heterogeneity attributed to individual households, seasonality, and environmental noise or measurement error. Here, 'partitioning' refers to attributing proportions of the total variability to individual factors. Partitioning the total variance of a response variable into its component sources is common in ecological studies^{28,29} as it can help inform a wide variety of research and management questions^{30,31}. For instance, the variance partitioning approach developed by [31](#) for a negative binomial mixed model was a useful method for assessing the response of a variance structure to large-scale ecological changes. On the other hand, as observed in our mosquito data, overdispersion is commonly exhibited in ecological count data, and can be modeled effectively using a variety of methods, including a negative binomial distribution.

The mosquito counts consisted of a large proportion of zero counts; 70% in Jinja, 54% in Kanungu, and 21% in Tororo; and showed some degree of overdispersion. As a general rule, the presence of over 30% of zeros in the data would require a zero-inflated model. Due to the presence of excess zeros in the Ugandan mosquito count data (a common feature of malaria vector data), a range of models capable of handling excess zeros were considered, including zero-inflated Poisson, zero-inflated negative binomial (ZINB), Poisson hurdle, and negative binomial hurdle regression models. A simulation study was designed to determine the most robust and best-fitted model for quantifying household-level heterogeneity in malaria vector density. The ZINB model³² outperformed the other three models in the simulation study. The details and the discussion of the results of the simulation study are described in [Supplementary File 2](#).

A Bayesian framework was adopted to estimate model parameters, using the integrated nested Laplace approximation (INLA) approach³³ implemented via the R-INLA package. By fitting the ZINB regression models to the mosquito counts, we estimated household biting propensities (ω), seasonal signal (S), and noise (e) at each of the study sites. For mosquito counts $Y = \{y_1, y_2, \dots, y_n\}$, the ZINB distribution can be written as

$$p(Y = y_i) = \begin{cases} \pi_i + (1 - \pi_i) \left(1 + \frac{\lambda_i}{\tau}\right)^{-\tau}, & y_i = 0, \\ (1 - \pi_i) \frac{\Gamma(y_i + \tau)}{y_i! \Gamma(\tau)} \left(1 + \frac{\lambda_i}{\tau}\right)^{-\tau} \left(1 + \frac{\tau}{\lambda_i}\right)^{-y_i}, & y_i > 0, \end{cases}$$

where $\tau > 0$ is a shape parameter which quantifies the amount of overdispersion. The mean and variance of the ZINB distribution are $E(Y_i) = (1 - \pi_i)\lambda_i$ and $\text{var}(Y_i) = \lambda_i(1 - \pi_i)(1 + \pi_i\lambda_i + \lambda_i/\tau)$, respectively.

For each of the study sites, the expected count for household j on day i , λ_{ij} , was modeled on a set of explanatory variables with the aid of a log link function, while the probability π_{ij} was modeled on the same set of explanatory variables using a logit link function. The covariates of interest included household identifiers (ID) and sampling days (t). Household identifiers were treated as fixed effects in the regression model, whereas daily seasonal signal and random noise were treated as random effects. A range of Bayesian prior distributions were imposed on sampling days (t), including first- and second-order random walks and autoregressive processes of order 1 and order 2³⁴. The random noise was assumed to be independent and identically distributed (i.i.d.), i.e., Gaussian i.i.d.. For consistency and convenience, the default prior specifications in R-INLA have been chosen for each of the prior distributions. The Watanabe-Akaike information criterion³⁵ was used as the model comparison criterion in order to select the model that produced the best estimates of ω , S , and e for each of the three study sites in Uganda. Note that the estimated ω for each of the following scenarios had been scaled to have a mean of one so that they reflected relative attractiveness of household mosquito biting with respect to other households. See [Supplementary File 3](#) for the R code for estimating household biting propensities and seasonal signal using a ZINB model.

The model for estimating the overall household biting propensities (ω_{overall}) is as follows,

$$\begin{aligned} \text{logit}(\pi_{ij}) &= a_j \cdot \text{ID}_j + f(t_i) + \epsilon_{ij}, \\ \log(\lambda_{ij}) &= a_j \cdot \text{ID}_j + f(t_i) + \epsilon_{ij}. \end{aligned}$$

Here a_j quantifies the effects of household biting. The random noise, ϵ_{ij} , are assumed to be i.i.d., and t_i , the temporally structured random effects, are assigned one of the four Bayesian temporal smoothing prior distributions. Using the estimated parameters, we obtained the household biting propensities, $\omega_{\text{overall}} = \exp(a_j)$ and the seasonal signal, $S = \exp(f(t_i))$. The noise, $e = \exp(\epsilon_{ij})$, accounted for additional variation among mosquito counts that was not accounted for by the covariates or by Poisson (random) variation around the mean ωS .

We quantified differences in household biting propensities due to different seasons and vector control interventions using Kendall's coefficient of concordance (or Kendall's W)³⁶. To study the impact of different seasons on household biting, we explored biting propensities during dry (ω_{dry}) and rainy seasons (ω_{rainy}) for all three sites. For Jinja and Kanungu, we evaluated changes in transmission in the first and last half of the study by estimating biting propensities in the first half period ($\omega_{\text{1st half}}$) and the second half period ($\omega_{\text{2nd half}}$) of surveillance. This provides a rough idea of how the use of LLIN over time has impacted on mosquito abundance in Jinja and Kanungu. For Tororo, the impact of IRS on household mosquito biting was examined by computing biting propensities before ($\omega_{\text{before IRS}}$) and after ($\omega_{\text{after IRS}}$) the deployment of IRS.

Household biting propensities during dry (ω_{dry}) and rainy (ω_{rainy}) seasons, during the first half period ($\omega_{\text{1st half}}$) and the second half period ($\omega_{\text{2nd half}}$) of surveillance, and before the deployment of IRS ($\omega_{\text{before IRS}}$) and after the deployment of IRS ($\omega_{\text{after IRS}}$) were estimated using one of the following sets of equations,

$$\begin{aligned} \text{logit}(\pi_{ij}) &= f(\text{ID}_{\text{dry}}) + f(\text{ID}_{\text{rainy}}) + f(t_i), \\ \log(\lambda_{ij}) &= f(\text{ID}_{\text{dry}}) + f(\text{ID}_{\text{rainy}}) + f(t_i), \quad \text{or} \\ \text{logit}(\pi_{ij}) &= f(\text{ID}_{\text{1st half}}) + f(\text{ID}_{\text{2nd half}}) + f(t_i), \\ \log(\lambda_{ij}) &= f(\text{ID}_{\text{1st half}}) + f(\text{ID}_{\text{2nd half}}) + f(t_i), \quad \text{or} \\ \text{logit}(\pi_{ij}) &= f(\text{ID}_{\text{before IRS}}) + f(\text{ID}_{\text{after IRS}}) + f(t_i), \\ \log(\lambda_{ij}) &= f(\text{ID}_{\text{before IRS}}) + f(\text{ID}_{\text{after IRS}}) + f(t_i). \end{aligned}$$

Here, ID_{dry} denote household IDs during the dry season, ID_{rainy} denote household IDs during the rainy season, and so forth. The random effects $f(\text{ID}...)$ was assumed to be i.i.d.. By modeling the biting propensities of different scenarios as random effects, it allowed for estimation and comparison of two scenarios within one model, instead of having to split the datasets into two subsets. The noise term, ϵ_{ij} , was not required in the model because the focus here was to estimate household-level effects. Hence, the biting propensities for different scenarios are such that: $\omega_{\text{dry}} = \exp(f(\text{ID}_{\text{dry}}))$; $\omega_{\text{rainy}} = \exp(f(\text{ID}_{\text{rainy}}))$; $\omega_{\text{1st half}} = \exp(f(\text{ID}_{\text{1st half}}))$; $\omega_{\text{2nd half}} = \exp(f(\text{ID}_{\text{2nd half}}))$; $\omega_{\text{before IRS}} = \exp(f(\text{ID}_{\text{before IRS}}))$; and $\omega_{\text{after IRS}} = \exp(f(\text{ID}_{\text{after IRS}}))$.

We also conducted hotspot analysis on ω using the Getis-Ord (G_i^*) statistic³⁷ for all three sites to identify hotspots of vector abundance. The G_i^* statistic is a local statistic that identifies variation across the study area, by focusing on individual features and their relationships to nearby features. It also identifies statistically significant hot spots and cold spots, given a set of weighted features. To be statistically significant, the hot spot or cold spot will have a high or low value and be surrounded by other features with high or low values. The value of the target feature is included in analysis. The G_i^* statistic is a z-score and given as:

$$G_i^* = \frac{\sum_{j=1}^n k_{i,j} \omega_j - \bar{W} \sum_{j=1}^n k_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n k_{i,j}^2 - \left(\sum_{j=1}^n k_{i,j}\right)^2}{n-1}}}$$

where ω_j is the attribute value for feature j (biting propensities in this case), $k_{i,j}$ is the spatial weight between feature i and j , n is the total number of features while:

$$\bar{W} = \frac{\sum_{j=1}^n \omega_j}{n} \quad \text{and} \quad S = \sqrt{\frac{\sum_{j=1}^n \omega_j^2}{n} - (\bar{W})^2}$$

The Getis-Ord test was performed using the `localG` function in the `spdep` package in R³⁸. At a significance level of 0.001, a

z-score would have to be less than -3.29 or greater than 3.29 to be statistically significant.

High resolution climatic and environmental covariates ($100\text{m} \times 100\text{m}$) known to interact with mosquito density, along with housing covariates for all the sampled households at the three study sites were assembled (Table 1). The housing covariates (roof type, wall type, floor type, eaves, air bricks, and house type) were all available as categorical variables. To explore the association of household biting with housing covariates, a linear regression model was fitted between ω and the housing covariates of various study sites and scenarios, which resulted in multiple correlation coefficient for each pair. For each of the environmental covariates, we calculated Pearson correlation coefficients for the covariate and ω of various scenarios to understand the association of a covariate with household biting propensities. The associations of household biting propensities with environmental and housing covariates were presented on a correlation heatmap.

For each study site and for all three study sites combined, a negative binomial generalized linear regression was used to model the relationship between the six categorical housing covariates and the total number of *Anopheles* mosquito caught per household by light trap catches, with the number of sampling nights included as an offset term in the model. The analyses were performed using the `glm.nb` function in the MASS package in R (R code available in Supplementary File 4). This allowed us to evaluate the overall effect of the various housing covariates at all study sites as well as their effects separately at each study site. In a separate analysis to assess the impact of environmental

factors on mosquito abundance, a Bayesian ZINB geostatistical spatio-temporal framework, along with environmental covariates, were used to predict mosquito density across the whole study site (details in 39).

Results

Quantifying heterogeneous household biting in mosquito abundance

Using pseudo-datasets that mimicked the Ugandan data (i.e., presence of excess zeros and Poisson distributed) and several model selection criteria in the simulation study, we found the ZINB model to be the most robust and best-fitted method for estimating household biting propensities (ω), seasonality (S), and environmental and measurement noise (e) for each study site (Figure 1).

The overall household biting propensities for the entire duration of surveillance were estimated for each of the study sites, denoted ω_{overall} . We found patterns among households as well as spatial patterns at a larger scale. Although household biting propensities are closely correlated (by design) with the average number of mosquitoes caught in a light trap per household per night at each study site, they are not identical (Figure 2). Households receiving a larger amount of mosquito bites tend to have a larger biting propensity but this is not always true due to other factors such as seasonality and the environment. We also note that there were differences in the distributions of biting propensities across sites. A large variation in biting propensities was observed in households across the three study sites, as shown in panel (a) of Figure 3, Figure 4, and Figure 5 for Jinja, Kanungu, and Tororo, respectively. Households with the largest

Table 1. Description of environmental and housing covariates.

Covariate	Description	Source
LST_day	Day time land surface temperature	MODIS derivative ⁴⁰
LST_delta	Diurnal difference in land surface temperature	MODIS derivative
LST_night	Night time land surface temperature	MODIS derivative
Precipitation	Climate Hazards Group InfraRed Precipitation with Station data	CHIRPS ⁴¹
EVI	Enhanced vegetation index	MODIS derivative
GUF	Global urban footprint: global map of human settlements	German Aerospace Center (DLR) ⁴²
NDVI	Normalized difference vegetation index	Landsat 8 composite ⁴³
Population	Estimated population per $100\text{m} \times 100\text{m}$ pixel	WorldPop ⁴⁴
TCB	Tasseled Cap Brightness; measure of land reflectance	MODIS derivative
TCG	Tasseled Cap Greenness	MODIS derivative
TCW	Tasseled Cap Wetness	MODIS derivative
Elevation	Elevation	MODIS derivative
DistToWater	GIS derived surface that measures distance to permanent and semi-permanent water based on presence of lakes, wetlands, rivers, streams and accounting for slope and precipitation	MAP (from WWF surfaces) ⁴⁵
Roof type	Tiles; metal; or thatch	Uganda study ²³
Wall type	Cement; mud; wood; or metal	Uganda study
Floor type	Earth and dung; earth or sand; cement/concrete; bricks; stones; or parquet or polished wood	Uganda study
Eaves	Open; or closed	Uganda study
Air bricks	No airbricks; unscreened airbricks; or screened airbricks	Uganda study
House type	Traditional; or others (cement/wood/metal wall, tiled/metal roof, closed eaves)	Uganda study

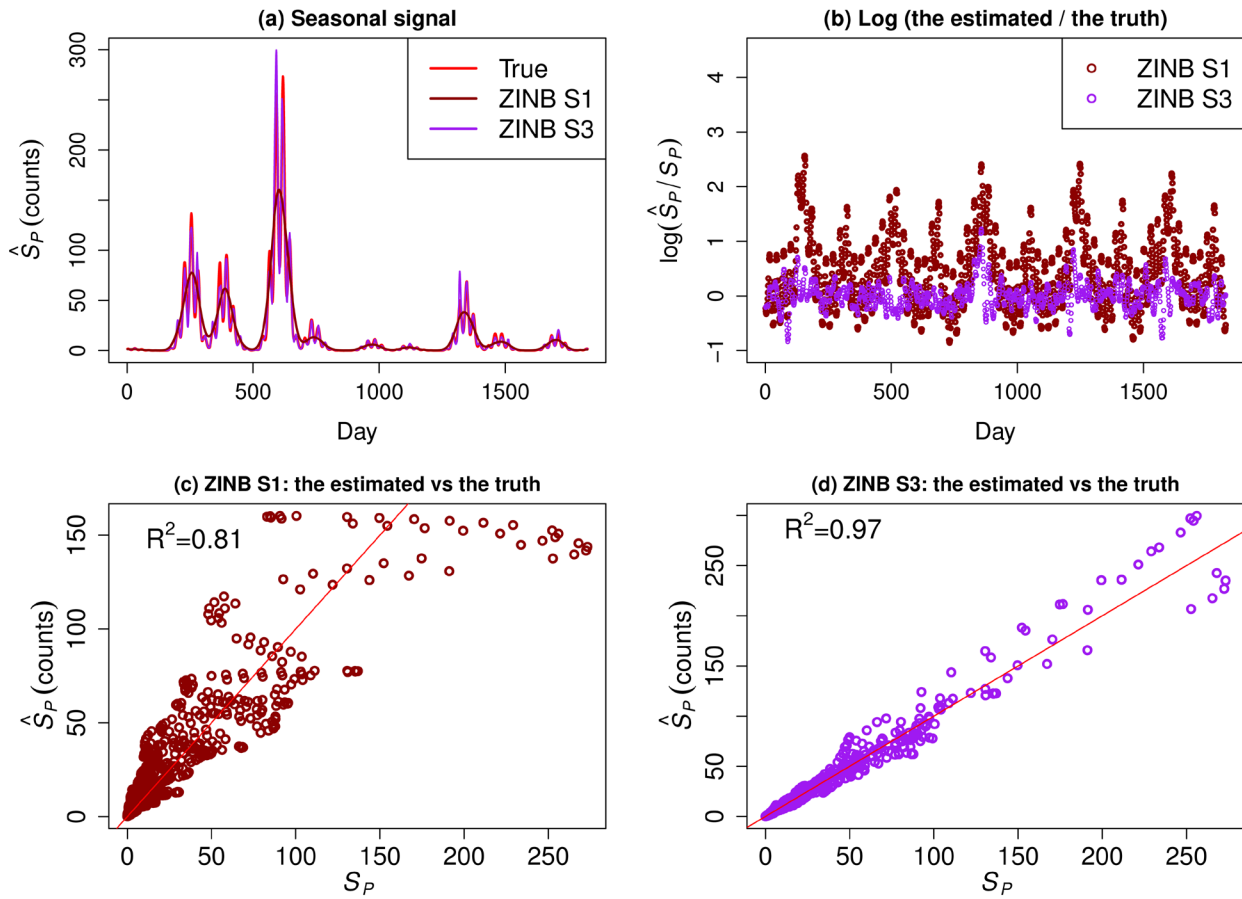


Figure 1. Simulation study: **(a)** The estimates of seasonal signal (\hat{S}_P) for pseudo-dataset D3 reconstructed using the best-fitted model, i.e. the zero-inflated negative binomial model. Smoothing technique S1 (a Gaussian kernel was used to smooth the counts prior to model fitting) produced the worst fit while smoothing technique S3 (a second-order random walk prior distribution imposed on temporal random effects) produced the best fit. **(b)** The scatter plot shows $\log(\hat{S}_P/S_P)$: ZINB S3 resulted in a much better fit of \hat{S}_P than ZINB S1. **(c)** The ZINB S1 model produced the worst fit when \hat{S}_P (estimated) is fitted against S_P (the truth) on a simple linear regression; coefficient of determination, R^2 , is around 80%. **(d)** The ZINB S3 model produced the best fit when \hat{S}_P is fitted against S_P on a simple linear regression; R^2 is close to 1.

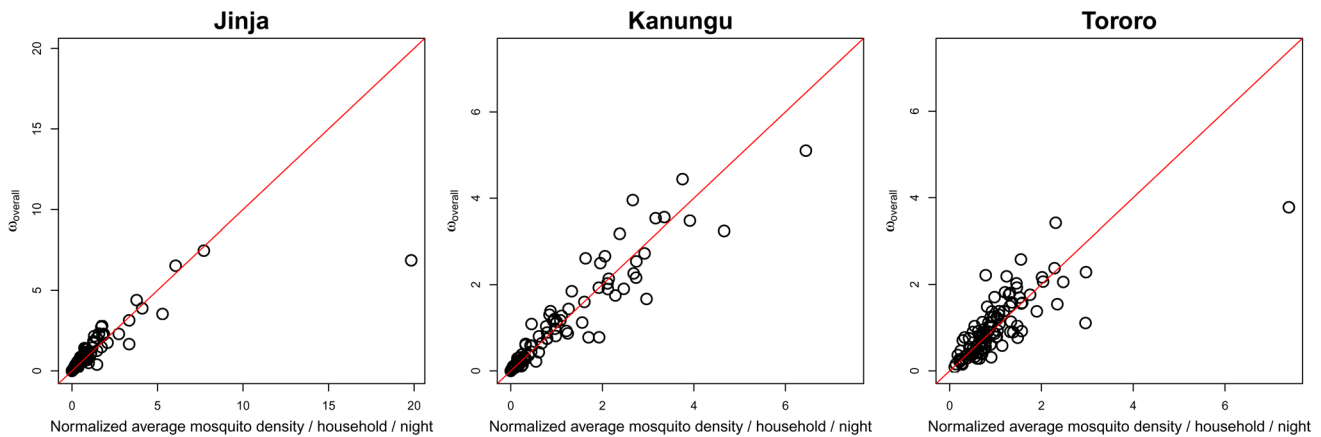


Figure 2. Household biting propensities for the entire duration of surveillance (ω_{overall}) are plotted against the normalized average mosquito density per household per night for each of the study sites. The red line denotes the 1:1 line. Evidently, ω_{overall} had a linear relationship with mosquito density at each study site.

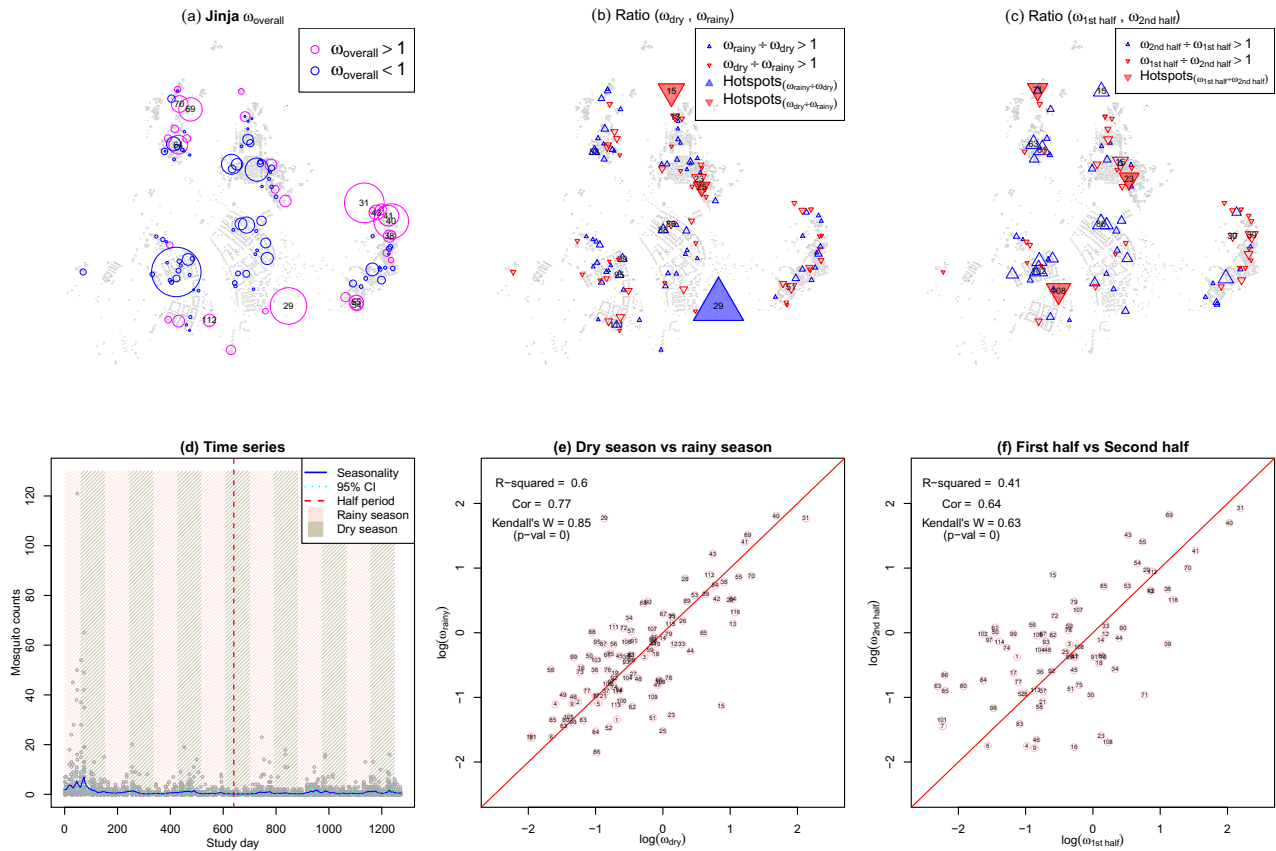


Figure 3. Jinja **(a)**: Household biting propensities for the entire duration of surveillance. Pink circles denote households with $\omega_{\text{overall}} > 1$ with the size of the circles showing the magnitude of ω_{overall} ; blue circles denote households with $\omega_{\text{overall}} < 1$ with the size of the circles showing the magnitude of $1/(3 \times \omega_{\text{overall}})$ to make their sizes comparable to the pink circles; 10% of households with the largest ω_{overall} are labelled. The largest pink circle indicates a household with the largest ω_{overall} while the largest blue circle indicates a household with the smallest ω_{overall} . Households with the largest ω_{overall} include HH31, HH29, and HH40. Grey dots in the background denote all enumerated households at Walukuba subcounty, Jinja District. **(b)** Ratios of household biting propensities during dry (ω_{dry}) and rainy seasons (ω_{rainy}). Each plotted upward pointing blue triangle represents a household with a larger ω during rainy season compared to dry season; a blue filled triangle denotes a hotspot for the ratio of $\omega_{\text{rainy}}/\omega_{\text{dry}}$. Each plotted downward pointing red triangle represents a household with a larger ω during dry season compared to rainy season; a red filled triangle denotes a hotspot for the ratio of $\omega_{\text{dry}}/\omega_{\text{rainy}}$. **(c)** Ratios of household biting propensities in the first half period of surveillance ($\omega_{\text{1st half}}$) and the second half period ($\omega_{\text{2nd half}}$). Each plotted upward pointing blue triangle represents a household with a larger ω in the second half period of surveillance compared to the first half period; there were no hotspots for the ratio of $\omega_{\text{2nd half}}/\omega_{\text{1st half}}$. Each plotted downward pointing red triangle represents a household with a larger ω during the first half period of surveillance compared to the second half period; a red filled triangle denotes a hotspot for the ratio of $\omega_{\text{1st half}}/\omega_{\text{2nd half}}$. **(d)** Time series of mosquito count data for the entire duration of surveillance, where each grey circle denotes an observation of a household. The blue solid line denotes the estimated seasonal signal and cyan dashed lines denote the 95% Bayesian credible interval for the seasonal signal. The red vertical dashed line denotes the cut-off for the first half and the second half periods. Rainy season is highlighted in pink and dry season is highlighted in light green. A weak seasonal signal was observed in Jinja. **(e)** Scatter plot of $\log(\omega_{\text{dry}})$ and $\log(\omega_{\text{rainy}})$ along with measures of R^2 , correlation, and Kendall's W . For ease of interpretation, we plotted the ω on the logarithmic scale. The log transformation also preserved the order of the observations while making outliers less extreme. **(f)** Scatter plot of $\log(\omega_{\text{1st half}})$ and $\log(\omega_{\text{2nd half}})$ along with measures of R^2 , correlation, and Kendall's W .

and smallest ω_{overall} in Jinja appeared to be located away from the centre of the study region (Figure 3(a)), which is highly urbanized⁸. Households at low elevations in Kanungu (northern part of the study region) generally had a larger ω_{overall} while households at high elevations and less rural part of Kanungu (southeast part of the study region) had a much smaller ω_{overall} (Figure 4(a)). In Tororo, some of the households with the largest and smallest ω_{overall} were found to be located around the border of the study region (Figure 5(a)).

Seasonally varying hotspots of mosquito abundance

In Figure S4 (Supplementary File 1), we plotted ω for different scenarios against ω_{overall} , changes of transmission over time was confirmed in Jinja while the application of IRS in Tororo had affected the ω . Using the G_i^* statistic, we found seasonal varying hotspots of mosquito abundance at the household level for different scenarios at the three sites. Here, a seasonal varying hotspot refers to a household (HH) with drastic differences in ω under different scenarios and was surrounded by other

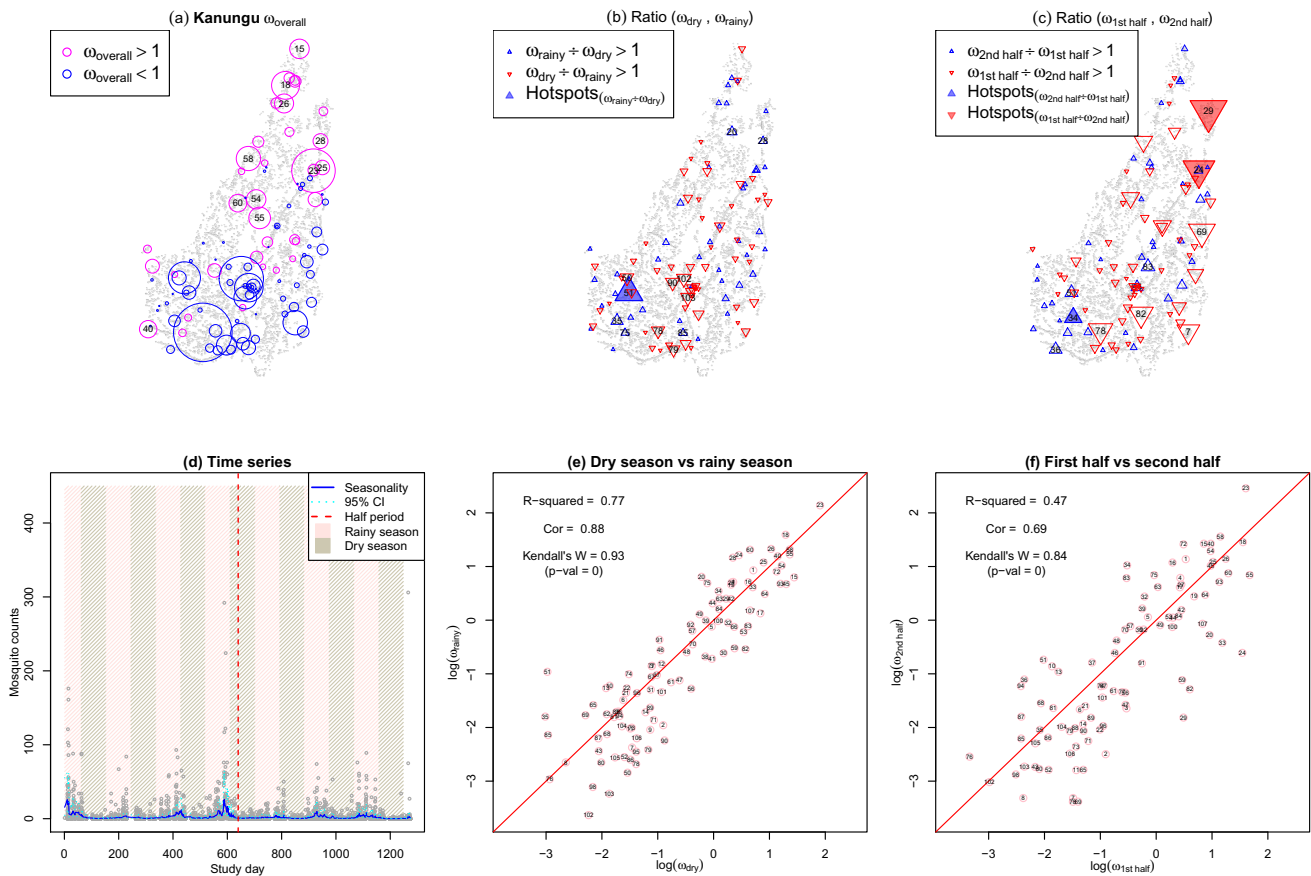


Figure 4. Kanungu. The layout of the figure is the same as in Figure 3 with some exceptions as follows. **(a):** Blue circles denote households with $\omega_{\text{overall}} < 1$ with the size of the circles showing the magnitude of $1/(5 \times \omega_{\text{overall}})$ to make their sizes comparable to the pink circles. Households with the largest ω_{overall} include HH23, HH18, and HH58. **(d)** A moderate seasonal signal was observed in Kanungu.

households with less drastic changes in ω . In Jinja, three households showed considerable differences during dry and rainy seasons while three households were identified as seasonal varying hotspots when compared in the first and last half of the study (Figure 3(b–c)). Being the study site with the lowest malaria transmission intensity, Jinja exhibited a weak seasonal signal in mosquito counts over time (Figure 3(d)). It is noteworthy that HH29, located next to a swampy area near Lake Victoria, had a much larger ω during the rainy season compared to the dry season. Most households in Jinja behaved similarly during dry and rainy seasons, whereas greater differences in ω were shown between the two time periods (Figure 3(e–f)).

In Kanungu, there was a seasonal varying hotspot for the dry vs. rainy seasons comparison while three households were identified as seasonal varying hotspots when compared in the first and second half of the study (Figure 4(b–c)). The intermediate transmission intensity found in Kanungu in previous studies^{23,46} could be explained by the moderate seasonal signal in mosquito counts over time (Figure 4(d)). Most households in Kanungu behaved very similarly during different seasons and time periods despite some slight differences in the first half and the second half periods of surveillance (Figure 4(e–f)).

In Tororo, a seasonal varying hotspot was identified for the dry vs. rainy seasons comparison while six households were identified as seasonal varying hotspots before the deployment of IRS but not after IRS (Figure 5(b–c)). Tororo, the study site with highest transmission intensity (add REF, please), exhibited a strong seasonal signal in mosquito counts over time (Figure 5(d)). It was clear that mosquito counts peaked at the end of the rainy season and at the very beginning of the dry season and reduced remarkably after the deployment of IRS. Most households in Tororo behaved very similarly during dry and rainy seasons but showed large differences in ω before the deployment of IRS and after (Figure 5(e–f)).

Association of heterogeneous household biting with environmental and housing covariates

The environmental covariates did not have strong associations with household biting heterogeneity at all three sites (Figure 6). Nevertheless, households in Kanungu had greater associations of biting propensities with environmental covariates including land surface temperature (day time, night time, and diurnal difference), precipitation, and elevation, compared to two other sites. Kanungu was a rural area of rolling hills in western Uganda located at an elevation of 1,310 m above sea

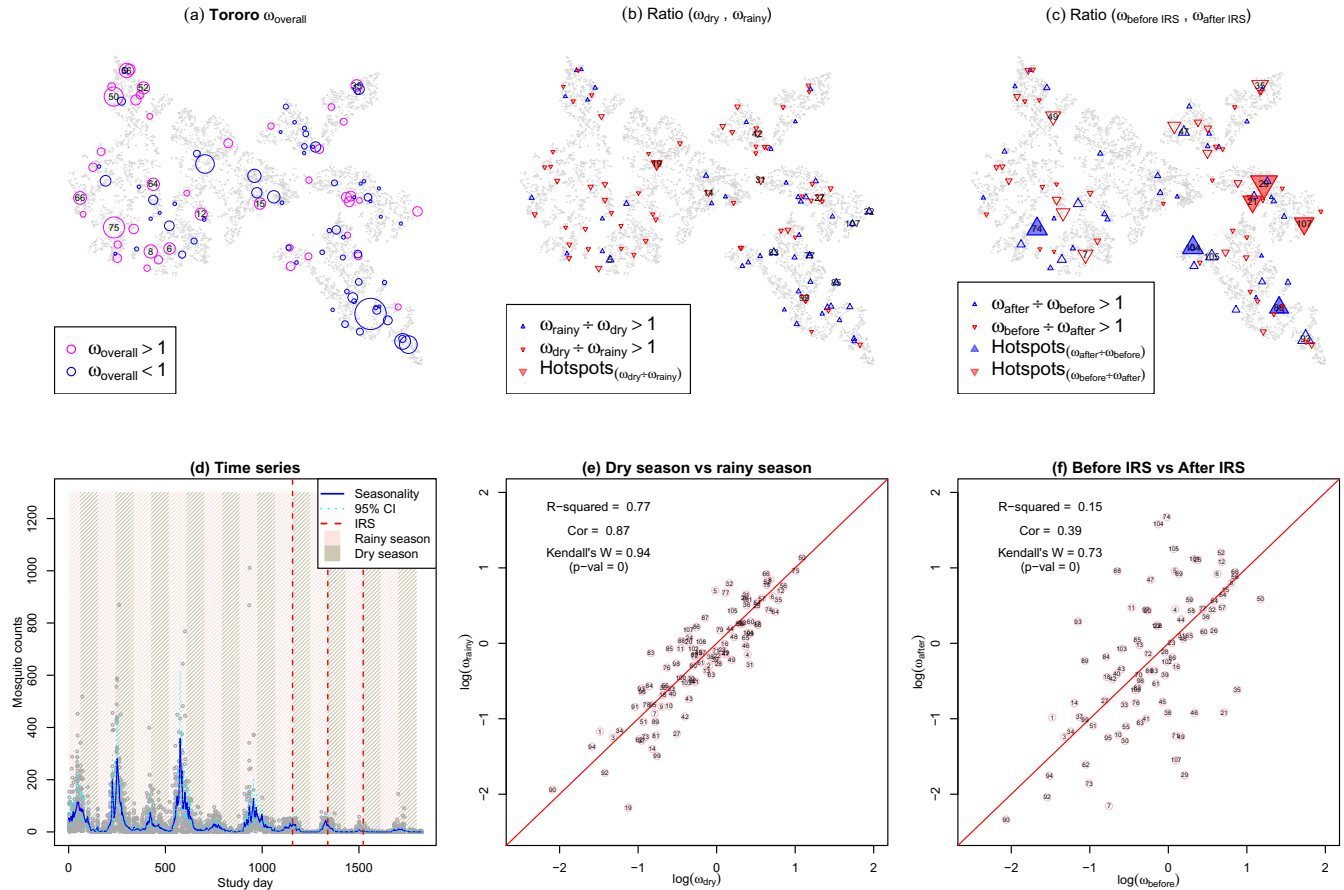


Figure 5. Tororo. The layout of the figure is the same as in [Figure 3](#) with some exceptions as follows. **(a):** Blue circles denote households with $\omega_{overall} < 1$ with the size of the circles showing the magnitude of $1/(2 \times \omega_{overall})$ to make their sizes comparable to the pink circles. Households with the largest $\omega_{overall}$ include HH75, HH50, HH56, and HH52. **(c)** Ratios of household biting propensities before the deployment of IRS ($\omega_{before\ IRS}$) and after the deployment of IRS ($\omega_{after\ IRS}$). Each plotted upward pointing blue triangle represents a household with a larger ω after the deployment of IRS compared to before the deployment of IRS; a blue filled triangle denotes a hotspot for the ratio of $\omega_{after\ IRS}/\omega_{before\ IRS}$. Each plotted downward pointing red triangle represents a household with a larger ω before the deployment of IRS compared to after the deployment of IRS; a red filled triangle denotes a hotspot for the ratio of $\omega_{before\ IRS}/\omega_{after\ IRS}$. **(d)** The red vertical dashed lines denote the deployment of IRS which were six months apart. A strong seasonal signal was observed in Tororo. **(f)** Scatter plot of $\log(\omega_{before\ IRS})$ and $\log(\omega_{after\ IRS})$ along with measures of R^2 , correlation, and Kendall's W .

level and it was the study site with the highest elevation and greatest within-site elevation change among the three sites²³. The greater variation in environmental characteristics across households in Kanungu due to its altitude gradient most likely contributed to the greater association between biting propensities and the environmental covariates.

[Figure S5 \(Supplementary File 1\)](#) shows the density of *Anopheles* mosquitoes captured by the light traps in Tororo before and after the deployment of IRS. [Figure S6 \(Supplementary File 1\)](#) shows the average predicted *Anopheles* mosquito (*An. gambiae* (*s.l.*) and *An. funestus* (*s.l.*)) density per household per night across the whole study region before and after the deployment of IRS, predicted using environmental covariates. Clearly, the environmental covariates did not account for a big portion of the heterogeneity in mosquito density, especially in the density of *An. gambiae* (*s.l.*) before the deployment of IRS (upper left panel of [Figure S6](#)).

A portion of the heterogeneity in mosquito biting may be attributed to household attractiveness caused by housing characteristics or structures, consistent with those reported previously²¹.

Among the three sites, the housing covariates had the strongest associations with ω in Tororo, followed by ω in Jinja, and ω in Kanungu ([Figure 6](#)). [Table 2](#) summarizes the incidence rate ratios of the best fitted model (retaining significant risk factors only) selected by the Akaike information criterion. In Jinja, houses with metal walls had a high biting propensity score compared to houses with cement walls; houses with earth or sand floors were also much more likely to attract mosquitoes compared to other types of floors. In Kanungu and Tororo, houses with mud floors had a higher mosquito abundance compared to houses with cement floors; houses with open eaves also attracted more mosquitoes compared to houses with closed eaves. On

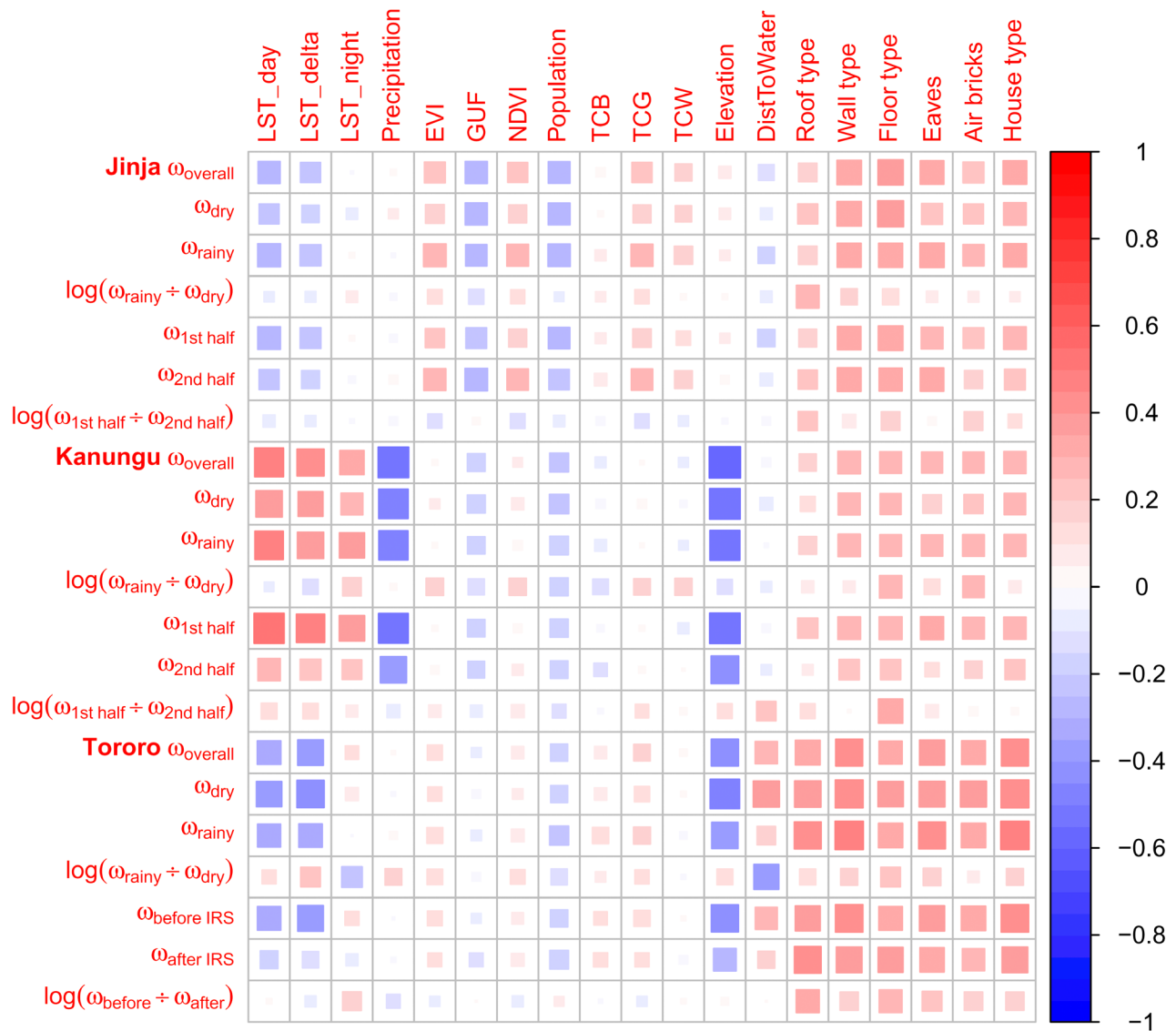


Figure 6. Correlation heatmap of household biting propensities and environmental and housing covariates. See Table 1 for descriptions of the covariates. The size of the squares and the shade of colors both illustrate the magnitude of the correlation. Red-shaded squares denote positive correlation whereas blue-shaded squares denote negative correlation.

the other hand, houses with screened airbricks in Tororo had a protective effect reducing the number of indoor mosquitoes which could result in a reduced malaria transmission. For all study sites combined, houses with wood walls, brick floors, screened airbricks, and metal or tiled roofs were protective against malaria vector. These findings are consistent with those reported in 21, but we found additionally that screened airbricks reduce number of indoor mosquitoes for all three sites combined.

Discussion

We have shown that it is possible to quantify heterogeneity in malaria vector density and its component parts: seasonality and household biting propensities. The proposed method

was successfully validated via an extensive simulation study. Household biting propensities, seasonality, and environmental factors could be quantified using a Bayesian ZINB regression model along with a temporal smoothing prior distribution for estimating seasonal signal. Using these biting propensities, we were able to identify the factors accounting for differences in mosquito abundance at three study sites in Uganda. Despite the variance introduced by seasonality, biting propensities tended to be strongly correlated with the average counts within a site, though there were some notable outliers.

While many studies of malaria transmission have focused on quantifying spatial patterns or seasonality, substantially less

Table 2. Association between household characteristics and the anophelines collected per household per night (total anophelines caught/total nights of collection) at three sites in Uganda, estimated with negative binomial generalized linear regression models. The best fitted model had identified four risk factors of household biting for all sites combined (wall type, floor type, airbricks, and roof type); two risk factors for Jinja (wall type and floor type); two risk factors for Kanungu (wall type and eaves); and three risk factors in Tororo (wall type, eaves, and airbricks). Note that interpretations of the IRRs are made for covariates significant at the 95% significance level (highlighted in boldface and labelled with an *), i.e., p-value <0.05. (IRR: incidence rate ratio; CI: confidence interval).

Characteristics	All sites combined		Jinja		Kanungu		Tororo	
	IRR (95% CI)	p	IRR (95% CI)	p	IRR (95% CI)	p	IRR (95% CI)	p
Wall type	Cement	1.00	-	-	-	-	1	-
	Mud	1.70 (0.99-2.92)	0.069	1.69 (0.80-3.58)	0.187	1.99 (1.16-3.36)	0.010*	1.96 (1.32-2.93)
	Wood	0.18 (0.07-0.58)	0.001*	2.27 (0.92-6.31)	0.067	-	-	-
	Metal	0.60 (0.20-2.31)	0.377	4.23 (1.65-12.42)	0.003*	-	-	-
Floor type	Earth or sand	1	-	1	-	-	-	-
	Earth and dung	2.23 (1.49-3.39)	<0.001*	0.94 (0.47-2.00)	0.865	-	-	-
	Parquet or polished wood	0.37 (0.04-44.68)	0.509	0.19 (0.03-4.80)	0.160	-	-	-
	Bricks	0.07 (0.01-1.28)	0.013*	-	-	-	-	-
	Cement/concrete	0.74 (0.45-1.24)	0.311	0.42 (0.20-0.86)	0.022*	-	-	-
	Stones	0.46 (0.09-7.73)	0.443	-	-	-	-	-
Eaves	Open	-	-	-	-	1	1	1
	Closed	-	-	-	-	0.56 (0.32-0.96)	0.038*	0.74 (0.58-0.97)
Airbricks	None	1.00	-	-	-	-	1	-
	Unscreened	1.12 (0.71-1.75)	0.592	-	-	-	1.17 (0.82-1.68)	0.355
	Screened	0.18 (0.09-0.42)	<0.001*	-	-	-	0.30 (0.13-0.79)	0.006*
Roof type	Thatch	1.00	-	-	-	-	-	-
	Metal or tiles	0.44 (0.27-0.70)	0.001*	-	-	-	-	-

attention has been paid to quantifying heterogeneous biting propensities among individuals or households or environmental noise. While some evidence suggests biting patterns are spatially heterogeneous due to environmental factors⁴⁷, these patterns had not been examined critically alongside other measures of heterogeneity, such as seasonality, household quality or other factors, and environmental noise. Our study, like others⁴, has shown that heterogeneity is operating at multiple scales. Large differences among households, which are partly explained by household quality, are also influenced by spatial patterns related to other environmental covariates (e.g. elevation (altitude gradient) and urbanicity), which are likely related to mosquito ecology. Our method quantifies heterogeneity at a small scale (i.e. at the household level), validated by simulation studies, giving us confidence in our analysis of complex spatial patterns collected in field studies. A related analysis on the same households shows that mosquito counts data are also crudely associated with epidemiological measures of malaria risk⁴⁶. While the results of this study describe patterns of putative mosquito exposure, we note that every individual in the study was given a bednet and the reported rates of use were quite high, so the association between mosquito hotspots and malaria hotspots in this study remains unclear.

Our work provided an improved understanding of heterogeneity in malaria vector density at the three study sites in Uganda and offered a valuable opportunity for assessing whether interventions could be spatially targeted to be aimed at households with the highest mosquito biting propensities or spatial areas, i.e., hotspots of vector abundance. While analysis using scan statistics can identify statistical anomalies, our method is better at quantifying patterns and identifying its causes. We are able to identify cold spots (households with very small biting propensities) as well as hot spots (households with very large biting propensities). The study showed that the presence of mosquito hotspots was associated with the environment (including urbanicity, distance to water bodies or aquatic habitats, rainy seasons, and elevation) as well as housing characteristics. It has been found that house design in rural Uganda was associated with additional reductions in mosquito density and parasite prevalence following the introduction of IRS⁴⁸. In the Ugandan context, we proposed that malaria control efforts should be targeted towards houses with features that led to an elevated possibility of malaria transmission due to indoor vector abundance, including houses with mud walls, earth or sand floors, earth and dung floors, open eaves, unscreened airbricks, and thatched roofs. Improvements on housing design such as house-screening, closing the eaves, or installing ceilings help prevent mosquitoes gaining access to houses and should therefore reduce malaria transmission and prevalence^{49,50}.

Differences in vector seasonality were apparent at the three study sites. Based on the results of our study, Tororo (the most rural site) had the strongest seasonality, while Jinja (where mosquitoes were sampled in the town) had the weakest seasonality. The magnitude of seasonality at the three sites appeared to correlate with mosquito densities over time. The dynamics and seasonal abundance of malaria vectors are known to be associated with microecology, rainfall, and temperature patterns⁵¹. The

households in Kanungu and Tororo mostly had similar biting propensities during different seasons while households in Jinja showed greater differences in ω during different seasons. Clearly, the landcover in Jinja town would not support the temporary breeding sites that follow rainy seasons in rural areas. The identification of three hotspots in Jinja compared to only one hotspot each in Kanungu and Tororo for the dry vs. rainy seasons comparison suggested that mosquito attractiveness of households in Jinja might be more susceptible to changes in seasons than in Kanungu and Tororo. This could possibly be attributed to the fact that the sampled households in Jinja were in close proximity to a swampy area near Lake Victoria, which acted as a critical water body (see [Figure S1](#)). Distance of houses to water bodies had long been known as a risk factor of malaria⁵². These findings suggested that it would be possible to target different households during different seasons for optimal control efforts.

Household mosquito counts recorded in Jinja in the second half period of surveillance (1,172 mosquitoes) were only about half of those recorded in the first half period (2,309 mosquitoes). Similarly, household mosquito counts recorded in Kanungu in the second half period (5,177 mosquitoes) were only about half of those recorded in the first half period (10,018 mosquitoes). Such observations suggested some changes in malaria transmission over time at both study sites, however²⁷, found no association of LLIN with malaria transmission in Jinja and only modest effect of LLIN in Kanungu. A remarkable reduction of mosquito abundance after the deployment of IRS was observed in Nagongera where IRS began in December 2014 (see [Figure S5](#) and [Figure S6](#)), which at the same time reduced mosquito biting propensities in about two-third of the households. There were three hotspots of vector abundance with a much larger ω before IRS compared to after IRS (HH21, HH29, and HH107); these households were located reasonably close to each other, i.e. in the middle right area of the study site. The three hotspots with a much larger ω after IRS compared to before IRS were HH74, HH88, and HH104 which were located across the lower area of the study site. Despite a drastic reduction in household mosquito counts after IRS in these three households, i.e. from 1,295 mosquitoes to 223 mosquitoes in HH74, from 999 mosquitoes to 109 mosquitoes in HH88, and from 1,211 mosquitoes to 177 mosquitoes in HH104, biting propensities in these households after IRS increased instead of decreased, relative to other households.

In conclusion, the study found that housing quality contributed to a portion of the heterogeneity in household mosquito abundance. Cement walls, brick floors, closed eaves, screened airbricks, and tiled roofs are features of a house that had shown to reduce indoor mosquitoes reducing exposure to the vector of their inhabitants. Household mosquito biting propensities in Jinja showed some important differences during the dry season and the rainy season, most likely due to the close proximity of the study site to a swampy area. Jinja and Kanungu had shown a reduction of mosquito densities in the second half period of surveillance. The application of IRS in Tororo had caused a massive reduction in mosquito abundance as well as reducing household biting propensities in two-third of the

households. Based on these findings, IRS was a successful measure of vector control interventions in Uganda and improvements in house quality should be recommended as a supplementary measure for malaria control.

Data availability

Figshare Dataset 1: The Ugandan mosquito count data, household-level housing covariates, environmental covariates, and a full list of enumerated households for the three study sites. <https://doi.org/10.6084/m9.figshare.6797408.v3>⁵³.

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

Grant information

Bill and Melinda Gates Foundation [OPP1159934; Su Yun Kang and David L. Smith], [OPP1152978; Katherine E. Battle],

[OPP1106023; Harry S. Gibson]; Clinton Health Access Initiative [Katherine E. Battle]; National Institutes of Health & International Centers of Excellence in Malaria Research (ICMER) program [U19AI089674; David L. Smith, Kilama Maxwell, Moses Kanya, Steven W. Lindsay, Grant Dorsey, Bryan Greenhouse, Isabelle Rodriguez-Barraquer].

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

The authors would like to thank those in the study communities who participated in this study or helped collect mosquitoes. The authors wish to acknowledge the Infectious Diseases Research Collaboration (IDRC) for administrative and technical support.

A previous version of this article is available on BiorXiv (<https://doi.org/10.1101/299529>).

Supplementary material

Supplementary File 1: Supplementary Figures S1 to S6.

[Click here to access the data.](#)

Supplementary File 2: Detailed description of simulation analyses for identifying the most robust method for quantifying heterogeneity in malaria vector density.

[Click here to access the data.](#)

Supplementary File 3: R code for estimating household biting propensities and seasonal signal using a zero-inflated negative binomial model.

[Click here to access the data.](#)

Supplementary File 4: R code for performing negative binomial generalized linear regression analyses to identify housing risk factors of household mosquito biting.

[Click here to access the data.](#)

References

- Bhatt S, Weiss DJ, Cameron E, *et al.*: **The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015.** *Nature*. 2015; **526**(7572): 207–211.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- WHO: **World Malaria Report 2015.** World Health Organization. 2016.
[Reference Source](#)
- MW Service, Townson H: **The *Anopheles* vector.** In DA Warrell and HM Gilles, editors, *Essential Malariaology*. London: Arnold, 4 edition, 2002; 59–84.
- Bejon P, Williams TN, Nyundo C, *et al.*: **A micro-epidemiological analysis of febrile malaria in Coastal Kenya showing hotspots within hotspots.** *eLife*. 2014; **3**: e02130.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Woolhouse ME, Dye C, Etard JF, *et al.*: **Heterogeneities in the transmission of infectious agents: implications for the design of control programs.** *Proc Natl Acad Sci USA*. 1997; **94**(1): 338–342.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Smith DL, McKenzie FE, Snow RW, *et al.*: **Revisiting the basic reproductive number for malaria and its implications for malaria control.** *PLoS Biol*. 2007; **5**(3): e42.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Tusting LS, Bottomley C, Gibson H, *et al.*: **Housing Improvements and Malaria Risk in Sub-Saharan Africa: A multi-country Analysis of Survey Data.** *PLoS Med*. 2017; **14**(2): e1002234.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

8. Kigozi SP, Pindolia DK, Smith DL, *et al.*: **Associations between urbanicity and malaria at local scales in Uganda.** *Malar J.* 2015; 14(1): 374.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Stresman GH, Giorgi E, Baidjoe A, *et al.*: **Impact of metric and sample size on determining malaria hotspot boundaries.** *Sci Rep.* 2017; 7: 45849.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Bousema T, Griffin JT, Sauerwein RW, *et al.*: **Hitting hotspots: spatial targeting of malaria for control and elimination.** *PLoS Med.* 2012; 9(1): e1001165.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Gething PW, Casey DC, Weiss DJ, *et al.*: **Mapping *Plasmodium falciparum* Mortality in Africa between 1990 and 2015.** *N Engl J Med.* 2016; 375(25): 2435–2445.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Cairns ME, Walker PG, Okell LC, *et al.*: **Seasonality in malaria transmission: implications for case-management with long-acting artemisinin combination therapy in sub-Saharan Africa.** *Malar J.* 2015; 14(1): 321.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Guelbéogo WM, Gonçalves BP, Grignard L, *et al.*: **Variation in natural exposure to anopheles mosquitoes and its effects on malaria transmission.** *eLife.* 2018; 7: pii: e32625.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Simmons RA, Mboera L, Miranda ML, *et al.*: **A longitudinal cohort study of malaria exposure and changing serostatus in a malaria endemic area of rural Tanzania.** *Malar J.* 2017; 16(1): 309.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Reiner RC, Le Menach A, Kunene S, *et al.*: **Mapping residual transmission for malaria elimination.** *eLife.* 2015; 4: pii: e09520.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Ruktanonchai NW, DeLeenheer P, Tatem AJ, *et al.*: **Identifying Malaria Transmission Foci for Elimination Using Human Mobility Data.** *PLoS Comput Biol.* 2016; 12(4): e1004846.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. Kulldorff M: **A spatial scan statistic.** *Commun Stat Theory Methods.* 1997; 26(6): 1481–1496.
[Publisher Full Text](#)
18. Kanya MR, Arinaitwe E, Wanzira H, *et al.*: **Malaria transmission, infection, and disease at three sites with varied transmission intensity in Uganda: implications for malaria control.** *Am J Trop Med Hyg.* 2015; 92(5): 903–912.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Lindsay SW, Armstrong Schellenberg JR, Zeiler HA, *et al.*: **Exposure of Gambian children to *Anopheles gambiae* malaria vectors in an irrigated rice production area.** *Med Vet Entomol.* 1995; 9(1): 50–58.
[PubMed Abstract](#) | [Publisher Full Text](#)
20. Midega JT, Smith DL, Olotu A, *et al.*: **Wind direction and proximity to larval sites determines malaria risk in Kilifi District in Kenya.** *Nat Commun.* 2012; 3: 674.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Wanzirah H, Tusting LS, Arinaitwe E, *et al.*: **Mind the gap: house structure and the risk of malaria in Uganda.** *PLoS One.* 2015; 10(1): e0117396.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Lindsay SW, Adiamah JH, Miller JE, *et al.*: **Variation in attractiveness of human subjects to malaria mosquitoes (Diptera: Culicidae) in The Gambia.** *J Med Entomol.* 1993; 30(2): 368–373.
[PubMed Abstract](#) | [Publisher Full Text](#)
23. Kilama M, Smith DL, Hutchinson R, *et al.*: **Estimating the annual entomological inoculation rate for *Plasmodium falciparum* transmitted by *Anopheles gambiae* s.l. using three sampling methods in three sites in Uganda.** *Malar J.* 2014; 13(1): 111.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Uganda Bureau of Statistics: **The national population and housing census 2014–main report.** Technical report, Kampala, Uganda, 2016.
[Reference Source](#)
25. Gillies MT, Coetzee M: **A supplement to the Anophelinae of Africa South of the Sahara.** Johannesburg, South Africa: The South African Institute for Medical Research. 1987.
[Reference Source](#)
26. Wirtz RA, Duncan JF, Njlesenani EK, *et al.*: **ELISA method for detecting *Plasmodium falciparum* circumsporozoite antibody.** *Bull World Health Organ.* 1989; 67(5): 535–42.
[PubMed Abstract](#) | [Free Full Text](#)
27. Katureebe A, Zinszer K, Arinaitwe E, *et al.*: **Measures of Malaria Burden after Long-Lasting Insecticidal Net Distribution and Indoor Residual Spraying at Three Sites in Uganda: A Prospective Observational Study.** *PLoS Med.* 2016; 13(11): e1002167.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. Fletcher JJ, Underwood AJ: **How to cope with negative estimates of components of variance in ecological field studies.** *J Exp Mar Biol Ecol.* 2002; 273(1): 89–95.
[Publisher Full Text](#)
29. Qian SS, Shen Z: **Ecological applications of multilevel analysis of variance.** *Ecology.* 2007; 88(10): 2489–2495.
[PubMed Abstract](#) | [Publisher Full Text](#)
30. Lindenmayer DB, Likens GE: **Adaptive monitoring: a new paradigm for long-term research and monitoring.** *Trends Ecol Evol.* 2009; 24(9): 482–486.
[PubMed Abstract](#) | [Publisher Full Text](#)
31. Irwin BJ, Wagner T, Bence JR, *et al.*: **Estimating spatial and temporal components of variation for fisheries count data using negative binomial mixed models.** *Trans Am Fish Soc.* 2013; 142(1): 171–183.
[Publisher Full Text](#)
32. Lambert D: **Zero-inflated Poisson regression, with an application to defects in manufacturing.** *Technometrics.* 1992; 34(1): 1–14.
[Publisher Full Text](#)
33. Rue H, Martino S, Chopin N: **Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations.** *J R Stat Soc Series B Stat Methodol.* 2009; 71(2): 319–392.
[Publisher Full Text](#)
34. Sorbye SH, Rue H: **Scaling intrinsic Gaussian Markov random field priors in spatial modelling.** *Spat Stat.* 2014; 8: 39–51.
[Publisher Full Text](#)
35. Watanabe S: **A widely applicable Bayesian information criterion.** *J Mach Learn Res.* 2013; 14: 867–897.
[Reference Source](#)
36. Field AP: **Kendall's coefficient of concordance.** *Encyclopedia of Statistics in Behavioral Science.* 2005.
[Publisher Full Text](#)
37. Getis A, Ord JK: **The analysis of spatial association by use of distance statistics.** *Geogr Anal.* 1992; 24(3): 189–206.
[Publisher Full Text](#)
38. Bivand R, Anselin L, Berke O, *et al.*: **spdep: Spatial dependence: weighting schemes, statistics and models.** 2011.
[Reference Source](#)
39. Alegana VA, Kigozi SP, Nankabiwa J, *et al.*: **Spatio-temporal analysis of malaria vector density from baseline through intervention in a high transmission setting.** *Parasit Vectors.* 2016; 9(1): 637.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Pagano TS, Durham RM: **Moderate resolution imaging spectroradiometer (MODIS).** In *Sensor Systems for the Early Earth Observing System Platforms.* International Society for Optics and Photonics, 1993; 1939: 2–18.
[Publisher Full Text](#)
41. Funk CC, Peterson PJ, Landsfeld MF, *et al.*: **A quasi-global precipitation time series for drought monitoring.** Technical report, US Geological Survey, 2014.
[Publisher Full Text](#)
42. Esch T, Heldens W, Hirne A, *et al.*: **Breaking new ground in mapping human settlements from space | The Global Urban Footprint.** *ISPRS J Photogramm Remote Sens.* 2017; 134: 30–42.
[Publisher Full Text](#)
43. Chander G, Markham BL, Helder DL: **Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors.** *Remote Sens Environ.* 2009; 113(5): 893–903.
[Publisher Full Text](#)
44. Farr TG, Rosen PA, Caro E, *et al.*: **The shuttle radar topography mission.** *Rev Geophys.* 2007; 45(2).
[Publisher Full Text](#)
45. Trabucco A, Zomer RJ: **Global aridity index (global-aridity) and global potential evapo-transpiration (global-PET) geospatial database.** *CGIAR Consortium for Spatial Information,* 2009.
46. Rodríguez-Barraquer I, Arinaitwe E, Jagannathan P, *et al.*: **Quantification of anti-parasite and anti-disease immunity to malaria as a function of age and exposure.** *eLife.* 2018; 7: pii: e35832.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
47. Takken TW, Lindsay SW: **Factors affecting the vectorial competence of *Anopheles gambiae*: a question of scale.** In *Ecological aspects for application of genetically modified mosquitoes.* Kluwer Academic Publishers, 2003; 2: 75–90.
[Reference Source](#)
48. Rek JC, Alegana V, Arinaitwe E, *et al.*: **Rapid improvements to rural Ugandan housing and their association with malaria from intense to reduced transmission: a cohort study.** *Lancet Planet Health.* 2018; 2(2): e83–e94.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
49. Njie M, Dilger E, Lindsay SW, *et al.*: **Importance of eaves to house entry by anopheline, but not culicine, mosquitoes.** *J Med Entomol.* 2009; 46(3): 505–510.
[PubMed Abstract](#) | [Publisher Full Text](#)
50. Lindsay SW, Emerson PM, Charlwood JD: **Reducing malaria by mosquito-proofing houses.** *Trends Parasitol.* 2002; 18(11): 510–514.
[PubMed Abstract](#) | [Publisher Full Text](#)
51. Dery DB, Brown C, Asante KP, *et al.*: **Patterns and seasonality of malaria transmission in the forest-savannah transitional zones of Ghana.** *Malar J.* 2010; 9(1): 314.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
52. Kleinschmidt I, Sharp BL, Clarke GP, *et al.*: **Use of generalized linear mixed models in the spatial analysis of small-area malaria incidence rates in Kwazulu Natal, South Africa.** *Am J Epidemiol.* 2001; 153(12): 1213–1221.
[PubMed Abstract](#) | [Publisher Full Text](#)
53. Kang SY, Battle KE, Gibson HS, *et al.*: **Heterogeneous Exposure and Hotspots for Malaria Vectors at Three Study Sites in Uganda.** *figshare.* Fileset. 2018.
<http://www.doi.org/10.6084/m9.figshare.6797408.v3>

Open Peer Review

Current Referee Status:  

Version 2

Referee Report 04 December 2018

<https://doi.org/10.21956/gatesopenres.13949.r26797>



Mark Janko 

Global Health Institute, Duke University, Durham, NC, USA

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 14 November 2018

<https://doi.org/10.21956/gatesopenres.13949.r26757>



Gillian Stresman 

Department of Immunology and Infection, London School of Hygiene and Tropical Medicine, London, UK

No further comments to make.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Referee Report 28 August 2018

<https://doi.org/10.21956/gatesopenres.13915.r26613>



Mark Janko 

Global Health Institute, Duke University, Durham, NC, USA

The study sets out to improve upon our understanding of malaria transmission heterogeneity using data from 3 study sites in Uganda. Specifically, they investigate household vector density and relate it to various features, with particular focus on household construction quality. Overall, I think this is a good

paper, although I have some comments which I note below:

Major comment:

1. Mosquito abundance and malaria risk are not the same thing. The authors appear to equate the two, and I think it is more appropriate to treat mosquito biting as a *risk factor* for transmission. In other words, it is a necessary but not sufficient condition (e.g. the mosquito has to have malaria to transmit).
2. The authors state that there is a linear relationship between biting propensity and mosquito density. Figure 2 does indeed show a linear trend, but I think it important to note that the figure shows variability in biting propensity increases as mosquito density increases, particularly in the Kanungu and Tororo sites. This makes the linear story a bit less clear.
3. The authors set out to identify hot spots, but it was not clear how hot spots may evolve over time. This certainly has implications for spatial targeting.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Referee Expertise: Malaria, spatial and spatio-temporal modeling, Bayesian statistics

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reader Comment 31 Oct 2018

Su Yun Kang, University of Oxford, UK

Response: We thank the reviewer for their kind words and for their thorough reading of our manuscript.

Response to Major comment #1: We have made clear that the results of our study are about

mosquito abundance, which is one of the factors affecting risk of malaria transmission. (Some of these have already been described in greater detail in our response to Reviewer Gillian Stresman.)

Response to Major comment #2: We added the sentence: "We also note that there were differences in the distributions of biting propensities across sites."

Response to Major comment #3: Our results showed that hotspot spatial patterns change through time. We identify areas where hotspots are more common through the different seasons. These should be the areas that are targeted by interventions.

Competing Interests: No competing interests were disclosed.

Referee Report 25 July 2018

<https://doi.org/10.21956/gatesopenres.13915.r26563>



Gillian Stresman 

Department of Immunology and Infection, London School of Hygiene and Tropical Medicine, London, UK

Overall: This manuscript analyzes vector density data per household to estimate households that have higher or lower exposure to the malaria vector than expected, controlling for seasonality and other factors. They also associate certain household characteristics with greater mosquito densities. Overall, this is a very interesting paper and comprehensive analysis. However, some comments should be addressed.

Major Comments:

The authors have data on the density of the malaria vector per household, over time, in three settings of varying transmission intensity. Vector density is used here as a proxy for malaria transmission and risk however, important terms such as malaria risk, biting, and density imply very different things and here appear to be used interchangeably. No information on the malaria incidence/prevalence in these households was provided, a measure of whether the mosquitoes caught were infectious for malaria, or whether they had fed on humans was provided, limiting their interpretation of the data to vector density or risk of mosquito exposure. For example, malaria hotspots are mentioned several times, but in fact, what the authors are able to measure with their data is hotspots of mosquito abundance whereas the jump from mosquito abundance to malaria abundance is an assumption, which is not always true and is not clearly detailed as such in the manuscript. Similarly, the term 'biting propensity' is a ratio of expected to observed mosquito density but no information is given on whether these mosquitoes are biting (likely not given the high bednet coverage in the villages). These nuances in language are critical to ensure that authors are not over-stating their findings.

Minor Comments:

Intro Paragraph 1: The role of 'hotspots' in low and elimination settings in maintaining transmission and the effectiveness of spatially targeted interventions to hotspots has had mixed results in terms of actual impact (see Bejon et al., PLoS Med, 2010; Bousema et al. PLoS Med, 2016; Stresman et al., Sci rep, 2017 etc.). Current WHO strategy focus on 'foci' of transmission, or the spatially-delineated regions of malaria transmission that remain in low-endemic settings. The introduction should be framed in the context of the current literature.

Intro Paragraph 2: “heterogeneity in individual or household biting” do you mean individual or household exposure to mosquito bites?

Intro Paragraph 4: “Individuals who receive the most bites are most likely to be infected.....” They can also absorb more of the infectious bites as they are bitten more frequently.

Results Paragraph 4: “Kanungu was the study site with intermediate transmission intensity and thus exhibited....” Implies that the seasonal pattern of mosquitoes was a consequence of malaria transmission when in fact the opposite is more plausible.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reader Comment 31 Oct 2018

Su Yun Kang, University of Oxford, UK

Response to the overall comment: We thank the reviewer for their kind words and for their thorough reading of our manuscript.

Response to Major comments: We have revisited the manuscript following the suggestion of the reviewer. We made clear that our results showed spatio-temporal heterogeneity of mosquito abundance and biting propensity and not malaria risk by replacing the phrase ‘malaria risk’ with ‘vector abundance’ or ‘mosquito abundance’ throughout the manuscript. However, mosquito abundance has an effect on risk of malaria transmission. Areas with low number of vectors have usually a very low level of malaria transmission.

We have also added two lines to the discussion: “A related analysis on the same households

shows that mosquito counts data are also crudely associated with epidemiological measures of malaria risk [40]. While the results of this study describe putative patterns of mosquito exposure, we note that every individual in the study was given a bednet and the reported rates of use were quite high, so the association between mosquito hotspots and malaria hotspots in this study remains unclear." We note the recently published eLife paper has been added to the citations.

Response to Minor comment #1: The last line of P1 previously read: "In countries with low endemic and elimination settings where transmission is maintained in hotspots, spatially targeted interventions are valuable tools in malaria elimination efforts."

We have changed it to now read:

"The role of heterogeneity has been investigated in several recent studies [4, 8, 9] including the potential for spatially targeting interventions in foci, or hotspots within foci [10] but there is no current consensus about what exactly is meant by a hotspot or a focus or how to identify one."

Response to Minor comment #2: We have changed with:

"Heterogeneous biting rates among people driven partly by household mosquito abundance (i.e., superspreading) tend to stabilize endemic transmission, but provide an opportunity to increase efficiency of malaria control by targeting interventions at those who are bitten most [5, 6], such as those with the poorest housing quality [7]."

Response to Minor comment #3: We have added the following text: " , which will tend to stabilize transmission at low intensity but reduce overall exposure at high intensity as many of the bites are absorbed on fewer individuals [6]."

Response to Minor comment #4: We have changed the sentence to make it clearer:

"The intermediate transmission intensity found in Kanungu in previous studies [23, 40] could be explained by the moderate seasonal signal in mosquito counts over time (Figure 4(d))."

Competing Interests: No competing interests were disclosed.