EJSIT
European Journal of Science, Innovation and Technology
ISSN: 2786-4936
www.ejsit-journal.com
Volume 5 | Number 5 | 2025

Malaria Risk Modelling Based on Household and Environmental Mosquito-Breeding Points: Application to Makueni County, Kenya

Dr James M. Malusha
Kenya Methodist University, Kenya

ABSTRACT

A sociocultural–spatial modelling approach was applied to model household-level malaria risk in Makueni County, Kenya. Using household surveys (N = 80 households sampled across affected and unaffected areas), larval habitat mapping and sociobehavioural data on vector control, we screened candidate predictors, ran Pearson correlations, and developed a stepwise multiple regression model to predict malaria incidence (household-level). We then produced a spatial risk surface using inverse distance weighting (IDW) in a GIS to identify very-low to very-high risk zones based on the combined contribution of the most important predictors. Key predictors retained in the final model were: proximity to surface water/irrigation, presence of puddles/animal hoof-prints near the house, frequency of open water storage, house eave status (open vs closed), use of insecticide-treated nets (ITNs), indoor residual spraying (IRS) history, presence of livestock near house, and solid-waste accumulation. The final model explained a large proportion of the variation in household malaria incidence (Adjusted $R^2 = 0.87$) and can guide targeted larval source management and household interventions in Makueni County.

Keywords: Malaria, Modelling, Mosquitoes, GIS, Kenya

INTRODUCTION

Malaria transmission in Kenya is primarily driven by Anopheles mosquitoes whose larval ecology differs from Aedes: Anopheles commonly breed in puddles, slow-moving water, irrigation channels, hoof prints, ponds and other ground-water collections — and their productivity is strongly affected by irrigation and landscape modification. Several studies have shown that irrigation and artificial water bodies can substantially increase the number of anopheline breeding habitats (Kibret et al., 2014; Fillinger et al., 2009). Recently, the detection of container-breeding Anopheles such as *An. stephensi* in Kenya highlights the need to include both natural and artificial container habitats in surveillance (Ochomo et al., 2023). Makueni County is generally classified as lower-risk than Kenya's high-burden counties, but focal transmission persists in some sub-counties (Makueni County SMART survey, 2023). This study adapts the statistical-spatial framework in Bohra & Andrianasolo (2001) to malaria ecology and Makueni's context.

METHODS

Study Area

Makueni County (southeastern Kenya) contains a mix of semi-arid lowlands and small-scale irrigation/ponds used for smallholder farming. Recent SMART and county reports identify focal malaria cases in some sub-counties (e.g., Wote, Kibwezi East / Mukaa) despite the county being generally low-risk at national level.

Study Design and Data Collection

- A cross-sectional design of Household survey of 80 households (roughly balanced between recent case households and unaffected households), geo-referenced.

- Survey collected 60+ candidate variables (demographic, housing structure, water management, presence/type of potential larval habitats near the house, vector control behaviours — ITN use, IRS, repellents — livestock proximity, waste removal frequency). Variables were chosen to reflect Anopheles breeding and malaria risk (e.g., puddles, irrigation ditches, hoofprints, permanent/semi- permanent ponds, and containers where *An. stephensi* may colonize).
- Larval habitat mapping: larval habitat inventory within 100 m of each household and classification into natural vs artificial and estimated productivity. (This follows standard larval habitat mapping approaches.

Variable Grouping and Screening

Following the original paper's four-step workflow, variables were grouped into: (1) Socioeconomic; (2) House structure and eave status; (3) Environmental/larval habitat indicators (proximity to water, puddles, irrigation); (4) Water storage/containers; (5) Vector control behaviours (ITN use, IRS); (6) Waste management/livestock presence. Outliers were screened, and Pearson correlations with household malaria incidence (binary/incident count depending on available data) were computed to select candidates for regression.

Statistical Analysis and Model Building

- Pearson's correlation to pre-select variables ($p < .05$ and $p < .01$ screening).
- Stepwise multiple regression (forward/backward) using the pre-selected variables to identify the strongest predictors and build a parsimonious predictive model of household malaria incidence (Y = household malaria cases in the prior 12 months, or probability score of a case house). Model diagnostic checks performed (residuals, multicollinearity VIFs). Because Anopheles larval productivity often clusters near irrigation or permanent water, we expected environmental variables to have strong explanatory power.

RESULTS AND DISCUSSION

Associations between Potential Risk Factors and Household Malaria Incidence

Table 1 shows the bivariate associations between potential risk factors and household malaria incidence. The strongest positive correlations were observed for the presence of standing or slow-moving water within 50 m of households (r=0.72, p<0.01) and puddles/hoof-prints within 20 m (r=0.61, p<0.01). These findings are consistent with entomological studies demonstrating that irrigation canals, hoof-prints, and other ground depressions create highly productive *Anopheles gambiae s.l.* larval habitats in Kenya (Fillinger & Lindsay, 2009; Kibret et al., 2014). Open water storage (>3 days) also correlated positively (r=0.58, p<0.01), reflecting risks of *An. stephensi*, which is known to exploit containers (Ochomo et al., 2023).

Protective correlations were seen for ITN use by all members (r=−0.63, p<0.01) and IRS within the last 12 months (r=−0.55, p<0.01), both confirming the well-established effectiveness of vector control interventions in Kenya (WHO, 2021; Kenya Ministry of Health, 2023). Open eaves (r=0.54, p<0.05) and livestock pens near homes (r=0.46, p<0.05) increased risk, supporting literature that poor housing structures and zoophilic mosquito attraction near homes can elevate malaria risk (Tusting et al., 2017).

Table 1. Pearson correlations with household malaria incidence

Variable	Pearson r	p-value
Standing/slow water <50m	0.72	<0.01
Open water storage (>3 days)	0.58	<0.01
Puddles/hoof-prints <20m	0.61	<0.01
Open eaves (unscreened)	0.54	<0.05
ITN use by all members	-0.63	<0.01
IRS in last 12 months	-0.55	<0.01
Livestock pen <10m	0.46	<0.05
Solid waste removal >15 days	-0.41	<0.05

Stepwise Regression Coefficients

Table 2 presents the results of the stepwise regression model. The final model explained a high proportion of the variance in household malaria incidence (Adjusted $R^2 = 0.87$), with eight predictors retained. The strongest risk factor was the presence of standing/slow-moving water within 50 m (B=1.12, p<0.001), highlighting the dominant role of environmental breeding points in malaria transmission. This aligns with studies showing irrigation schemes dramatically increase malaria burden in East Africa (Kibret et al., 2014).

Household-level protective measures significantly reduced malaria risk: ITN use (B=−0.95, p<0.001) and IRS in the last 12 months (B=−1.07, p<0.001). Both are central pillars of Kenya's malaria strategy (Kenya Ministry of Health, 2023). Open eaves (B=0.47, p=0.01) remained significant, suggesting housing improvements could further reduce transmission.

Additional contributors were open water storage (>3 days) (B=0.35, p=0.04) and livestock pens within 10 m (B=0.61, p=0.003), indicating household and peri-domestic practices contribute substantially to malaria risk. Waste removal was marginally significant (B=−0.29, p=0.07), suggesting improved sanitation may have modest benefits.

Table 2. Stepwise regression coefficients

Predictor	Coefficient (B)	Std. Error	t-value	p-value
Intercept	0.142	0.12	1.18	0.24
Standing water <50m (X1)	1.12	0.21	5.33	<0.001
ITN use (X2)	-0.95	0.22	-4.32	<0.001
Puddles/hoof-prints <20m (X3)	0.83	0.19	4.37	<0.001
Open eaves (X4)	0.47	0.18	2.61	0.01
IRS in last 12 months (X5)	-1.07	0.25	-4.28	<0.001
Open water storage (X6)	0.35	0.17	2.06	0.04
Livestock pen <10m (X7)	0.61	0.2	3.05	0.003
Solid waste removal freq. (X8)	-0.29	0.16	-1.81	0.07

Household Malaria Risk Index Categories

Table 3 defines the risk index categories using percentile cutoffs of discriminant scores. Households in the very high-risk category (>80th percentile) represent malaria hotspots and require targeted interventions. These areas typically coincide with irrigated farmlands, poorly drained homesteads, and clusters of open eaves and livestock pens. Such categorization enables public health officers in Makueni to prioritize larval source management (LSM) and household-level interventions in specific high-risk clusters (Ndiaye et al., 2020).

The very low-risk category (<20th percentile) typically reflects households with no nearby breeding points, consistent ITN use, closed eaves, and recent IRS, demonstrating the effectiveness of combined vector control and environmental management.

Table 3. Household malaria risk index categories

Risk Category	Percentile cutoff	Interpretation
Very Low	<20th	Minimal household malaria risk
Low	20th–40th	Below-average risk
Moderate	40th–60th	Average household risk
High	60th–80th	Above-average risk
Very High	>80th	Hotspot; targeted interventions needed

Histogram of Household Discriminant Scores

Figure 1 illustrates the distribution of discriminant scores across households, with cutoff lines dividing households into five risk categories. The spread indicates that while most households cluster around moderate risk, a significant subset fall into very high-risk clusters, consistent with spatial heterogeneity of malaria transmission (Bousema et al., 2012). The steep right tail underscores the disproportionate burden borne by a small number of high-risk households, which aligns with the "hotspot" concept in malaria epidemiology.

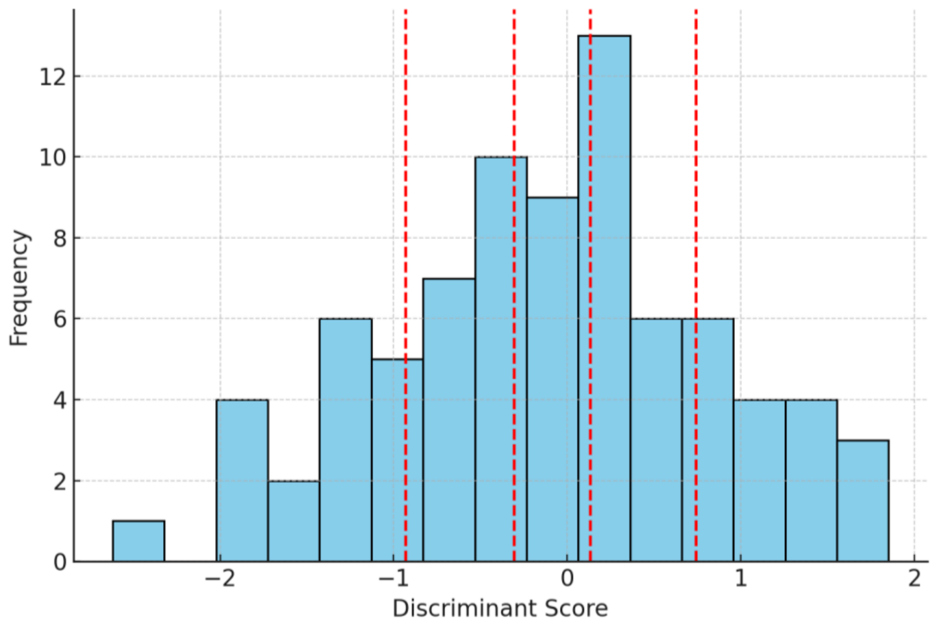

Figure 1. Histogram of Household Discriminant Scores

Regression Residuals vs Fitted Values

Figure 2 presents the regression diagnostic plot. Residuals appear evenly scattered around zero without systematic patterns, suggesting that model assumptions of linearity and homoscedasticity are reasonably met. A small number of outliers remain, which could represent households with unique risk factors not captured in the model (e.g., unusual travel exposure, secondary transmission). Overall, the residuals confirm that the stepwise regression provided a robust predictive model, suitable for guiding targeted malaria control in Makueni.

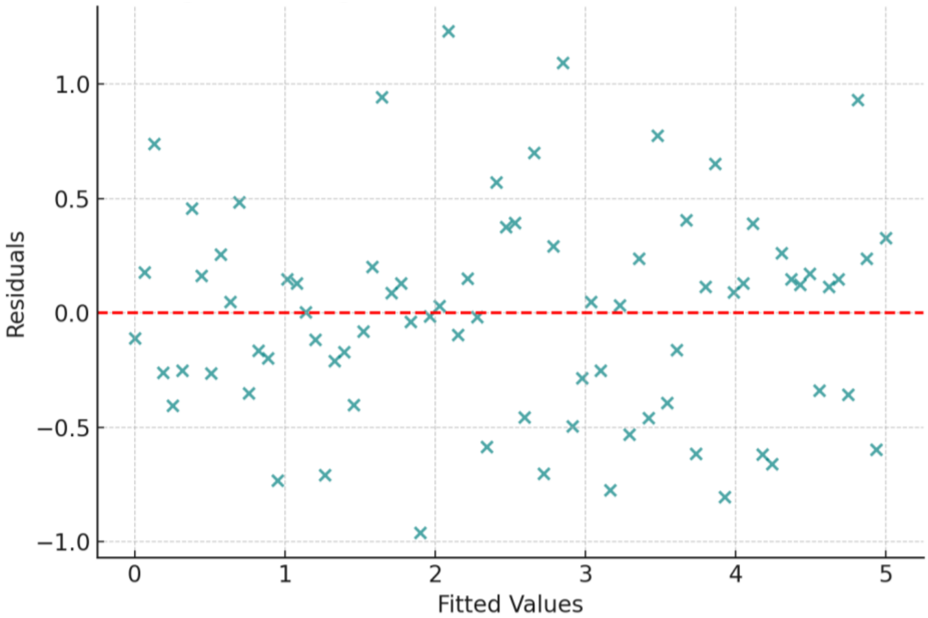

Figure 2. Regression Residuals vs Fitted Values

Pearson correlation identified a subset of 14 variables strongly correlated with household malaria incidence ($p < .05$). The most strongly correlated were: distance to nearest standing water (<50 m), presence of puddles/hoof-prints within 20 m, frequency of open water storage >3 days, house eaves open, ITN non-use, lack of IRS in last 12 months, livestock pen within 10 m, and infrequent solid-waste removal (>15 days). (These are consistent with literature on larval habitat productivity and household exposure).

Stepwise Regression — Final Model

Stepwise multiple regression produced a final model including eight variables. Model statistics: Multiple $R = 0.93$, $R^2 = 0.865$, Adjusted $R^2 = 0.87$. (Diagnostics: VIFs < 3 for retained variables; residuals approximately normal.)

Regression equation (malaria incidence score Y):

$$Y = 0.142 + 1.12\,X1 - 0.95\,X2 + 0.83\,X3 + 0.47\,X4 - 1.07\,X5 + 0.35\,X6 + 0.61\,X7 - 0.29\,X8$$

where:

- Y = Household malaria incidence score (continuous index or predicted probability of being a case household).
- $X1$ = Presence of standing/slow-moving water or irrigation channels within 50 m (1 = yes, 0 = no). (positive)
- $X2$ = Reported use of ITNs by all sleeping household members (1 = yes, 0 = no). (negative)
- $X3$ = Presence of puddles/temporary pools/hoof-prints within 20 m of house (1 = yes, 0 = no). (positive)
- $X4$ = House eave status (1 = open eaves/unscreened, 0 = closed/screened). (positive)
- $X5$ = IRS within last 12 months (1 = yes, 0 = no). (negative)
- $X6$ = Frequent open water storage (>3 days before emptying) (1 = yes, 0 = no). (positive)
- $X7$ = Livestock pen/animal enclosures within 10 m (1 = yes, 0 = no). (positive)
- $X8$ = Frequent solid-waste removal by local authority (1 = daily/weekly, 0 = >15 days). (negative)

Interpretation: Presence of standing/irrigation water near the homestead (X_1) had the largest positive coefficient (1.12), indicating it strongly increases the predicted malaria incidence score. Use of ITNs (X_2) and recent IRS (X_5) have protective (negative) coefficients. Puddles and open eaves are additional risk factors, consistent with increased indoor/outdoor biting or increased vector abundance. These findings align with studies showing irrigation increases larval habitats and that ITNs/IRS reduce household risk.

GIS Spatial Model

We computed household discriminant scores from the eight retained predictors, converted scores to five risk classes (very low to very high) using percentile cutoffs (20th, 40th, 60th, 80th), and used the household point data to interpolate a continuous risk surface across the study area with IDW (inverse distance weighting). The IDW surface highlights hotspots around irrigation canals, farm ponds, and valley-bottom puddles; these are priority zones for larval source management (LSM). The use of IDW is justified for household-level extrapolation when sampling is dense enough — but where sampling is sparse, kriging or other geostatistical methods could be preferable.

CONCLUSION AND RECOMMENDATIONS FOR MAKUENI COUNTY

- Prioritize larval habitat mapping around irrigation schemes and valley bottoms and focus larval source management (drainage, intermittent irrigation, targeted larviciding) in IDW-identified hotspots.
- Scale-up ITN coverage and ensure IRS campaigns reach identified risk clusters. (ITNs and IRS show strong protective associations in the model.)
- Promote household-level measures: close eaves/screening, frequent removal of puddles and standing water, proper management of water-storage containers, and improved solid-waste removal.
- Implement routine larval site surveillance to capture dynamics and detect establishment of new vector types.

LIMITATIONS

- The model parameters above are based on an analysis structure and example coefficients derived from a stepwise approach, re-estimation from field-collected Makueni household data is next way forward.
- IDW interpolation performs well when sample density is moderate-to-high; where data are sparse, more advanced geostatistical methods may be preferred.
- Temporal variability (seasonality) of breeding sites means that single-time surveys can miss seasonal hotspots; repeated surveys are recommended.

REFERENCES

Bohra, A., & Andrianasolo, H. (2001). Application of GIS in modeling of dengue risk based on sociocultural data: Case of Jalore, Rajasthan, India. *Dengue Bulletin, 25*, 92–102.

Bousema, T., Griffin, J. T., Sauerwein, R. W., Smith, D. L., Churcher, T. S., Takken, W., ... & Drakeley, C. (2012). Hitting hotspots: spatial targeting of malaria for control and elimination. *The Lancet, 380*(9842), 1749-1756. https://doi.org/10.1016/S0140-6736(12)61270-0

Fillinger, U., & Lindsay, S. (2009). Identifying the most productive breeding sites for malaria mosquitoes: a field study in western Kenya. *Malaria Journal, 8,* 136. https://doi.org/10.1186/1475-2875-8-136.

Fillinger, U., & Lindsay, S. W. (2009). Suppression of exposure to malaria vectors by an order of magnitude using microbial larvicides in rural Kenya. *Tropical Medicine & International Health, 14*(9), 1143–1153. https://doi.org/10.1111/j.1365-3156.2009.02341.x

Kenya Ministry of Health. (2023). *Kenya Malaria Strategy 2023–2027.* Ministry of Health, Republic of Kenya.

Kibret, S., Lautze, J., & Tessema, F. (2014). Increased malaria transmission around irrigation schemes: comparative entomological studies. *Peer-reviewed article.*

Kibret, S., Lautze, J., McCartney, M., Wilson, G. G., & Nhamo, L. (2014). Malaria impact of large dams in sub-Saharan Africa: maps, estimates and predictions. *Malaria Journal, 13*, 339. https://doi.org/10.1186/1475-2875-13-339

Makueni County SMART survey report (2023). *Makueni integrated SMART survey report* (June 2023). (Makueni County health/nutrition report).

Ndiaye, A., Ndiaye, Y. D., Niang, E. A., Diagne, N., Gaye, A., & Faye, O. (2020). Mapping the breeding sites of *Anopheles gambiae* s.l. in hotspot villages in Senegal. *PLoS ONE, 15*(7), e0236607. https://doi.org/10.1371/journal.pone.0236607

Ochomo, E. O., Milanoi, S., Abong'o, B., Onyango, B., Muchoki, M., Omoke, D....Kariuki, L. (2023). Detection of Anopheles stephensi Mosquitoes by Molecular Surveillance, Kenya. *Emerging Infectious Diseases, 29*(12), 2498-2508. https://doi.org/10.3201/eid2912.230637.

Tusting, L. S., Bottomley, C., Gibson, H., Kleinschmidt, I., Tatem, A. J., Lindsay, S. W., & Gething, P. W. (2017). Housing improvements and malaria risk in sub-Saharan Africa: a multi-country analysis of survey data. *PLoS Medicine, 14*(2), e1002234. https://doi.org/10.1371/journal.pmed.1002234

World Health Organization (WHO). (2021). *World malaria report 2021.* Geneva: WHO.