

# Bayesian Spatio-Temporal Models for the Incidence of Malaria Using Time Dependent Covariates

Evalyne Nduvi Musyoka<sup>\*</sup>, Samuel Mwalili, Boniface Malenje

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya

## Email address:

evemusioka@gmail.com (Evalyne Nduvi Musyoka), smusili@jkuat.ac.ke (Samuel Mwalili), bonniemiya@gmail.com (Boniface Malenje)

<sup>\*</sup>Corresponding author

## To cite this article:

Evalyne Nduvi Musyoka, Samuel Mwalili, Boniface Malenje. Bayesian Spatio-Temporal Models for the Incidence of Malaria Using Time Dependent Covariates. *International Journal of Data Science and Analysis*. Vol. 9, No. 3, 2023, pp. 60-66.

doi: 10.11648/j.ijdsa.20230903.12

**Received:** October 23, 2023; **Accepted:** November 3, 2023; **Published:** November 11, 2023

---

**Abstract:** This research study focuses on the Spatial and temporal Modelling of malaria incidences in Kenya, taking into account Time- dependent covariates. Malaria remains a significant public health concern in Kenya, with varying rates of infection across its 47 counties. Environmental factors such as temperature, rainfall, humidity and elevation play a crucial role in influencing Malaria transmission. Despite numerous malaria control efforts and initiatives the burden of the disease persist. The main objective of this study was to formulate Bayesian Spatio-temporal models for malaria incidence, with a particular emphasis on incorporating time-dependent covariates. The availability of data collected over time from various counties, as provided by the malaria project Atlas, was essential for achieving this goal. The Besag-York-Mollié (BYM) Spatio-temporal Model were formulated and implemented using Bayesian approach. Bayesian inference technique, coupled with Markov Chain Monte Carlo (MCMC) algorithms, was used to fit the models to the data. We also conducted convergence diagnostic of MCMC algorithm in order to check if the algorithm has converged and how reliable the posterior estimates are. In the analysis under Bayesian model choice and comparison of spatio-temporal model, spatial model with time dependent covariates and Spatio-temporal model with time dependent covariate were fitted. We found out that Spatio-temporal model with Time Dependent covariates was the best model. The resulting model and maps will be valuable for identifying disease hotspots, allocating resources for disease prevention and mitigation, and guiding policy decisions to reduce the burden of malaria. To ensure the validity of the Bayesian analysis, MCMC diagnostics were applied, including the Geweke Test, Gelman-Rubin statistics, and trace plots. These tests confirmed that the MCMC chains had converged to a common distribution, indicating the reliability of the obtained results.

**Keywords:** Spatial Model, Spatio-Temporal Model, MCMC Convergence, Gelman Rubins Statistics, Malaria-Incidences, Geweke Test

---

## 1. Introduction

Malaria remains one of the deadliest infections in the world. It is a mosquito-borne protozoan disease that is caused by Plasmodium parasites. These parasites exist in five species: Plasmodium falciparum, P. malariae, P. ovale, P. vivax, and P. knowlesi [3]. It is estimated that malaria is a risk to almost half of the world's population in nearly 100 countries and territories [6]. There were estimated 247 million cases of malaria in 2021 reported in 84 malaria-endemic countries [9]. It resulted in approximately 619,000 deaths during the same

year. The statistics show an increase in malaria from 245 million in 2020. Most of the increase in the cases are from countries in the African region.

Over the years, tremendous steps have been taken to enhance malaria control and research programs. Organizations like the World Health Organization (WHO) and Global Technical Strategy (GTS) in the past years stipulated a sum of \$6.4 billion annually to achieve a 90% decrease in malaria incidences and mortality rates by 2023 [8]. There has been the use of intervention efforts, including indoor residual spraying (IRS), artemisinin-based combination therapy (ACT), and insecticide-treated bed nets

(ITNs) [3]. However, there are still many incidences despite these investments, among other eradication strategies that the WHO initiated [8].

Globally, malaria cases incidence (cases per 1000 population at risk) reduced from 82 in 2000 to 52 in 2019 [9]. Later the incidence increased to 59 in 2020. Further, between 2020 and 2021, there was no change in case incidence. The change in 2020 came from the COVID-19 pandemic, which disrupted services worldwide. COVID-19 pandemic disruption resulted in an estimated additional 13.4 million cases globally between 2019 and 2021.

Kenya is in Eastern Sub-Saharan Africa, and its malaria incidence is also high. Malaria in Kenya is a risk to about 70% of the country's 47 million inhabitants [11]. Besides, 13 million people stay in endemic areas and 19 million in highland epidemic-prone and seasonal transmission areas [9]. The western Kenya region has the highest burden of infection. Malaria infection and transmission in this geographical region are mainly determined by altitude, temperature, and rainfall patterns. It leads to a variation of the malaria prevalence by season and across the geographical zones in the country. The disease accounts for about 30% of outpatients' attendance in private and public health facilities across the country [11].

It is reported that *Plasmodium falciparum* causes the most severe form of the disease. This parasite accounts for 99% of infections in Kenya. Also, it is home to four *Plasmodium* parasites that infect humans. The number of malaria cases varies across the years. The malaria incidence cases are approximated to averagely being 6.8 million in 2000 and 3.4 million in 2021 [9]. Malaria incidence cases are highly experienced in endemic regions like Lake Victoria and coastal regions [8]. This region has an ambient temperature suitable for malaria transmission. The region also experiences necessary and long seasonality of rainfall. Temperature, rainfall, and humidity are determinants of perennial transmission of malaria.

Other regions that experience malaria incidences include the highland epidemic-prone areas, semi-arid, seasonal malaria transmission areas, and low-risk malaria areas (KMIS, 2020). The western highland of Kenya is seasonal, and they have epidemic malaria events that favour climatic conditions favouring the sustainability of malaria. The temperatures experienced in the region are a minimum of 18°C. The temperatures are experienced during periods of short and long rains favouring sustainable breeding and sporogony, thus occasion malaria transmission. On the other hand, the semi-arid areas of the northeastern, southeastern, and northern parts of Kenya experience short periods of intense transmission of malaria during the rainfall season. During the events of normal rainfalls, the temperatures are usually high, and water pools are always available for breeding habitats of malaria vectors. Malaria can also be introduced in low-risk malaria areas if there are changes in the hydrological cycle due to climate change. These areas include the central highlands of Kenya, including Nairobi.

Malaria is still a public health threat for locals and

travelers globally in Africa, Kenya, and other malaria-endemic regions. There is still a high burden of infections despite the important role played by WHO GTS [8] and the Kenyan Ministry of Health [11] in scaling down malaria incidences. There is a need for more studies to be able to achieve the goal of eliminating malaria. Eradication of malaria would therefore mean a more healthy nation and so growth of the country's social and economic status. There will be an achievement of sustainable development goals and universal health coverage [16]. It is also important in attaining Kenya's vision 2030 and achieving the Government's six pillars of bottom-up economic plan since Health care is among them.

In order to prevent or cure the disease in the context of its epidemics, decision-makers need to be aware of the risk the epidemic in space and time [15, 2]. Understanding spatial and temporal distributions of a disease is often accomplished by applying statistical methods to surveillance data and generating a map that describes the variations in risk [10]. Spatial statistics provides tools to analyse spatially and/or temporally distributed data, capitalizing on the correlation between incidence to interpolate and delineate areas with high disease risk. Geostatistics is a powerful spatial technology which contributes immensely to prediction of random process distributed in space or time, and has increasingly been applied in epidemiological studies facilitating quantification of spatial features of disease's transmission and its interpolation within the environment [13].

Despite the efforts that have been put in place, the burden of Malaria is still being felt at varied rate across many counties in Kenya. Several studies have shown that malaria infection is influenced by environmental factors such as temperature, rainfall, humidity and elevation. Even though there is quit a number of spatio-temporal models for malaria but they do not consider clustering and Time Dependent covariates. The main aim of this study is to formulate the spatio-temporal models for malaria incidence varying over time in Kenya using Time Dependent covariates.

## 2. Study Methods

### 2.1. Data and Study Variables

The study obtained publicly available secondary Malaria data from The Malaria Atlas Project Data Platform.<sup>1</sup> The platform provides malaria data at varying levels of detail to suite different needs. We will only consider the *Plasmodium* parasite data from Kenya. The main intervention is insecticide-treated bed nets (ITNs), which are a form of personal protection that has been shown to reduce malaria illness, severe disease, and death due to malaria in endemic regions the response variable of interest is number of newly diagnosed *plasmodium falciparum* malaria cases, on a given year, from 2010 to 2020. The Time Dependent covariates include;

- 1) Insecticide-treated bed nets access

2) Insecticide-treated bed nets use.

## 2.2. Spatial-Temporal Model

We assumed that  $y_i$  represents the number of incidence cases of malaria in each county  $i$ , which follows a Poisson distribution. Let  $Y_{it}$ , denote the number of Malaria incidence cases at country  $i$  and year  $t$ . Malaria data is available for different time periods  $t = 1, \dots, T$ . The linear predictor  $\eta_{it}$ , will be decomposed additively into components depending on space, time, or both. For the spatial component we will adopt the standard [1] model with a spatially unstructured  $u$  and structured component  $v$ . We thus consider a random effects model with mean  $\eta$ : We assume that

$$Y_{it} \sim \text{Pois}(\eta_{it}) \quad (1)$$

$$\log \eta_{it} = \log(e_{it}) + \log(r_{it}) \quad (2)$$

$$\log \eta_{it} = \log(e_{it}) + \beta_0 + \beta_i x_i + u_{it} + v_{it} + p_t \quad (3)$$

Where the term  $p_t$ , represents the effect of the  $t^{th}$  period, modelled as an autoregressive conditional random term. The unstructured random effects  $u_i$ ,  $i = 1, \dots, n_i$ , are assumed to be independent mean-zero normally distributed with unknown variance  $\sigma_u^2$ . To account for the fact that geographically close regions often have similar incidence rates the spatially structured component  $v$  is modeled as an intrinsic Gaussian Markov random field [14].

## 2.3. Spatial Model with Time Dependent Covariates

We initially assumed that  $y_i$  represents the number of incidence cases of malaria in each county  $i$ , which follows a Poisson distribution. Since the covariates considered are area-specific and time-dependent they can be included in the model equivalent to the BYM. With time dependent covariate  $x_{it}$ , the model is now given by

$$\log \eta_{it} = \log(e_i) + \beta_0 + u_i + v_i + \beta_i x_{it} \quad (4)$$

## 2.4. Spatio-Temporal Model with Time Dependent Covariates

We introduce the effect of the  $t^{th}$  period  $p_t$ , into the model, since the covariates considered are area-specific and time-dependent their inclusion in the model equivalent to the BYM: interaction among space and time random effects. With time dependent covariate  $x_{it}$ , the model is now given by;

$$\log \eta_{it} = \log(e_{it}) + \beta_0 + u_{it} + v_{it} + p_t + \beta_i x_{it} \quad (5)$$

## 2.5. Bayesian Estimation Methods

In Bayesian inference, the parameters within the likelihood model are allowed to be stochastic, that is, to have distributions. These distributions are called prior distributions and are assigned to the parameters before seeing the data. This allowance also makes the parameters in the prior distributions of the likelihood parameters to be stochastic. By so doing, hierarchical models are obtained. These models form the basis of inference under the Bayesian paradigm. The product of the

likelihood (data) and the prior distributions for the parameter gives the so-called posterior distribution. This distribution describes the behavior of the parameters after observing the data and making the necessary prior assumptions. However, most disease mapping models are complex and the resulting posterior distributions are not analytically tractable. Hence it is often not possible to derive simple estimators for parameters such as the relative risk. In this case posterior distribution is obtained via posterior sampling i.e., using simulation methods to obtain samples from the posterior distribution which then can be summarized to yield estimates of relevant quantities. Markov chain Monte Carlo (MCMC) methods are a set of methods which use iterative simulation of parameter values within a Markov chain. The theory of MCMC was first developed as a tool for Bayesian posterior sampling starting in the early 1990s [17]. Then using Bayes theorem the posterior distribution of  $\theta$  is given by:

$$\pi(\theta|y) = \frac{p(y|\theta) \times p(\theta)}{p(y)} \quad (6)$$

Here  $p(\theta)$  is the prior probability distribution of  $\theta$  which represents the prior belief on  $\theta$ ;  $p(y|\theta)$  is the likelihood function which specifies the distribution of the data  $y$  given the prior belief;  $p(y)$  is the marginal distribution of the data which is independent  $\theta$  and is treated as just a normalization constant. Thus the posterior distribution of  $\theta$  is often stated as:

$$\pi(\theta|y) \propto p(y|\theta \times p(\theta)) \quad (7)$$

### Markov Chain Monte Carlo

The aim of MCMC procedures is to generate random variables with stationary distributions that are similar to certain target distributions having probability distribution function  $\pi(y)$ . The target distributions in the Bayesian inference technique is often the posterior distribution  $p(\theta|y)$ .

### Gibbs Sampler

Gibbs Sampler was first developed by Geman and Geman (1984) for Bayesian image re-construction and later proposed by Gelfand and Smith (1990) as a sampling procedure for simulating marginal distributions in a Bayesian estimation context. This algorithm is structured as follows [7].

Set the initial values  $\theta^t = (\theta_1^{(t)}, \dots, \theta_p^{(t)})$ , for all the parameters and set  $t=1$ ,

Draw  $\theta^t = (\theta_1^{(t)}, \dots, \theta_p^{(t)})$ , by

$$\theta_1^{(t)} \sim p(\theta_1 | \theta_2^{(t-1)}, \dots, \theta_p^{(t-1)})$$

$$\theta_2^{(t)} \sim p(\theta_2 | \theta_1^{(t-1)}, \dots, \theta_p^{(t-1)})$$

$$\theta_d^{(t)} \sim p(\theta_d | \theta_2^{(t-1)}, \dots, \theta_{p-1}^{(t-1)})$$

Increase  $t$  by 1 that is let  $\theta^{t+1} = (\theta_1^{(t+1)}, \dots, \theta_p^{(t+1)})$

The Gibbs Sampler has attracted a lot of interest and attention in disease mapping and other epidemiological research due to the accessibility of cutting-edge software like Win BUGS, which has enabled its implementation and application in a variety of issues conceivable.

## 2.6. MCMC Convergence

Convergence is diagnosed when the chains have ‘forgotten’ their initial values, and the output from all chains is indistinguishable. Geweke proposed a convergence diagnostic for Markov chains based on a test for equality of the means of the first and last part of a Markov chain (by default we use the first 10% and the last 50%) [5]. Geweke’s approach involves calculation of the sample mean and asymptotic variance in each window, the latter being determined by spectral density estimation. His convergence diagnostic  $Z$  is the difference between these two means divided by the asymptotic standard error of their difference. As the chain length  $\rightarrow \infty$ , the sampling distribution of the chain has converged. Hence values of  $Z \rightarrow (0, 1)$  which fall in the extreme tails of a standard normal distribution,  $\pm 2$ , suggest that the chain has not fully converged.

Geweke and Gelman proposed a general approach to monitoring convergence of MCMC output in which two or more parallel chains are run with starting values that are over dispersed relative to the posterior distribution [4, 5]. Convergence for multiple chains may be assessed using Gelman-Rubin scale factor reduction factors that compare variation of the samples parameter values within and between chains. It is based on a comparison of within-chain and between-chain variances, and is similar to a classical analysis of

variance. To measure the variability of sample  $\theta_j^t$ , within the chain ( $j = 1, \dots, J$ ) define

$$V_j = \sum_{t=T+1}^{T+M} \frac{(\theta_j^t - \bar{\theta})^2}{M-1} \quad (8)$$

Over  $M$  iterations after an initial burn-in of  $T$  iterations, where  $\bar{\theta}$  is the average of  $\theta_j^{(t)}$  ( $t = T+1, \dots, T+M$ ). Ideally, the burn-in period is the initial set of samples where the effect of initial parameter values tails off. Convergence is therefore assessed from  $T+1$  to  $T+M$ . Variability within chains  $V_W$  is the average of  $V_j$ ’s, between chain variance is measured by

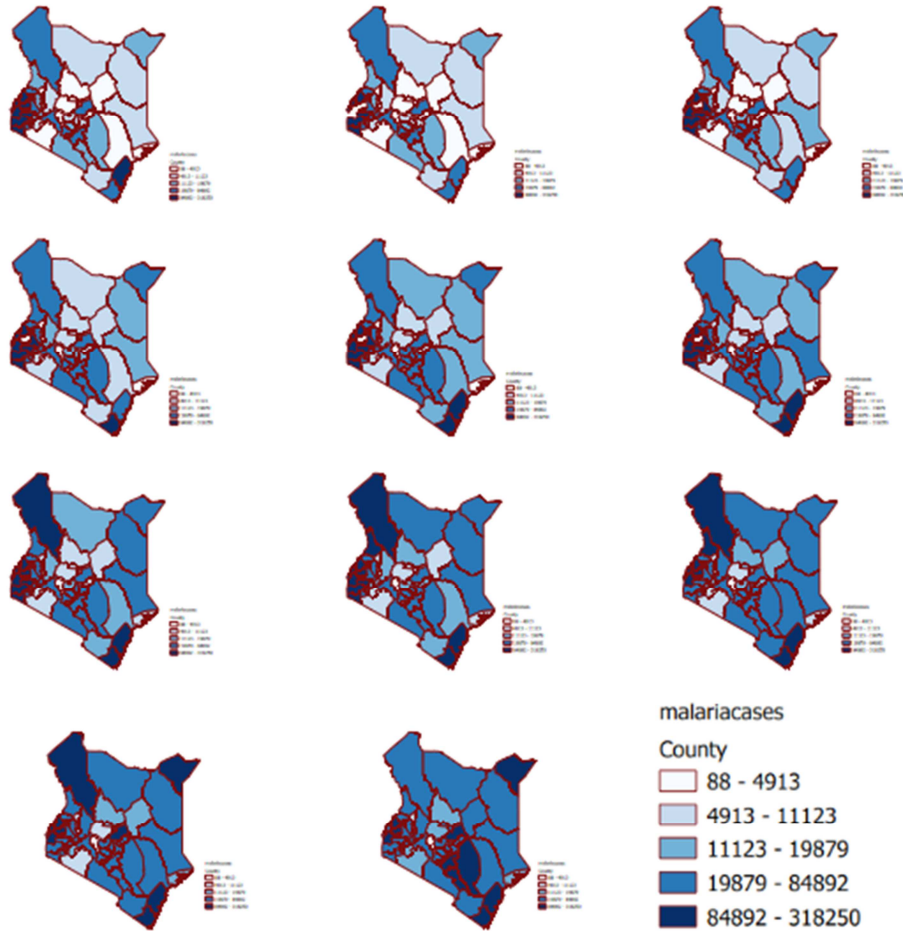
$$V_B = \sum_{t=T+1}^{T+M} (\bar{\theta}_j - \bar{\theta})^2 \quad (9)$$

Where  $\bar{\theta}$  is the average of  $\bar{\theta}_j$ ’s. The scale factor reduction (SRF) compares a pooled estimator of  $(\theta)$ , given by  $V_P = \frac{V_B}{M} + \frac{MV_W}{M-1}$ , to  $V_W$ . More specifically,  $SRF = \sqrt{\frac{V_P}{V_W}}$  with values under 1.2 [2] indicating convergence.

## 3. Results

The figure below represents malaria cases distribution in Kenya from 2010 to 2020.

*Observed malaria incidence cases maps*



**Figure 1.** Shows how malaria inciden cases are distributed among Kenyan counties from 2010 to 2020 respectively.

### 3.1. Spatial and Spatio Temporal Output

**Table 1.** Parameter estimates from Spatio-temporal model.

Parameter	Estimate	SD	2.5%	97.5%
$\beta_0$	-0.2255	0.3106	-0.5585	-0.1169
$\beta_1$	-0.0009	0.0003	-0.0014	-0.0002
$\beta_2$	0.0009	0.0004	0.0004	0.0022
$\sigma_p$	0.3496	0.1073	0.1941	0.5975
$\sigma_u$	0.5521	0.0277	0.5035	0.6077
$\sigma_v$	3.8520	0.4175	3.1400	4.7660

In table 1, we considered the time effect, covariates and the county random effect. In the outputs subscript 1 and 2 represents access to ITN and use of ITN respectively. Since the value of  $\sigma_v$  is large, this implies that the random variability are related to the spatial location of the data points. Malaria cases are varying from one county to another and the malaria cases of one county affect the malaria cases of the neighbouring county. However the coefficient of  $\sigma_p$  is small this implies that there is no temporal variability of malaria cases over time. Access to ITN has a negative coefficient which shows that access to ITN helps to mitigating malaria cases.

**Table 2.** Parameter estimates from spatial model with TVC.

Parameter	Estimate	SD	2.5%	97.5%
$\beta_0$	-2.3170	2.0470	-4.4050	-0.2544
$\beta_1$	-0.0005	0.0015	-0.0011	-0.0001
$\beta_2[1]$	0.0470	0.0552	-0.0093	0.1034
$\beta_2[2]$	0.0382	0.0428	-0.0060	0.08367
$\beta_2[3]$	0.0322	0.0346	-0.0031	0.0678
$\beta_2[4]$	0.0381	0.0358	0.0010	0.0751
$\beta_2[5]$	0.0354	0.0307	0.0032	0.0664
$\beta_2[6]$	0.0301	0.0277	0.0018	0.0585
$\beta_2[7]$	0.0315	0.0299	0.0007	0.0636
$\beta_2[8]$	0.0283	0.0288	-0.0012	0.0576
$\beta_2[9]$	0.0276	0.0288	-0.0032	0.0573
$\beta_2[10]$	0.0334	0.0338	-0.0013	0.0680
$\beta_2[11]$	0.0575	0.0575	-0.0017	0.1164
$\sigma_u$	0.9063	0.3588	0.5157	1.3340
$\sigma_v$	3.6980	0.4597	2.9130	4.7130

Table 2 reveals the influence of covariates on malaria incidences over time. A positive correlation exists between the use of insecticide-treated bed nets (ITN) and malaria cases, indicating that regions with higher malaria burdens tend to employ ITNs more frequently. Conversely, access to ITNs displays a negative coefficient, suggesting that areas with better ITN availability experience lower malaria incidence rates. The substantial variance in the structured component underscores the strong spatial dependence of malaria cases between neighboring counties. In essence, this highlights the role of ITNs in malaria prevention and the significant spatial dynamics of malaria transmission.

**Table 3.** Parameter estimates from spatio-temporal with TVC.

Parameter	Estimate	SD	2.5%	97.5%
$\beta_0$	-2.3820	2.3670	-4.7670	0.0173
$\beta_1$	-0.0002	0.0012	-0.0017	0.0013
$\beta_2[1]$	0.0428	0.0535	-0.0129	0.1000
$\beta_2[2]$	0.0358	0.0445	-0.0103	0.0825

Parameter	Estimate	SD	2.5%	97.5%
$\beta_2[3]$	0.0481	0.0235	0.0236	0.0721
$\beta_2[4]$	0.0445	0.0233	0.0193	0.0690
$\beta_2[5]$	0.0255	0.0402	-0.0154	0.0663
$\beta_2[6]$	0.0295	0.0300	-0.0011	0.0603
$\beta_2[7]$	0.0286	0.0315	-0.0036	0.0607
$\beta_2[8]$	0.0277	0.0303	-0.0034	0.0587
$\beta_2[9]$	0.0371	0.0229	0.0135	0.0606
$\beta_2[10]$	0.0515	0.0215	0.0294	0.0734
$\beta_2[11]$	0.0515	0.0677	-0.0183	0.1214
$\sigma_p$	1.0770	0.2715	0.6829	1.7260
$\sigma_u$	0.7356	0.1298	0.5618	0.9158
$\sigma_v$	3.8000	0.4249	3.0770	4.7450

In this model we considered the time effect, time-dependent covariates and the county random effect.

Table 3 shows that the use of insecticide-treated bed nets (ITN) is positively associated with malaria cases, indicating that regions with higher malaria burdens rely more on ITNs. Conversely, access to ITNs has a negative correlation, implying that areas with greater ITN availability experience fewer malaria cases. The substantial temporal variance suggests significant fluctuations in malaria cases over time, while the high structural random effect variance underscores the influence of neighboring counties on each other's malaria incidences.

### 3.2. Bayesian Model Comparison

The analysis gave the following parameter estimates and the goodness of fit measures, as presented in Table 3.

**Table 4.** Bayesian model comparison.

Model	$\bar{D}$	$\bar{D}$	$\rho D$	DIC
Spatio-temporal model	6756.524	6240.678	515.846	7272.370
Spatial model with TDC	6755.206	6238.602	516.604	7271.810
Spatio-temporal Model with TDC	6753.570	6236.630	516.940	7270.510

For model comparison, the effective number of parameters ( $\rho D$ ) and the deviance information criterion (DIC) were computed. The best fitting model is one with the smallest DIC value. From the DIC values in Table 3, it is clear that models with time dependent covariates is the best model since it has the smallest DIC. This confirms that Bayesian spatio temporal model with time-dependent covariate produces better results.

### 3.3. MCMC Convergence

#### 1) Geweke Test

Geweke test was performed on the two sets of Markov Chain Monte Carlo chains. The test is used to assess convergence of Markov Chain Monte Carlo chains.

**Table 5.** Results from Geweke Test.

chain 1 output
Fraction in 1 <sup>st</sup> window = 0.1
Fraction in 2 <sup>nd</sup> window = 0.5

**Table 6.** Results from Geweke Test.

chain 2 output
Fraction in 1 <sup>st</sup> window = 0.1
Fraction in 2 <sup>nd</sup> window = 0.5

The test provides a fraction for two windows. In this case, the fraction is 0.1 in the first window and 0.5 in the second window for both Chains. This shows that the algorithm used converged.

### 2) Gelman-Rubins statistics

**Table 7.** Gelman & Rubin's variance inflation factor.

Parameter	Point est.	Upper C.I.
$\beta_1$	1.00	1.00
$\beta_{2[11]}$	1.00	1.00
$\sigma_p$	1.02	1.02
$\sigma_U$	1.01	1.01
$\sigma_V$	1.00	1.00

From Table 5, it appears that for all parameters, the upper bound of the inflation factor is less than 1.1. There is no large deviation for the variance within-chain and between chain. This indicates that there's high confidence that the algorithm has converged.

### 3) Trace Plots.

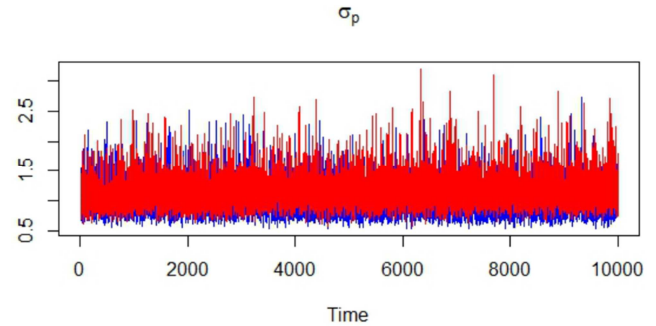
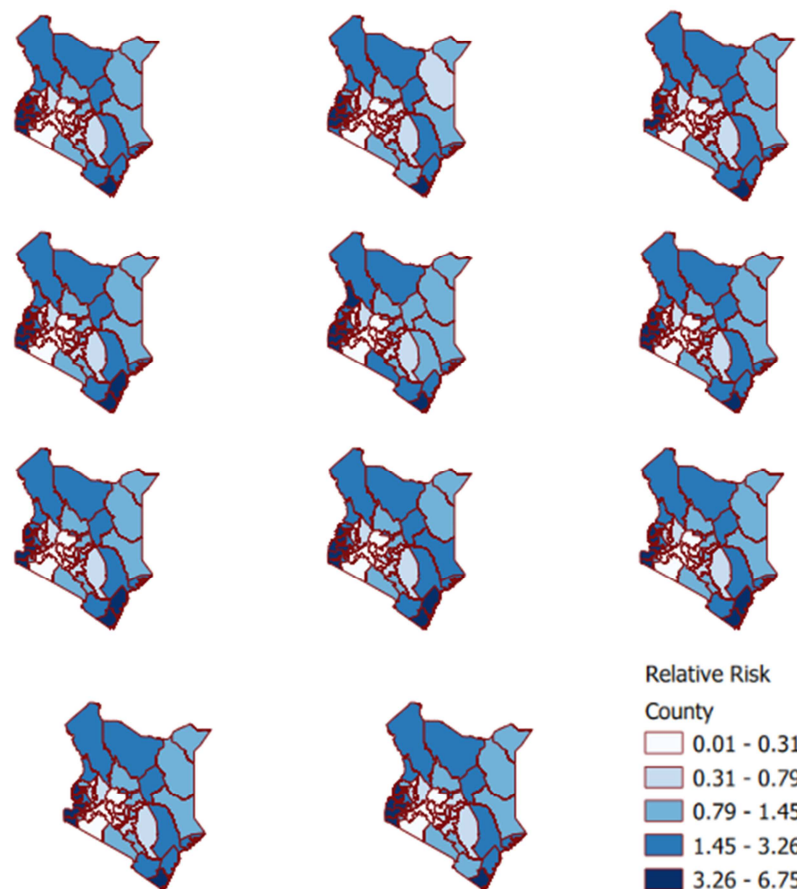
**Figure 2.** History of the MCMC chain for variance of parameter.

Figure 2, represents trace plot that displays the evolution of a specific parameter from two different Markov Chain Monte Carlo (MCMC) chains over iterations. The trace for the two chains is seen simultaneously in different colors. The output from this function can provide insights into the convergence of the chains. We can thus conclude that both chain1 and chain 2 have converged to a common distribution.

### The relative Risk map

The relative risk map highlights the varying incidence of malaria across different counties. In areas with an increased risk of malaria, the incidence of the disease is notably higher. These regions are marked in dark blue colors on the map. Overtime most counties kept a constant incidence of malaria.

## Relative Risk for malaria incidence

**Figure 3.** Shows how the risk of malaria is distributed among Kenyan counties from 2010 to 2020 respectively.



## 4. Conclusion

This study explored the significance of disease maps in spatial epidemiology, specifically focusing on malaria incidence in Kenyan counties from 2010 to 2020. These maps serve various purposes, including identifying high-incidence areas and optimizing healthcare resource allocation. The visual analysis revealed the dynamic nature of malaria incidence across years, with numerous counties experiencing consistently high rates. Relative risk maps highlighted the variations in risk levels among counties, singling out Kwale, Homa Bay, Kisumu, Siaya, Busia, and Mombasa as high-risk areas. To model spatial random effects, a Conditional Autoregressive (CAR) prior was used, and Bayesian analysis was carried out using WinBUGS. Furthermore, a spatio-temporal CAR model was employed to account for both spatial and temporal dependencies and capture interactions among various influencing factors.

We examined malaria in Kenya, emphasizing the significance of spatio-temporal models with time-dependent covariates. It shows that malaria incidence positively correlates with insecticide-treated bed net (ITN) use, while areas with ITN access experience lower incidence. Among three models compared using Deviance Information Criterion, the spatio-temporal model with time-dependent covariates was the most effective. Markov Chain Monte Carlo diagnostics confirmed convergence and result reliability. The study underscores the importance of proper modeling to manage malaria, with ITN access and use playing key roles.

This research contributes to a better understanding of the spatial distribution of malaria in Kenya and supports evidence-based public health interventions. The increased malaria risk coincided with lower ITN coverage, emphasizing the importance of promoting ITN use in high-risk malaria areas for effective disease prevention.

## Acknowledgments

I thank the Almighty God for the knowledge, strength and good health during the study period. I wish to extend this gratitude to my supervisors, Prof. Mwalili and Dr. Malenje for their continued support and guidance throughout this journey. I also sincerely thank my friends and family for supporting, advising and encouraging me throughout.

## Conflicts of Interest

The authors declare that they have no conflicts of interests.

## References

- [1] Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, 43: 1–20.
- [2] Congdon, P. D. (2010). *Applied Bayesian hierarchical methods*. CRC Press.
- [3] Elagali, A., Ahmed, A., Makki, N., Ismail, H., Ajak, M., Alene, K. A., Weiss, D. J., Mohammed, A. A., Abubakr, M., Cameron, E., et al. (2022). Spatiotemporal mapping of malaria incidence in sudan using routine surveillance data. *Scientific Reports*, 12(1): 1–13.
- [4] Gelman, A. and Rubin, D. B. (1996). Markov chain monte carlo methods in biostatistics. *Statistical methods in medical research*, 5(4): 339–355.
- [5] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian statistics*, 4: 641–649.
- [6] Liu, Q., Jing, W., Kang, L., Liu, J., and Liu, M. (2021). Trends of the global, regional and national incidence of malaria in 204 countries from 1990 to 2019 and implications for malaria prevention. *Journal of Travel Medicine*, 28(5): taab046.
- [7] Margaritella, N. (2020). Parameter clustering in bayesian functional pca of neuroscientific data.
- [8] Nkiruka, O., Prasad, R., and Clement, O. (2021). Prediction of malaria incidence using climate variability and machine learning. *Informatics in Medicine Unlocked*, 22: 100508.
- [9] Organization, W. H. et al. (2022). *World malaria report 2022*. World Health Organization.
- [10] Osnas, E. E., Heisey, D. M., Rolley, R. E., and Samuel, M. D. (2009). Spatial and temporal 57 patterns of chronic wasting disease: fine-scale mapping of a wildlife epidemic in wisconsin. *Ecological Applications*, (5): 1311–1322.
- [11] Otambo, W. O., Onyango, P. O., Ochwedo, K., Olumeh, J., Onyango, S. A., Orondo, P., Atieli, H., Lee, M.-C., Wang, C., Zhong, D., et al. (2022). Clinical malaria incidence and health seeking pattern in geographically heterogeneous landscape of western kenya. *BMC Infectious Diseases*, 22(1): 768.
- [12] Robert, C. and Casella, G. (2011). A short history of mcmc: Subjective recollections from incomplete data. *Handbook of markov chain monte carlo*, 49.
- [13] Saxena, R., Nagpal, B., Srivastava, A., Gupta, S., and Dash, A. (2009). Application of spatial technology in malaria research & control: some new insights. *Indian Journal of Medical Research*, 130(2): 125–132.
- [14] Schrödle, B. and Held, L. (2011). Spatio-temporal disease mapping using inla. *Environ-metrics*, 22(6): 725–734.
- [15] Thomson, M. C. and Connor, S. J. (2001). The development of malaria early warning systems for africa. *Trends in parasitology*, 17(9): 438–445.
- [16] Were, V., Buff, A. M., Desai, M., Kariuki, S., Samuels, A., Phillips-Howard, P., Ter Kuile, F. O., Kachur, S., and Niessen, L. W. (2019). Trends in malaria prevalence and health 58 related socioeconomic inequality in rural western kenya: results from repeated household malaria cross-sectional surveys from 2006 to 2013. *BMJ open*, 9(9): e03388.