

Investigating Spatio-Temporal Pattern of Relative Risk of Tuberculosis in Kenya Using Bayesian Hierarchical Approaches

Abdul-Karim Iddrisu¹, Abukari Alhassan², Nafiu Amidu³

¹Department of Mathematics and Statistics, University of Energy and Natural Resources, Sunyani, Ghana

²Department of Statistics, University for Development Studies, Navrongo, Ghana

³Department of Biomedical Laboratory Sciences, School of Medical and Health Sciences, University for Development Studies, Tamale, Ghana

Email: karim@aims.ac.za

How to cite this paper: Iddrisu, A.-K., Alhassan, A. and Amidu, N. (2018) Investigating Spatio-Temporal Pattern of Relative Risk of Tuberculosis in Kenya Using Bayesian Hierarchical Approaches. *Journal of Tuberculosis Research*, 6, 175-197. <https://doi.org/10.4236/jtr.2018.62017>

Received: February 8, 2018

Accepted: June 26, 2018

Published: June 29, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Proper understanding of global distribution of infectious diseases is an important part of disease management and policy making. However, data are subject to complexities caused by heterogeneities across host classes and space-time epidemic processes. This paper seeks to suggest or propose Bayesian spatio-temporal model for modeling and mapping tuberculosis relative risks in space and time as well identify risks factors associated with the tuberculosis and counties in Kenya with high tuberculosis relative risks. In this paper, we used spatio-temporal Bayesian hierarchical models to study the pattern of tuberculosis relative risks in Kenya. The Markov Chain Monte Carlo method via WinBUGS and R packages were used for simulations and estimation of the parameter estimates. The best fitting model is selected using the Deviance Information Criterion proposed by Spiegelhalter and colleagues. Among the spatio-temporal models used, the Knorr-Held model with space-time interaction type III and IV fit the data well but type IV appears better than type III. Variation in tuberculosis risk is observed among Kenya counties and clustering among counties with high tuberculosis relative risks. The prevalence of HIV is identified as the determinant of TB. We found clustering and heterogeneity of TB risk among high rate counties and the overall tuberculosis risk is slightly decreasing from 2002-2009. We proposed that the Knorr-Held model with interaction type IV should be used to model and map Kenyan tuberculosis relative risks. Interaction of TB relative risk in space and time increases among rural counties that share boundaries with urban counties with high tuberculosis risk. This is due to the ability of models to borrow strength from neighboring counties, such that nearby counties have similar

risk. Although the approaches are less than ideal, we hope that our study provide a useful stepping stone in the development of spatial and spatio-temporal methodology for the statistical analysis of risk from tuberculosis in Kenya.

Keywords

Bayesian Hierarchical, Deviance Information Criterion, Hot Classes, Heterogeneity, Markov Chain Monte Carlo, Relative Risk, Spatial, and Spatio-Temporal

1. Introduction

Studying the spatio-temporal pattern of infectious diseases is an important health problem in biomedical research. This is because infectious diseases adversely affect the health of the human population as well as of domestic and wild animals. This calls for the need to employ appropriate approaches to explain the spatio-temporal pattern of such diseases. This will enable us to have a proper understanding of the sources of infectious diseases as well as how such diseases spread nationally or globally. The data used to study the spread of infectious diseases are complicated due to heterogeneities that could be of spatial or non-spatial in nature [1] [2] [3]. Hence, methods of analysis are required to account for such complexities in order to produce valid statistical inferences [1] [2].

Various authors have proposed methods to study the spread of infectious diseases [4] [5] [6] [7] [8]. The spatial and spatio-temporal patterns of a disease can be studied using methods that account for the spatial and the spatio-temporal patterns. The methods as well, will be able to create relative risk (RR) maps that visually display spatial and spatio-temporal variations of disease risk [2] [3]. However, disease counts maps have various issues such as the modifiable areal unit problem (MAUP) [9]; which is a source of statistical bias that can affect the outcome of statistical hypothesis. One special case of MAUP is ecological or medical bias where the question of interest is whether inference can be made at the individual level from aggregate data [9].

There are various approaches to handle the MAUP [9]. For instance, one can scale up the data to ensure smoothing or averaging of data and making statistical inferences at high aggregate level. Also, one can scale down to enable inferences at lower level than that used in the analysis. Multi-scale analysis can also be employed to handle MAUP [9]. The multi-scale analysis handles spatial units that are completely matched when aggregated [10]. Another important issue is that differences in populations between regions results to differences in variation of regional estimates. We can account for population difference using hierarchical Bayesian approaches [3] [9].

Application of Bayesian methods in disease modeling and mapping has received great attention in biomedical research. For instance, Clayton and col-

leagues [11] applied the Empirical Bayes (EB) method to obtain smooth relative maps of lip cancer. They assumed the multivariate normal for the log relative risks and addressed the spatial correlation between neighboring regions by using the conditional autoregressive model [12]. However, the EB approach is not a fully Bayesian approach, since a quadratic approximation was used for the likelihood and this did not take into account the uncertainty in the estimates of the hyper-parameters in the specified model. Besag and Molié [12] first proposed the full Bayesian approach for disease modeling and mapping. The convolution prior model described in Section 3.1.1 was used to model the log relative risk. Their study revealed that the model shrunk extreme disease rates towards the mean and is also able to detect spatial association that was apparent in the raw data and noted that the fully Bayesian model yield more precise parameter estimates than that produced by EB approach.

Several spatio-temporal models have been developed and applied to count data. Bernardineli and colleagues [4] approach assumed that the count data follow the Poisson distribution where the log of the relative was the focus of modeling. Waller and colleagues [5] proposed a spatio-temporal model which consist of some components from the [12] approach. This model uses the convolution prior and allows each time point to have separate spatial and non-spatial effect. They assumed that the risk factors are constant over time and that disease counts followed a Poisson distribution. Waller and colleagues [5] model was applied to lung cancer deaths in 88 Ohio counties for the year 1968-1988. Each year was modeled as a separate time point. Their study revealed that lung cancer deaths increase and there is spatial clustering and variation of relative risks over time. There was increasing evidence of clustering of relative risks among the high rate counties, but with higher rates increasing and lower rates were constant. This suggests that there is increasing variability (heterogeneity) in relative risks over the study period [13]. Knorr-Held and colleagues [14] proposed a spatio-temporal model which is composed of both the convolution prior on space and also a similar prior for temporal trends. Their approach is an extension of the [5] model by assuming that the spatial terms were constant over time. This formulation allows for exploration of additional varying risk factors within each year. This model was applied to the same Ohio lung cancer data, but it was not clear that it revealed additional features of the data [13].

Iddrisu and Amoako [3] paper focus was on studying the spatial pattern of relative risk of tuberculosis in Kenya with main focus on finding suitable spatial model for modeling and mapping such relative risks. In this paper, we investigate spatio-temporal pattern of the RR of TB in Kenya and then proposed a best fitting Bayesian hierarchical approach for studying the spatio-temporal pattern of tuberculosis in Kenya.

We introduce the data used in this paper in Section 2. We discuss the methods used in Section 3. Section 3.1 discusses spatial distribution of diseases with much focus on the conditional autoregressive (CAR) model formulation for modeling

spatial distribution of diseases [12]. The CAR model allows us to explain clustering of disease risk between neighboring regions/counties/countries. This discussion is followed by a description of the Besag, York and Mollié (BYM) [12] model formulation in Section 3.1.2. The BYM formulation allows us to capture both heterogeneity (variability) and clustering of diseases risk simultaneously. The discussions on the CAR and the BYM formulation are necessary because they form the components of the spatio-temporal models adopted in this paper. The methodologies of the spatio-temporal models are discussed in Section 3.2, and then methods applied to the Kenya TB data and the results presented in Section 4. In Section 5, we give a discussion and concluding remarks in Section 6.

2. Data Description

The data used in this paper are obtained from Kenya DHS. The map of Kenya showing the counties is displayed in **Figure 1**. Kenya is located in the Eastern part of Africa and is divided into 8 provinces and 47 administrative counties. Kenya share borders with Tanzania at the south, Uganda at the West, Ethiopia at the North, Somalia at the North-East and Southern Sudan at the North-West. The data contains records of Kenya's population size, tuberculosis cases, and some potential determinants of tuberculosis for each period from 2002-2009 and for each 67 districts. To study the risk of TB infection in each county, the data from the 67 districts were aggregated to provide county level summaries. Some of the determinants of TB that were recorded are: HIV prevalence, poverty prevalence, illiteracy, population less than 5 km to health facility, firewood usage, altitude, and mean house hold size. Summaries of all variables that constitute the data are presented in **Table 1**. The mean and the median estimates of TB cases and illiteracy revealed that most of their values are concentrated at the high scale. Following the same analogy, the HIV prevalence, proportion of poor people in the population and the mean house hold size variables have almost

Table 1. Summary statistics of explanatory variables.

Variables	No. of counties	Mean	SD	Median	Min	Max	95% CI
TB Cases	47	17830	22348.07	12531	1348	149600	(7200, 20560)
HIV Prevalence (%)	47	4.289	2.797	3.800	1.000	16.430	(2.950, 4.700)
Proportion of poor	47	0.5196	0.184	0.5013	0.1157	0.9434	(0.3778, 0.6369)
Illiteracy (%)	47	24.47	19.958	16.00	2.80	77.30	(12.10, 29.80)
House hold 5km away from Hospital (%)	47	77.76	16.399	80.80	19.20	99.00	(72.05, 86.72)
Firewood (%)	47	78.52	20.175	84.60	1.80	96.70	(74.95, 90.65)
Altitude (m)	47	1361	602.2214	1432	151	2274	(1138, 1813)
Mean House Hold Size	47	5.383	0.799	5.250	3.800	6.900	(4.775, 6.050)



Figure 1. Map of Kenya showing counties and borders.

equal values at the low and high scale. These variables are almost symmetric or normally distributed. Finally, the variables, percentage of people who are at 5km distance away from hospital, and the percentage of those who use firewood have most of their data values concentrated at the low scale.

Table 2 presents summaries of Kenya's population size for 2002-2009. The mean county level population size stood out to be 685586(407424,869752) in 2002 increasing to 838793(506907,1031868) in 2009. The mean and the median showed that most of the data values for 2002-2009 are concentrated at the high scale. **Table 3** presents summaries of TB cases in Kenya for 2002-2009. The mean county level TB cases is estimated at 1747(679.5,2006.5) in 2002 and increased to 2006 at an estimated value of 2452(1044,2844). TB cases decreased from 2007-2008 with corresponding values of 2416(966.5,2786.5) and 2291(850.5,2623.0) respectively. TB cases slightly increased in 2009, estimated at 2346(896,2,724).

3. Methods

In this section we discuss the Bayesian spatial and spatio-temporal models used

Table 2. County Level Population Size Summaries for 2002-2009.

Year	N	Mean	SD	Median	Min	Max	95% CI
2002	47	685,586	435,819.5	604,298	83,985	3,034,397	(407424, 869752)
2003	47	705,776	450,910.0	625,506	86,443	2,600,859	(420676, 890554)
2004	47	727,220	466,775.6	647,886	89,058	2,710,706	(434932, 912918)
2005	47	747,631	481,824	669,403	91,549	2,815,838	(449244, 934009)
2006	47	768,909	497,822	691,637	94,150	2,924,309	(462472, 956223)
2007	47	791,147	514,520	714,590	96,851	3,034,397	(475354, 979870)
2008	47	814,422	531,548	738,321	99,662	3,146,303	(490594, 1005057)
2009	47	838,793	548,905	762,870	102,593	3,260,124	(506907, 1031868)

Table 3. Summarises Statistics of County Level TB Cases From 2002-2009.

Year	N	Mean	SD	Median	Min	Max	95% CI
2002	47	1747	2409.296	1,053	111.0	15979.0	(679.5, 2006.5)
2003	47	2028	2761.937	1,355	154.0	18360.0	(746.5, 2386.0)
2004	47	2249	2954.713	1,497	138.0	19871.0	(928.5, 2620.0)
2005	47	2302	2913.039	1,549	124.0	19487.0	(938.5, 2737.5)
2006	47	2452	2934.737	1,757	172	19,472	(1044, 2844)
2007	47	2416	2836.123	1,988	177.0	18901.0	(966.5, 2786.5)
2008	47	2291	2790.387	1,676	223.0	18589.0	(850.5, 2623.0)
2009	47	2346	2843.675	1,700	249	18,984	(896, 2724)

in disease modeling and mapping.

3.1. Bayesian Hierarchical Spatial Model for Disease Mapping

Spatial data are associated with a given location on the surface of the earth where such models allow for *borrowing* of strength between neighboring locations. In this way, neighboring locations/countries/regions/counties will have similar risks while distant counties are expected to show variation in risks. Waldo Tobler's first law of spatial analysis states that "*everything is related to everything else but near-by things are more related than distant things*" [15].

3.1.1. Conditional Autoregressive (CAR) Model

The conditional autoregressive (CAR) models were introduced by Besag and Molié [12] and they are used to identify and detect clustering of diseases risk. In these models, risks of disease at any given area are influenced by the risk in the neighboring areas. The CAR model is often referred to as the structured or the correlated heterogeneity (CH) model. This is because estimation of risk in any given area is to some extent affected by the diseases risk at neighboring areas [16]. In the CAR model, the distances or boundaries between the regions/counties are often used in determining the neighborhood properties [17].

Application of the CAR model can be found in [18] [19] [20]. Let $S = \{1, 2, \dots, n\}$ represents the area to be studied. Let $N_i = \{j \in S : i \in j\}$ denote the set of all regions. Let $v_i, i \in S$ be a random variable. We define the corresponding random field v as the vector $v = (v_1, v_2, \dots, v_n)$. In the normal CAR model, it is often assumed that each observation of the outcome variable v_i has a conditional distribution defined as

$$v_i | v_{j \neq i} \sim N \left(\sum_{j \in N_i} \Phi_{ij} v_j, \tau_i^2 \right).$$

These are full conditionals where Φ_{ij} is the weight of each observation on the mean of v_i and also denotes the spatial dependence parameter. The Φ_{ij} is non zero only if $j \in S$. Conventionally, we set $\Phi_{ij} = 0$ since we do not want to regress any observation on itself. Hence no region is a neighbor of itself. The v_j denotes a vector of all observation except v_i . Note that v_i depends only on a set neighbours v_j only if location j is a neighborhood set N_i of v_i . The τ_i^2 is a potential unique variance for v_i . For instance, if state i has M neighbours and $\Phi_{ij} = \frac{1}{M}$ for every state that is a neighbour, and $\Phi_{ij} = 0$ otherwise, then the conditional expectation of a state's observation is the mean of all neighbours observations [20].

The CAR model is defined by mean and covariance function [21]. Assuming that each conditional distribution is Gaussian, we will need the mean and the variance-covariance to define the CAR model. The mean and variance-covariance are respectively defined as

$$E[v_i | v_{j \neq i}] = \mu_i + \sum_{j \in N_i} \Phi_{ij} [v_j - \mu_j]$$

and

$$\text{var}(v_i | v_j) = \tau_i^2$$

Detailed information on the conditional probability density function of a CAR random variable v_i can be found in [3] [12]. Thus, a conditional autoregressive model v has a probability density function defined as

$$v_i | v_{j \neq i} \sim N \left[\mu_i + \rho \sum_{j \in N_i} \Phi_{(ij)} (v_j - \mu_j), \tau_i^2 \right], i \in S$$

and the joint probability density is

$$v \sim N(\mu, B^{-1} \Sigma_D). \quad (1)$$

The necessary and sufficient condition for (1) to be a valid joint probability density function is that its covariance matrix should be symmetric and positive definite (eigenvalues $\lambda_i > 0, i, \dots, n$) [20]. The v_i is a Gaussian random variable with Σ_D as a symmetric matrix (See [3] [18]).

The τ_v^2 controls the overall variability of v_i and ρ describes the overall effect of spatial dependence. The value of ρ is should be chosen carefully [20].

Choose ρ such that $\Sigma_{D_v}^{-1}$ is non singular and preferably with $\rho \in \left(\frac{1}{\lambda_1}, \frac{1}{\lambda_n}\right)$.

Detailed description of how to choose ρ values and parameter estimation in the CAR model can be found in [3] [20]. Application of the CAR model to the Kenya TB data can be found in [3].

3.1.2. The Besag, York and Mollié (BYM) Model

The BYM is also a spatial model for risk smoothing which have appeared in literature and has received much attention [12]. The BYM model is composed of the CAR model component v_i discussed in Section 3.1.1 and the unstructured heterogeneity (UH) component u_i with normal distribution [3] [12]. The component u_i the prior distribution for the area-specific random effect. The BYM model was introduced by [11] and latter developed by [12]. The BYM or convolution model is defines as $\eta_i = \mu + u_i + v_i$ and the data y_i follow the Poisson distribution as $y_i \sim \text{Poisson}(E_i \exp(\eta_i))$, where $\mu_i = E_i \exp(\eta_i)$ is the mean of the Poisson distribution. The linear link function $\eta_i = X' \beta + u_i + v_i$. The log relative risk $\log(\mathcal{G}_i) = \eta_i$. Therefore, the relative risk for the i area is defined by $\mathcal{G}_i = \exp(X' \beta + u_i + v_i)$. The log log-link function is defined as

$$\log(\mu_i) = \log(E_i) + \exp(X' \beta + u_i + v_i),$$

where β, E and \mathcal{G} are vectors of the covariate, the associated parameters, the expected number of cases, and the relative risks of TB prevalence respectively. The u_i is the county level random effect capturing the residual log RR of disease in county i . The u_i (UH) is sometime thought of as a latent variable which captures the effect of unknown or unmeasured area level covariates and v_i has a CAR model structure. Detailed information on formulation of the BYM model, parameter estimation and application to the Kenya TB data can be found in [3] [12]. We now turn our attention to the spatio-temporal models, the focus of this paper.

3.2. Bayesian Hierarchical Spatio-Temporal Models Disease Mapping

Many disease mapping models are restricted to identification of spatial heterogeneity and clustering of diseases risk which are in fact constrained to a single time period. However, most data in public health are often in the form of time window for many years. Therefore, there is the need to consider models which account for spatial and temporal pattern of diseases risks. Several methods have been proposed to account for spatial and temporal pattern of diseases risks [4] [5] [12] [14].

We now present the structure of data in space and time. Consider the case where a given region of interest is divided into N areas (regions, districts, counties or municipalities) indexed by $i = 1, 2, \dots, n$. Let the temporal dimension be indexed by $t = 1, 2, \dots, T$, representing each period of time under study. Let n_{it} be the number of persons-times at risk in region i at period t and y_{it} be the

corresponding observed cases which are counts TB2. The observed data y_{it} depends on N_{it} , the number of people at risk in region i and period t in the study population observed. Let N_{it}^s be the number of people at risk in the standard population, y_{it}^s be the observed TB cases in the standard population and C_{it}^s be the crude rate of TB cases in the standard population. Therefore, the crude rate for region i and period t , C_{it}^s is defined by $C_{it}^s = \frac{y_{it}^s}{N_{it}^s}$. It follows

that the number of TB cases expected in region i and period t , E_{it} is defined by $E_{it} = C_{it}^s N_{it} = \frac{y_{it}^s}{N_{it}^s} N_{it}$. Therefore, the overall crude rate of TB cases is defined by

$C = \sum_i^n \sum_t^T \frac{y_{it}^s}{N_{it}^s}$ and the overall number of expected TB cases is defined by

$E = \sum_i^N \sum_t^T C_{it}^s N_{it} = \sum_i^N \sum_t^T \frac{y_{it}^s}{N_{it}^s} N_{it}$. We assumed that y_{it} follows the Poisson

distribution with expectation $E(y_{it}) = \mu_{it} = E_{it} \theta_{it}$, where θ_{it} denotes the disease risk in region i , at period t . The distribution of y_{it} is

$y_{it} \sim \text{Poisson}(E_{it} \exp(\eta_{it}))$, where $\eta_{it} = \mu + Z_i + A_t + ZA_{it} + u_{it}$ is a linear predictor, μ denotes the grand mean, Z_i the main effect of region i , A_t the temporal trend effect in period t , ZA_{it} is interaction of risk in space and time and u_{it} is the unstructured random effect. The contribution of a given term may serve to increase or decrease the risk of disease. The intercept or μ gives a background amount of risk shared by all regions and periods. Most often, an unstructured extra variability term u_{it} is included in the model so as to capture the overall effect of the other unaccounted and unobserved effects. The random effect u_{it} is defined as

$$u_e | \tau_u^2 \stackrel{iid}{\sim} N(0, \tau_u^2), e \in \{i, t, it\}$$

and A_t is often modeled as a structured random effect.

In the following sections, we describe the spatio-temporal models used and parameter estimation. Markov Chain Monte Carlo via Gibbs sampling is used to obtain parameter estimates under each model [16].

3.2.1. Spatio-Temporal Model I

This spatio-temporal model is based on the mode proposed by Bernardineli and colleagues [4]. This is a kind of parametric model in which the risk in a given region is considered as a linear or quadratic function of time. This model takes into account spatio-temporal interaction where temporal trend in risk may differ for spatial locations and may even have a spatial structure [16]. All the temporal trends are assumed to be linear and information is shared in both space and time [16]. This approach is used to investigate statistical linear rise in the reported TB prevalence in Kenya from 2002 to 2009.

Assume that the TB counts $y_{it} \sim \text{Poisson}(E_{it} \exp(\eta_{it}))$. According to Bernardineli and colleagues [4], the linear predictor η_{it} is defined by

$\eta_{it} = \mu + u_i + v_i + (\varpi + \delta_i) \times \mathfrak{T}_t$, where $u_i + v_i$ follows the BYM specifications [12] (See Section 3.1.2) and $\varpi \mathfrak{T}_t$ linear trend in time \mathfrak{T}_t , δ_i is the interaction random effect between region and time and ϖ is the overall linear time trend. The log relative risk for area i and period t is $\log(\mathcal{G}_{it}) = \eta_{it}$. Therefore the relative risk of disease is given by

$$\mathcal{G}_{it} = \exp(\eta_{it}) = \exp(\mu + u_i + v_i + (\varpi + \delta_i) \times \mathfrak{T}_t).$$

The log of the Poisson mean

$$\mu_{it} = E_{it} \exp(\mu + u_i + v_i + (\varpi + \delta_i) \times \mathfrak{T}_t)$$

is therefore given by

$$\log(\mu_{it}) = \log(E_{it}) + \mu + u_i + v_i + (\varpi + \delta_i) \times \mathfrak{T}_t.$$

3.2.2. Parameter Estimation

Given that $y_{it} \sim \text{Poisson}(E_{it} \exp(\eta_{it}))$ with likelihood function denoted by

$$\Pr(\mathbf{y}, \mathbf{E}, \mathbf{g} | \mathbf{u}, \mathbf{v}, \boldsymbol{\delta}, \varpi, \tau_u^2, \tau_v^2, \tau_\delta^2). \quad (2)$$

The prior distribution $P(\mathbf{u})$ of \mathbf{u} follows a normal distribution defined as

$$\Pr(\mathbf{u}) = \left(\frac{1}{2\pi}\right)^{n/2} \left(\frac{1}{\tau_u}\right)^n \exp\left(-\sum_{i=1}^n \frac{u_i^2}{2\tau_u^2}\right)$$

and prior distribution $P(\mathbf{v})$ of \mathbf{v} has CAR structure (Section 3.1.1). Also, δ_i is modeled as a CAR structure with prior distribution denoted by $P(\boldsymbol{\delta})$ and $\varpi \sim N(0, 0.005)$ with prior distribution $P(\varpi)$. The overall mean is defined as $\mu \sim N(0, 0.01)$. Therefore, the posterior distribution is defined as

$$\begin{aligned} & \Pr(\mathbf{u}, \mathbf{v}, \boldsymbol{\delta}, \varpi, \tau_u^2, \tau_v^2, \tau_\delta^2 | \mathbf{y}, \mathbf{E}, \mathbf{g}) \\ & \propto \Pr(\mathbf{y}, \mathbf{E}, \mathbf{g} | \mathbf{u}, \mathbf{v}, \boldsymbol{\delta}, \varpi, \tau_u^2, \tau_v^2, \tau_\delta^2) \times \Pr(\mathbf{u}) \Pr(\mathbf{v}) P(\boldsymbol{\delta}) P(\varpi). \end{aligned} \quad (3)$$

One limitation of the model is the assumption of a linear time trend in each region. This limitation is resolved by [8] model (discuss in Section 3.2.5). Parameters estimation from Equation (3) was carried out using MCMC via Gibbs sampling.

3.2.3. Spatio-Temporal Model II

The spatio-temporal Approach used in this section was developed by Waller and colleagues [21]. In this model, spatial effects are observed as a set of spatial models, one for each period of time, with almost no relation between them, except possibly for some restriction in their precision parameters [16]. Under this model, the hierarchical specification is applied to each time point separately [12] [16]. This model does not have a single spatial main effect and does allow spatial pattern at each time point to be completely different [2] [16].

Assume that $y_{it} \sim \text{Poisson}(E_{it} \exp(\eta_{it}))$, where $\eta_{it} = u_i^{(t)} + v_i^{(t)}$ and $\mu_{it} = E_{it} \exp(\eta_{it})$ is the Poisson mean. The log relative risk for area i and period t is $\log(\mathcal{G}_{it}) = \eta_{it}$. Therefore relative risk of disease is given by

$$\mathcal{G}_{it} = \exp(\eta_{it}) = \exp(u_i^{(t)} + v_i^{(t)}), \quad (4)$$

where for each period t , the model term $u_i^{(t)} + v_i^{(t)}$ follows the BYM specification (See Section 3.1.2) with different precision parameter $\tau_u^{(t)}$ and $\tau_v^{(t)}$ for each period of time. The log of the Poisson mean $\mu_{it} = E_{it} \exp(u_i^{(t)} + v_i^{(t)})$ is therefore given by

$$\log(\mu_{it}) = \log(E_{it}) + u_i^t + v_i^t \quad (5)$$

The $u_i^{(t)}$ and $v_i^{(t)}$ are respectively uncorrelated and correlated heterogeneity terms which may vary with time. This approach results in spatio-temporal model where the spatial dimension is nested within time; thus in effect a spatial model is fitted for each period.

3.2.4. Parameter Estimation

Given that $y_{it} \sim \text{Poisson}(E_{it} \exp(\eta_{it}))$ with likelihood function $\Pr(y, E, \mathcal{G} | u, v, \tau_u^2, \tau_v^2)$, the prior distribution $\Pr(u)$ of u follows a normal distribution and prior distribution $\Pr(v)$ of v has CAR structure (See Section 3.1.1). Therefore, the posterior distribution is defined as

$$\Pr(u, v, \tau_u^2, \tau_v^2 | y, E, \mathcal{G}) \propto \Pr(y, E, \mathcal{G} | u, v, \tau_u^2, \tau_v^2) \Pr(u) \Pr(v). \quad (6)$$

Estimation of parameters from Equation (6) was achieved through Bayesian MCMC via Gibbs sampling.

3.2.5. Spatio-Temporal Model III

Knorr-Held and Rasser [8] proposed this approach which is a type of smooth temporal evolution model where the evolution of the estimated risk in each region is a smooth function of time. Knorr-Held and Rasser [8] proposed this model to overcome the limitation suffered by the Bernardineli and colleagues [4] model. Assume that $y_{it} \sim \text{Poisson}(E_{it} \exp(\eta_{it}))$. Knorr-Held and Rasser defined the linear predictor η_{ij} as

$$\eta_{it} = \mu + u_i + v_i + \mathfrak{I}_t + \psi_{it},$$

where the term $u_i + v_i$ follows the BYM specification. The parameter \mathfrak{I}_t represents an unstructured or structured temporal effect and the parameter ψ_{it} is the space-time interaction. The log relative risk for area i and period t is $\log(\mathcal{G}_{it}) = \eta_{it}$. Therefore the relative risk of disease is given by

$$\mathcal{G}_{it} = \exp(\mu + u_i + v_i + \mathfrak{I}_t + \psi_{it}).$$

The log of the Poisson mean

$$\mu_{it} = E_{it} \exp(\mu + u_i + v_i + \mathfrak{I}_t + \psi_{it})$$

is therefore given by

$$\log(\mu_{it}) = \log(E_{it}) + \mu + u_i + v_i + \mathfrak{I}_t + \psi_{it}.$$

It should be noted that u, v and \mathfrak{I} are the main effects whiles ψ is the space-time interaction term. This model is used to study smooth temporal evolution of the estimated relative risk of TB prevalence in Kenya in each region at

given point in time.

3.2.6. Parameter Estimation

Given that $y_{it} \sim \text{Poisson}(E_{it} \exp(\eta_{it}))$ with likelihood function

$$\Pr(\mathbf{y}, \mathbf{E}, \boldsymbol{\vartheta} | \mathbf{u}, \mathbf{v}, \boldsymbol{\psi}, \boldsymbol{\vartheta}, \tau_u^2, \tau_v^2, \tau_\psi^2, \tau_\vartheta^2),$$

the prior distribution $\Pr(\mathbf{u})$ of \mathbf{u} has a normal distribution and prior distribution $\Pr(\mathbf{v})$ of \mathbf{v} has CAR structure (See chapter 3.1.1). According to [8], if the main temporal random effect $\boldsymbol{\vartheta}_t$ assumes unstructured random effect, then its prior distribution would be $\boldsymbol{\vartheta}_t | \tau_\vartheta^2 \sim N(0, \tau_\vartheta^2)$ and if it assumes structured random effect, then its prior density follows a first order random walk defined by

$$\Pr(\boldsymbol{\vartheta} | \tau_\vartheta^2) \propto \exp \left[-\frac{\tau_\vartheta^{-2}}{2} \sum_{t=2}^T (\boldsymbol{\vartheta}_t - \boldsymbol{\vartheta}_{t-1})^2 \right]. \quad (7)$$

According to [8], prior specification for the interaction term $\boldsymbol{\psi}$ depends on the spatial and temporal main effect which are assumed to interact. Different types interactions ψ_{it} were classified by [22] with prior distribution denoted by $\Pr(\boldsymbol{\psi})$ and precision variance denoted by τ_ψ^2 . Therefore, the posterior distribution for the relative risk $\boldsymbol{\vartheta}$ is defined by

$$\begin{aligned} & \Pr(\mathbf{u}, \mathbf{v}, \boldsymbol{\psi}, \boldsymbol{\vartheta}, \tau_u^2, \tau_v^2, \tau_\psi^2, \tau_\vartheta^2 | \mathbf{y}, \mathbf{E}, \boldsymbol{\vartheta}) \\ & \propto \Pr(\mathbf{y}, \mathbf{E}, \boldsymbol{\vartheta} | \mathbf{u}, \mathbf{v}, \boldsymbol{\psi}, \boldsymbol{\vartheta}, \tau_u^2, \tau_v^2, \tau_\psi^2, \tau_\vartheta^2) \times \Pr(\mathbf{u}) P(\mathbf{v}) \Pr(\boldsymbol{\psi}) P(\boldsymbol{\vartheta}). \end{aligned}$$

The interaction type depends on which of the two possible type of temporal effects (unstructured or structured) interacts with the two main effects (u_i and v_i). Knorr-Held and Rasser defined four interaction types. Each of the four type of interactions has different prior interrelationships involving the interaction term ψ_{it} TB1.

1) **Interaction type I:** If the unstructured main effects ($\boldsymbol{\vartheta}_t$ and u_i) are expected to interact, then the distribution of the interaction parameter ψ_{it} is defined as

$$\Pr(\boldsymbol{\psi} | \tau_\psi^2) \propto \exp \left[-\frac{\tau_\psi}{2} \sum_{i=1}^n \sum_{t=1}^T (\psi_{it})^2 \right].$$

This may be considered as an independent unobserved covariate for each combination of region and period (i, t) , thus without any structure TB1, TB2. On the other hand, if spatial and temporal main effects are present in the model, then the interaction effect only denote independence in the deviations from them. The main effects can cause contribution to risk in neighboring regions or in consecutive period of time to be highly correlated. This is a global space-time heterogeneity effect and it is often modeled as $\psi_{it} \sim N(0, \tau_\psi^2)$. This interaction type has independent prior with no structure in space-time interaction TB1, TB2.

2) **Interaction type II:** This interaction effect is distributed as a random walk independently of other counties if we modelled $\boldsymbol{\vartheta}_t$ as a random walk [14]. The

prior distribution for this interaction is defined by [8] [10] [14]

$$[\psi | \tau_\psi] \propto \exp \left[-\frac{\tau_\psi}{2} \sum_{i=1}^n \sum_{t=2}^T (\psi_{it} - \psi_{i,t-1})^2 \right]$$

This type of interactions has no structure in space TB4. This implies that each region has a specific evolution structure that is independent of that in the neighbouring region TB2, TB1.

3) **Interaction type III:** If we assumed that the unstructured temporal main effect (\mathfrak{I}_t) and the spatially correlated or structured main effect (v_i) interact, then the interaction effect parameter $\psi_t = (\psi_{1t}, \dots, \psi_{nt})$, $t = 1, \dots, T$ follows an independent Intrinsic autoregressive distribution defined as [8] [10] [14]

$$[\psi | \tau_\psi] \propto \exp \left[-\frac{\tau_\psi}{2} \sum_{t=1}^T \sum_{i \sim \ell} (\psi_{it} - \delta_{it})^2 \right]$$

This interaction is assumed to have a spatial structure for each period, independent of adjacent periods (its neighbors in time). This interaction type is analogous to the clustering effect, which is often modeled as a CAR distribution (section 3.1.1) for each period [1] [4]. Here we implicitly assumed that each specific region may have a slight deviation from the global trend, but that this deviation is likely to be identical to that in the neighboring regions, while at the same time, independent of that in that in the previous period of time [1] [8].

4) **Interaction type IV:** Type IV is completely dependent on space and time theoretically [14]. Hence the effect can no longer be factorized into independent blocks if \mathfrak{I}_t is modeled as a random walk allowing interaction with the structured main effect (v_i). [8] defined type IV interaction as

$$[\psi | \tau_\psi] \propto \exp \left[-\frac{\tau_\psi}{2} \sum_{t=2}^T \sum_{i \sim \ell} (\psi_{it} - \psi_{it} - \psi_{i,t-1} + \psi_{\ell,t-1})^2 \right]$$

[8] stated that, type IV is the most interesting type of interaction that occurs when deviation from the global trends are highly correlated with their neighbors, both in space and time. Here, hidden factors whose effects exceed the limits of one or more regions and also persistent for one or more period of time can be modeled. This is also an efficient way of obtaining information from data, particularly in the case of rare diseases or less populated regions, since the risk estimation for the region-period is not performed on the basis of only locally observed data but also on that in the neighboring regions and periods TB2.

The hyper-prior distribution for $\tau_{\mathfrak{I}}^2$ and τ_ψ^2 are modeled as gamma distribution. Estimation of all parameters was achieved with Bayesian MCMC via Gibbs sampling.

4. Results

Best fitting spatio-temporal model to Kenya TB data was selected from the above candidate models based on their respective model's DICs and pDs presented in **Table 4**. Note that these values are subject to Monte Carlo error, which is difficult to quantify. We have therefore chosen a very long run of which convergence

was reached at 1,200,200 after a burn-in period of 100,000 and thinning of every 30th element of the chain for each model shown in **Figure 2**.

From **Table 4**, \bar{D} is the mean of the posterior deviance, pD is the effective number of parameters and $DIC = \bar{D} + pD$ proposed by [22]. Among the spatio-temporal models presented in **Table 4**, the model with the lowest DIC (4194.30) and lowest pD (375.306) is the Knorr-Held and Rasser [8] model. We therefore recommend model (7) as the best fitting space-time model to Kenya TB data for 2002-2009. **Table 5** summaries our effort to identify appropriate interaction type. The **Table 5** presents deviance summary of the interaction types after MC convergence at 1,200,000 and a burn-in period of 100,000 for each model. The models fit with interaction type III and IV fit the data well but type IV seems better than type III since it yields the lowest pD (362.494) and DIC (419.410). We now consider analysis and interpretation of results under the model with interaction type IV since it gives best fit of the Kenya TB data. In **Table 6**, the overall mean relative risk μ is insignificant and the precision variance parameters τ_v^2 and τ_u^2 indicates significance of clustering and heterogeneity of relative over the studied period respectively. The precision of the variance parameter indicates significance of TB relative risk interaction in space-time.

Figure 3 displays the distribution of the posterior relative risk for 2002-2009. Generally, the pattern does not change much over the study period. However, some counties have interesting time trends. For instance, in **Figure 3**, the two

Table 4. Spatio-Temporal models Deviance Summaries.

Model indicators	Bernardinelli <i>et al.</i> , 1995	Waller <i>et al.</i> ,1997	Knorr-Held <i>et al.</i> , 20000
\bar{D}	135492	4039.320	3818.720
pD	487.536	594.296	375.306
DIC	135979	4633.62	4194.03

Table 5. Interaction Types Deviance Summaries.

Model indicators	Type I	Type II	Type III	Type IV
\bar{D}	3820.060	3818.600	3827.920	3826.910
pD	376.637	374.377	363.510	362.494
DIC	4196.700	4192.970	4191.430	4189.410

Table 6. Interaction Type IV Posterior Statistics.

Model indicators	estimates	95% Credible Interval
μ	-0.22	(-0.54,0.70)
τ_v^2	10.8	(1.39,50.10)
τ_u^2	9.17	(3.57,28.60)
τ_{ψ}^2	11.3	(9.55,13.20)

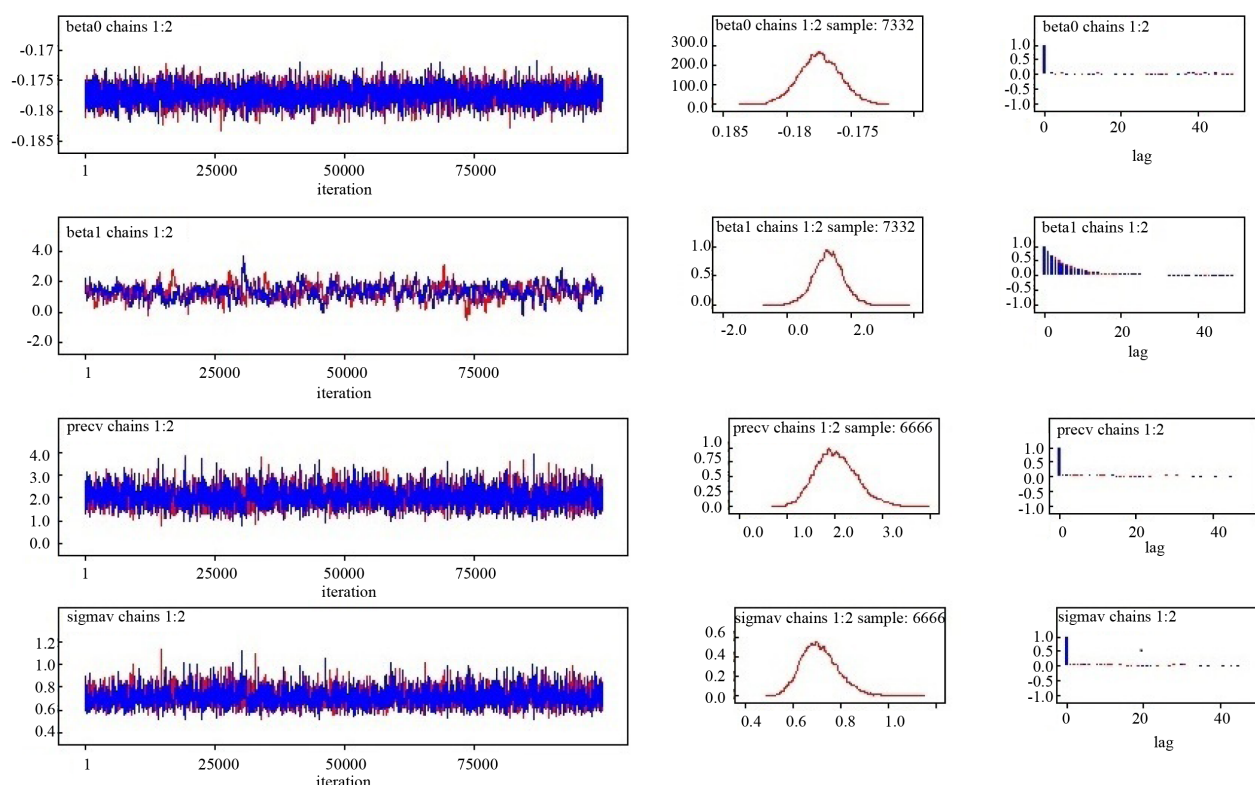


Figure 2. Convergence diagnostics of Markov Chain Carlo.

adjacent counties (Nairobi and Kiambu) in the central part of Kenya show opposite trend in disease risk. This may be due to the fact that type IV interaction borrows strength from neighboring counties, hence, the decreasing trend in Nairobi county causes the estimated increase in Kiambu which is less populated than Nairobi, to be less pronounced. Again, high risk of TB prevalence is observed in the North, West, North-West and the Central counties and low risk in the South-East counties for 2002-2009.

Figure 4 displays decreasing temporal trend of posterior relative parameters for some highly urbanized counties such as Mombasa and Nairobi. In contrast, pronounced increasing trends were observed for most rural counties such as Nandi and Kiambu. More information on temporal trend behavior of posterior relative risk for the rest of the counties can be found in **Appendix**. The left panel of the **Figure 5** displays a slight increasing temporal trend from 200-2004, slight decrease from 2004-2005, increase in 2006, a slight decrease in 2007-2008 and slight increase in 2009. The temporal trend effect does not change much for 2002-2009. The right panel of the **Figure 5** displays area relative risk of the type IV interaction. Again high risk of TB is observed in the North, West, North-West and Central counties of Kenya and low risk in the South-East counties. Interaction type IV was identified by [23] as the best fitting spatio-temporal model for modeling and mapping salmonellosis counts in cattle in Switzerland, 1991-2008. Knorr-Held fitted the four different types of interaction to the 21

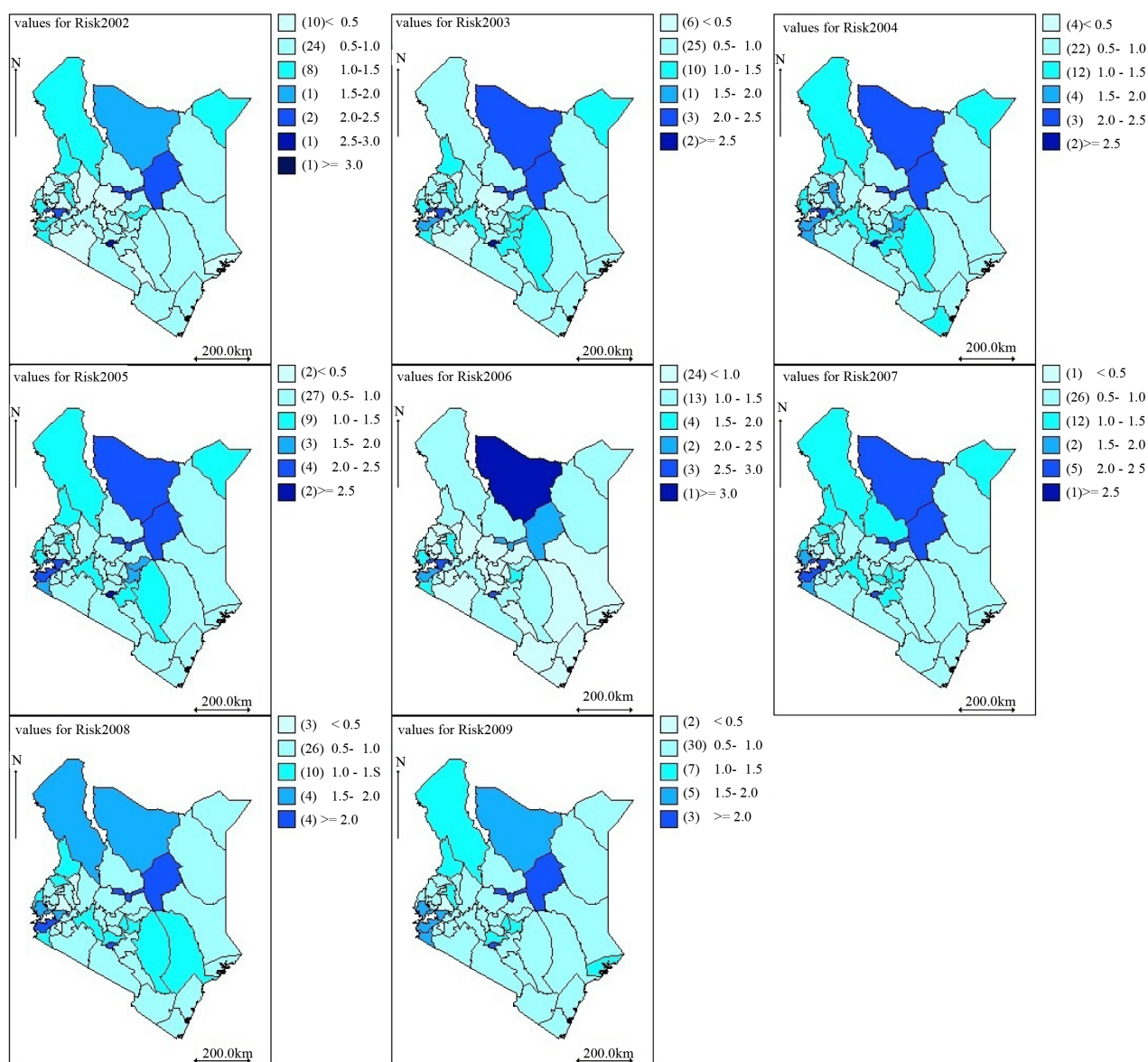


Figure 3. Type IV interaction posterior mean of the relative risk maps for 2002-2009.

years Ohio respiratory cancer dataset and found that interaction type II was appropriate; offering lowest deviance [16]. Depending on the data you are dealing with, any of the four interaction types can yield best fit for the data [8]. Spatio-temporal Parametrization of log relative risk can take a variety of forms and it is not clear yet which form is most appropriate [16].

5. Discussion

This paper explores application of spatio-temporal models used in disease modeling and mapping of TB relative risk in Kenya. These models were fitted to Kenya's TB prevalence data from 2002-2009. Markov Chain Monte Carlo via Gibbs sampling was used for simulation of parameters from posterior distributions. Rubin and Gelman convergence diagnostics test was used to confirm

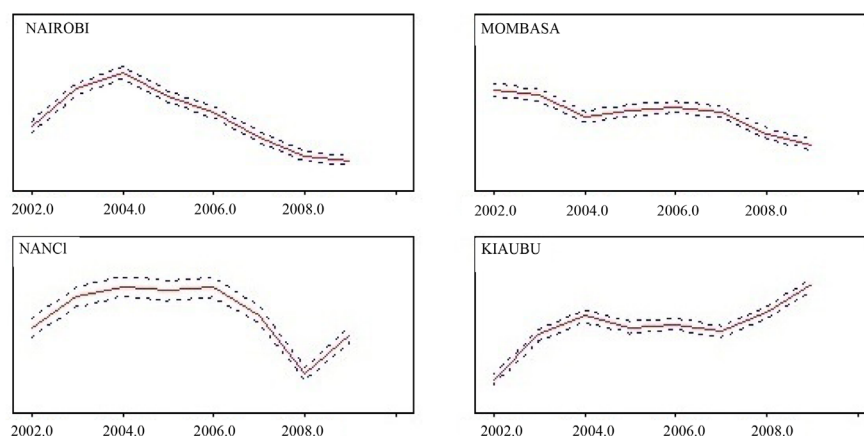


Figure 4. Type IV interaction posterior mean of the relative risk for 2002-2009.

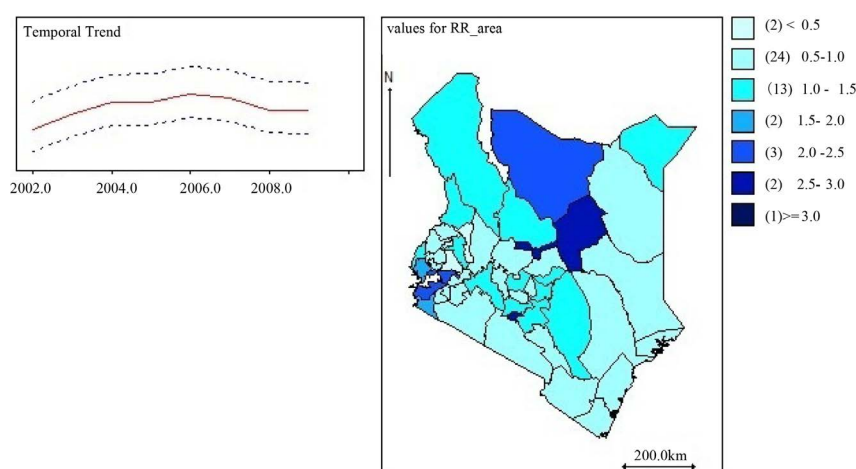


Figure 5. Type IV temporal trend and area posterior mean relative risk.

convergence of the Markov Chain. Thinning the Markov Chain and the over-relax algorithm though slow the speed of the MCMC but significantly reduces autocorrelation and number of iterations. Long-run MCMC iterations and high thinning sample size k is required for spatio-temporal models used in fitting Kenya's TB data. The DIC of each model were compared and best model selected from the set of candidate models used in fitting Kenya's TB prevalence data. Among the spatio-temporal models considered, the model proposed by Knorr-Held and Leonaard [8] with space-time interaction type III and IV fit the data well but type IV appears better than type III. The variation in TB risk is observed among Kenya counties and clustering among counties with high TB relative risk (RR). The prevalence of HIV is identified as the dominant determinant of TB. We found clustering and heterogeneity of risk among high rate counties and the overall TB risk is slightly decreasing from 2002-2009. Interaction of TB relative risk in space and time is found to be increasing among rural counties that share boundaries with urban counties with high TB risk. This is as a result of the ability of models to borrow strength from neighboring counties,

such that near by counties have similar risk.

6. Conclusions

We recommend the Knorr-Held and colleagues model [8] with interaction type IV structure for modeling and mapping Kenya TB relative risk. Generally clustering of risks and elevated risks is observed in the North, West, North-West and the central counties of Kenya and low clustering and elevated risk in the South-West counties. We have discovered an interesting association between temporal trends of interaction parameters and urbanization in Kenya, which might set a framework for further epidemiological research. Modeling of risk in space and time is quite a challenging task. Although these approaches are less than ideal, we hope that our formulations provide a useful stepping stone into the development of spatio-temporal methodology for modeling and mapping Kenya's TB prevalence data.

We are satisfied that the models selected in this paper are from an appropriate class that led to the analysis of the Kenya's TB data for 2002-2009. Further research is required for a standard or acceptable distribution type for space-time interaction ψ_{ij} to be identified since comparing posterior deviance from interaction type that assumed t_j should be modeled as structured could lead to one or more deficiencies to a given interaction type.

Competing Interests

The authors declare that he has no competing interests.

Authors' Contributions

AI performed the literature review, statistical analyses, and wrote the article. AA and NA contributed in reviewing the paper contributed to the writing and reviewing the manuscript and also provided consultation regarding analysis and interpretation of findings. The final version has been approved by the authors.

Acknowledgments

The author would like to acknowledge Dr. Thomas Noel Achia, for mentoring me on application of Bayesian methods to disease modeling and mapping.

Availability of Data and Materials

The data used are from Kenya Demographic and Health Survey and can be found on

<https://dhsprogram.com/What-We-Do/Protecting-the-Privacy-of-DHS-Survey-Respondents.cfm>

Funding

The study receives no funding.

Ethics Approval (and Consent to Participate)

Ethics approval and consent to participate statements can be found on <https://dhsprogram.com/What-We-Do/Protecting-the-Privacy-of-DHS-Survey-Respondents.cfm>. Procedures and questionnaires for standard DHS surveys have been reviewed and approved by the ICF International Institutional Review Board (IRB). Additionally, country-specific DHS survey protocols are reviewed by the ICF IRB and typically by an IRB in the host country. The ICF International IRB ensures that the survey complies with the U.S. Department of Health and Human Services regulations for the protection of human subjects (45 CFR 46), while the host country IRB ensures that the survey complies with laws and norms of the nation. Before each interview or biomarker test is conducted, an informed consent statement is read to the respondent, who may accept or decline to participate. A parent or guardian must provide consent prior to participation by a child or adolescent. DHS informed consent statements provide details regarding: the purpose of the interview/test, the expected duration of the interview, interview/test procedures, potential risks to the respondent. Potential benefits to the respondent, contact information for a person who can provide the respondent with more information about the interview/test. Most importantly, the informed consent statement emphasizes that participation is voluntary; that the respondent may refuse to answer any question, decline any biomarker test, or terminate participation at any time; and that the respondent's identity and information will be kept strictly confidential.

References

- [1] López-Qulez, A. and Munoz, F. (2009) Review of Spatio-Temporal Models for Disease Mapping.
- [2] Wakefield, E., Best, N.G., Briggs, D.J., *et al.* (2000) Spatial Epidemiology: Methods and Applications. Oxford University Press, Oxford.
- [3] Lawson, A.B. and Lawson, A.B. (2001) Statistical Methods in Spatial Epidemiology. John Wiley, Hoboken.
- [4] Currie, C.S.M., Williams, B.G., Cheng, R.C.H. and Dye, C. (2003) Tuberculosis Epidemics Driven by HIV: Is Prevention Better than Cure. *Aids*, **17**, 2501-2508. <https://doi.org/10.1097/00002030-200311210-00013>
- [5] Lawson, A.B. (2013) Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. CRC Press, Boca Raton.
- [6] Lawson, A.B. (2008) Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology. Volume 20, Chapman and Hall/CRC, New York. <https://doi.org/10.1201/9781584888413>
- [7] Norton, J.D. (2008) Spatiotemporal Bayesian Hierarchical Models, with Application to Birth Outcomes. The Florida State University, Tallahassee.
- [8] Waller, L.A., Carlin, B.P., Xia, H. and Gelfand, A.E. (1997) Hierarchical Spatio-Temporal Mapping of Disease Rates. *Journal of the American Statistical Association*, **92**, 607-617. <https://doi.org/10.1080/01621459.1997.10474012>
- [9] Miller, H.J. (2004) Tobler's First Law and Spatial Analysis. *Annals of the Associa-*

- tion of American Geographers*, **94**, 284-289.
<https://doi.org/10.1111/j.1467-8306.2004.09402005.x>
- [10] Kyung, M. and Ghosh, S.K. (2009) Bayesian Inference for Directional Conditionally Autoregressive Models. *Bayesian Analysis*, **4**, 675-706.
<https://doi.org/10.1214/09-BA425>
- [11] Smith, R.L. (2001) Environmental Statistics.
- [12] Mariella, L. and Tarantino, M. (2010) Spatial Temporal Conditional Auto-Regressive Model: A New Autoregressive Matrix. *Australian Journal of Statistics*, **39**, 223-244. <https://doi.org/10.17713/ajs.v39i3.246>
- [13] Iddrisu, A.-K. and Amoako, Y.A. (2016) Spatial Modeling and Mapping of Tuberculosis Using Bayesian Hierarchical Approaches. *Open Journal of Statistics*, **6**, 482-513. <https://doi.org/10.4236/ojs.2016.63043>
- [14] Bernardinelli, L., Clayton, D., Pascutto, C., Montomoli, C., Ghislandi, M. and Songini, M. (1995) Bayesian Analysis of Space-Time Variation in Disease Risk. *Statistics in Medicine*, **14**, 2433-2443.
- [15] Knorr-Held, L., Besag, J., *et al.* (1998) Modelling Risk from a Disease in Time and Space. *Statistics in Medicine*, **17**, 2045-2060.
[https://doi.org/10.1002/\(SICI\)1097-0258\(19980930\)17:18<2045::AID-SIM943>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1097-0258(19980930)17:18<2045::AID-SIM943>3.0.CO;2-P)
- [16] Besag, J. and York, M. and Mollié, A. (1991) Bayesian Image Restoration, with Two Applications in Spatial Statistics. *Annals of the Institute of Statistical Mathematics*, **43**, 1-20. <https://doi.org/10.1007/BF00116466>
- [17] Lawson, A.B., Browne, W. and Rodeiro, C.L. (2003) *Vidal. Disease Mapping with WinBUGS and MLwiN*. Wiley and Sons, Hoboken.
- [18] Clayton, D. and Kaldor, J. (1987) Empirical Bayes Estimates of Age-Standardized Relative Risks for Use in Disease Mapping. *Biometrics*, **43**, 671-681.
<https://doi.org/10.2307/2532003>
- [19] Besag, J. (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems (with Discussion). *Journal of the Royal Statistical Society: Series B*, **36**, 192-236.
- [20] Waller, L.A. and Gotway, C.A. (2004) *Applied Spatial Statistics for Public Health Data*. Wiley-Interscience, Hoboken.
- [21] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Van Der Linde, A. (2002) Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 583-639.
- [22] Knorr-Held, L. (1999) Bayesian Modelling of Inseparable Space-Time Variation in Disease Risk.
- [23] Schrödle, B. and Held, L. (2011) Spatio-Temporal Disease Mapping Using INLA. *Environmetrics*, **22**, 725-734. <https://doi.org/10.1002/env.1065>

Appendix: Supplementary Materials

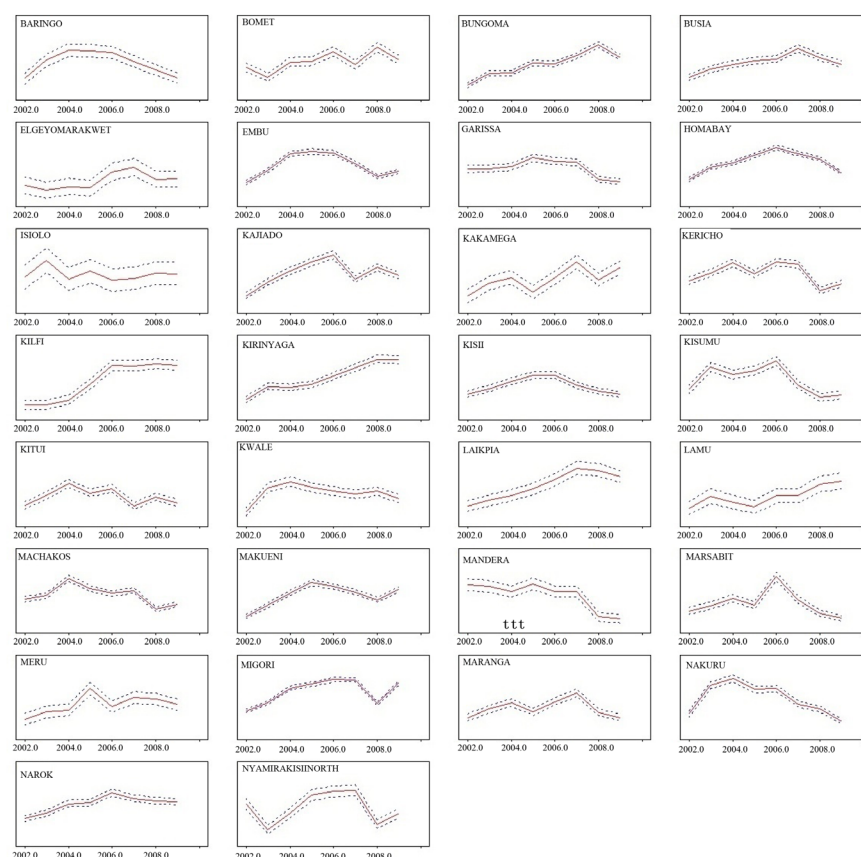


Figure S1. Interaction type IV posterior mean relative risk temporal trend by counties over the study period 2002-2009.

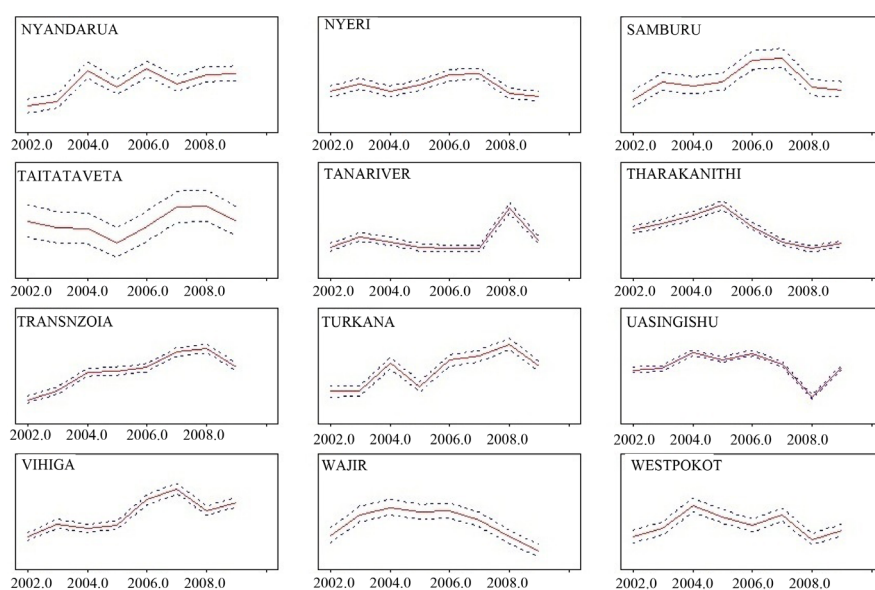


Figure S2. Interaction type IV posterior mean relative risk temporal trend by counties over the study period 2002-2009.

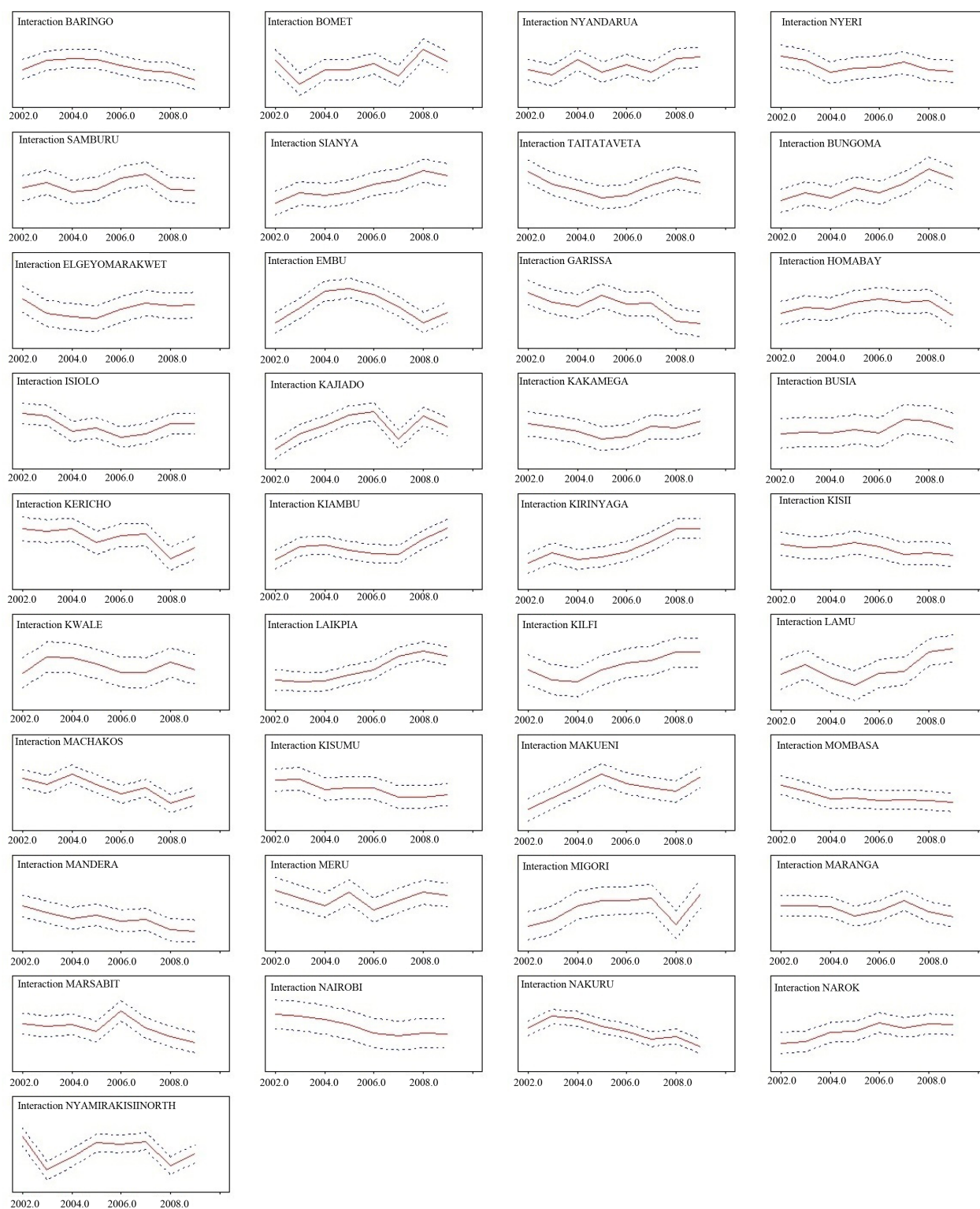


Figure S3. Interaction type IV temporal trend of TB prevalence by counties over the study period 2002-2009.

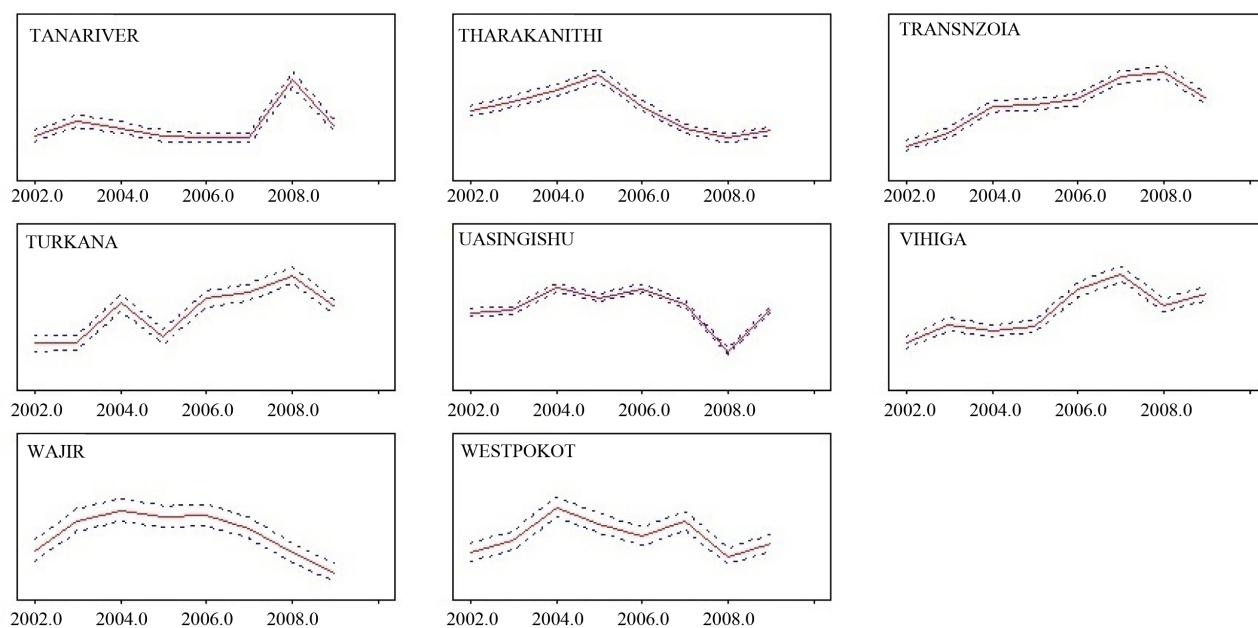


Figure 9. Interaction type IV temporal trend of TB prevalence by counties over the study period 2002-2009.