RESEARCH ARTICLE

Dynamic spatiotemporal modeling of the infected rate of visceral leishmaniasis in human in an endemic area of Amhara regional state, Ethiopia

Anteneh Asmare Godana [1]*, **Samuel Musili Mwalili**[2], **George Otieno Orwa**[2]

1 Pan African University Institute for Basic Science, Technology and Innovation (PAUSTI), Nairobi, Kenya,
2 Jomo Kenyatta University of Agriculture and Technology, Department of Statistics and Actuarial Sciences, Nairobi, Kenya

* antenehstat1988@gmail.com

Abstract

Visceral Leishmaniasis is a very dangerous form of leishmaniasis and, shorn of appropriate diagnosis and handling, it leads to death and physical disability. Depicting the spatiotemporal pattern of disease is important for disease regulator and deterrence strategies. Spatiotemporal modeling has distended broad veneration in recent years. Spatial and spatiotemporal disease modeling is extensively used for the analysis of registry data and usually articulated in a hierarchical Bayesian framework. In this study, we have developed the hierarchical spatiotemporal Bayesian modeling of the infected rate of Visceral leishmaniasis in Human (VLH). We applied the Stochastics Partial Differential Equation (SPDE) approach for a spatiotemporal hierarchical model for Visceral leishmaniasis in human (VLH) that involves a GF and a state process is associated with an autoregressive order one temporal dynamics and the spatially correlated error term, along with the effect of land shield, metrological, demographic, socio-demographic and geographical covariates in an endemic area of Amhara regional state, Ethiopia. The model encompasses a Gaussian Field (GF), affected by an error term, and a state process described by a first-order autoregressive dynamic model and spatially correlated innovations. A hierarchical model including spatially and temporally correlated errors was fit to the infected rate of Visceral leishmaniasis in human (VLH) weekly data from January 2015 to December 2017 using the R package R-INLA, which allows for Bayesian modeling using the stochastic partial differential equation (SPDE) approach. We found that the mean weekly temperature had a significant positive association with infected rate of VLH. Moreover, net migration rate, clean water coverage, average number of households, population density per square kilometer, average number of persons per household unit, education coverage, health facility coverage, mortality rate, and sex ratio had a significant association with the infected rate of visceral leishmaniasis (VLH) in the region. In this study, we investigated the dynamic spatiotemporal modeling of Visceral leishmaniasis in Human (VLH) through a stochastic partial differential equation approach (SPDE) using integrated nested Laplace approximation (INLA). Our study had confirmed both metrological, demographic, sociodemographic and geographic covariates

had a significant association with the infected rate of visceral leishmaniasis (VLH) in the region.

Introduction

Visceral leishmaniasis is the second most parasitic killer in the globe next to malaria, responsible for an estimated 200,000 to 400,000 infections each year globally [1, 2]. Globally about 90-95% of visceral leishmaniasis infections is contributed by most Sub-saharan countries, Asian and Latin American, According to the study by [3, 4]. Based on the species of Leishmania parasites and other immunological and epidemiological aspects, Leishmania infection can lead to cutaneous (CL), mucocutaneous (MCL) or visceral leishmaniasis (VL) [5]. Leishmaniasis is a group of diseases caused by protozoan parasites of the genus Leishmania that are transmitted between humans and other mammalian hosts by phlebotomine sand flies [5], this human infection caused by over 20 species [5, 6]. Furthermore, Fig 1 below shows that the life cycles of parasite of Leishmaniasis diseases. Visceral Leishmaniasis is a very dangerous form of leishmaniasis and, shorn of appropriate diagnosis and handling, it leads to death and physical disability [7]. Depicting the spatiotemporal pattern of disease is important for disease regulator
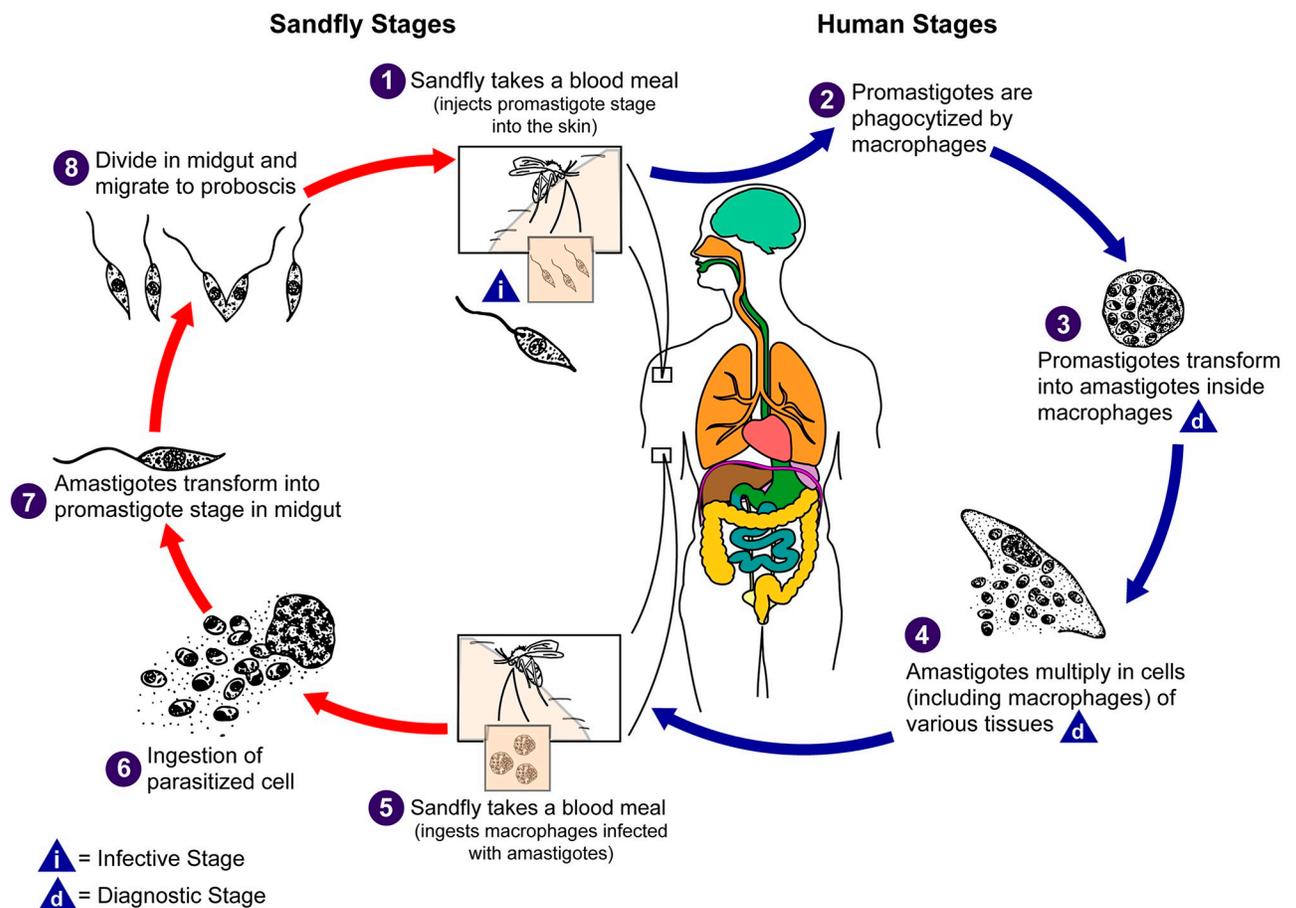
Fig 1. Life cycles of leishmaniasis. (**Source**:https://www.cdc.gov/parasites/leishmaniasis/biology.html, [6].

https://doi.org/10.1371/journal.pone.0212934.g001

and deterrence strategies. Spatiotemporal modeling has distended broad veneration in recent years. Spatial and spatiotemporal disease modeling is extensively used for the analysis of registry data and usually articulated in a hierarchical Bayesian framework. Integrated nested Laplace approximations (INLA) would be used as an apparatus for Bayesian inference. INLA is a favorable substitute to Markov chain Monte Carlo (MCMC) methods which deliver very exact results within little computational period [8], [9]. In the spatiotemporal scheme, data coming from different locations are assumed to be the realization of a contentiously indexed spatial process or random field changing in time; Thus;

$$Y(S:t) : S \in D_s \in R^m, t \in D_t \in R \tag{1}$$

$Y(S:t)$ denote a spatiotemporal random process, where D_s is the spatial domain of interest, D_t is the temporal domain of interest, s is a spatial index and t a time index. Gaussian Fields (GFs) is a prominent application in spatial statistics and spatiotemporal modeling [8], [10], [11], and it is a basic form of constructing hierarchical spatiotemporal Bayesian modeling [8]. Let $S \in D \subset R^m$ is a continuous indexed Gaussian Field (GF) for all finite collection of $[X(s_i)]$ are jointly Gaussian distribution. The Gaussian field specified by its mean function, $\mu = \mu(Y(s;t))$ and the spatiotemporal covariance function $\Sigma = Cov(Y(s;t), Y(s';t')) = \sigma^2 C((s;t),(s';t'))$ with $(s;t), (s';t') \in R^2 \times R$. The spatiotemporal process is second-order stationary if its mean is constant and the spatiotemporal covariance function depends on the locations and time points only through the spatial distance vector $d = s - s' \in R^2$ and the temporal lag $l = t - t' \in R$ [8]. The covariance function is isotropic and stationary, the covariance function is stationary only a function of a relative position of two locations and isotropic a covariance function depends only on the euclidian distance between two locations. Since a covariance matrix is positive definite, the spatiotemporal covariance function must be a positive definite function. Although Gaussian fields are expedient from both an analytical and a practical approach, the computational issues have always been complex. Although Gaussian fields are expedient from both an analytical and a practical approach, the computational issues have always been complex. This is due to the general computational cost of $O(n^3)$ to factorize the dense $n \times n$ spatiotemporal covariance matrices. to overcome 'the big n problem', by doing exact computations on a simplified Gaussian model of low rank [8], [12], [13], applied covariance tapering to zero-out parts of the covariance matrix to improve a computational speed. However, the sparsity pattern will depend on the range of the GFs [14], thus to avoid this complexity the GF is substituted by a Gaussian Markov random field (GMRF). The GMRF illustration can be constructed plainly by using a certain stochastic partial differential equation (SPDE) [8] which has GFs with Matern covariance function as the solution when driven by Gaussian white noise. The result is a basis function representation with piecewise linear basis functions and Gaussian weights with Markov dependences determined by a general triangulation of the domain. The critical lump is that the spatiotemporal covariance function and the dense covariance the matrix of a GF are substituted, respectively, by a neighborhood structure and by a sparse precision matrix that together defines a GMRF. Indeed, GMRFs are defined by a precision matrix with a sparse structure for which it is possible to use computationally effective numerical methods, especially for fast matrix factorization [14]. Moreover, when working with Bayesian inference for GMRFs, it is thinkable to make use of the Integrated Nested Laplace Approximation (INLA) procedure proposed by [15] as an alternative to MCMC methods for latent Gaussian field models. The best marvelous edge of INLA is computational because it produces almost immediately precise approximations to posterior distributions, also in the case of complex models. Thus, the joint use of the SPDE approach together with the INLA method is a candidate for being a powerful solution in overcoming the computational issues related to GF

modeling. The main aim of this paper is the spatiotemporal modeling of the infected rate of visceral leishmaniasis in human (VLH) over a stochastic partial differential equation (SPDE) approach using Integrated nested Laplace approximations (INLA) technique.

$$Y(s;t) = \beta Z(s,t) + \Phi(s,t) + \epsilon(s,t) \tag{2}$$

$$\Phi(s,t) = \lambda \Phi(s,t-1) + \gamma(s,t) \tag{3}$$

The spatiotemporal model above defines a hierarchical model described by a Gaussian field (GF), $Y(s;t)$ assembled from external covariate information $Z(s,t)$, microscale spatiotemporal variation, $\epsilon(s,t)$ and a first-order autoregressive a dynamic model for the latent process $\Phi(s,t)$ with spatially correlated innovations $\gamma(s,t)$. This kind of model has extensively applied in the diseases infection and disease modeling literature appreciations towards its suppleness in modeling the effect of relevant external variables (i.e. meteorological, sociodemographic and ecological variables) as well as time and space dependence. Thus, spatiotemporal modeling of the infected rate of visceral leishmaniasis in human in Amhara regional state, Ethiopia, using weekly data from first week of January 2015 to last week of December 2017 with a total of 156 weeks of 10 different locations in the area a total of 1560 spatiotemporal data for estimation and we used 5 validation location with a total of 780 spatiotemporal data.

Materials and methods

The main aim of this study is to identify the determinates of the infected rate of visceral leishmaniasis in Humans by using dynamic spatiotemporal modelling.

Hierarchical spatiotemporal modeling of the infected rate of visceral leishmaniasis(VLH)

Leishmaniases are an assembly of diseases triggered by more than 20 different species of the protozoan genus Leishmania that is spread between a human being and other mammalian hosts by phlebotomine sandflies [7], [16]. Visceral leishmaniasis (VL), which affects interior body part, for instance, the liver and spleen, if it is not identified and cured in the preliminary phases, it usually leads towards death [17]. The infection endangers about 350 million persons in 98 countries, most of them in the poorer regions of the globe [1], [6]. In Ethiopia, an estimated over 2500 to 4000 new VL cases happen per annum and over 3.2 million persons are at threat of infection of VL [16], [18]. It widely spread over the arid and semi-arid parts of the country. Despite the symptom of visceral leishmaniasis in human (VLH) is fundamentally similar to malaria and enteric fever, reliable laboratory methods become mandatory for the accurate diagnosis of the infection.VL in Ethiopia is caused by L. donovani with an anthroponotic transmission. Leishmania infections do not always equate with clinical illness. The ratio of incident asymptomatic infections to incident clinical cases varies among geographic regions. Most infections in immuno-competent individuals remain asymptomatic [19]. In addition, untreated individuals might act as reservoirs putting the community at risk of ongoing transmissions.

Fig 2 below shows that the flowchart of VLH infection from the first week of January 2015 (week 1) to the last week of December 2017(week 156) at every 15 stations. We collect the records of VLH infection in a spatiotemporal manner by considering all of 15 stations in the endemic regions. Furthermore, we divided these stations into estimation and validation sites, by selecting 10 of stations for estimations and the remaining 5 stations for validation in a random fashion. In this study, VLH infection was tested at each treatment site in the endemic area of Amhara regional state, based on the symptom of the patient and its traveling history
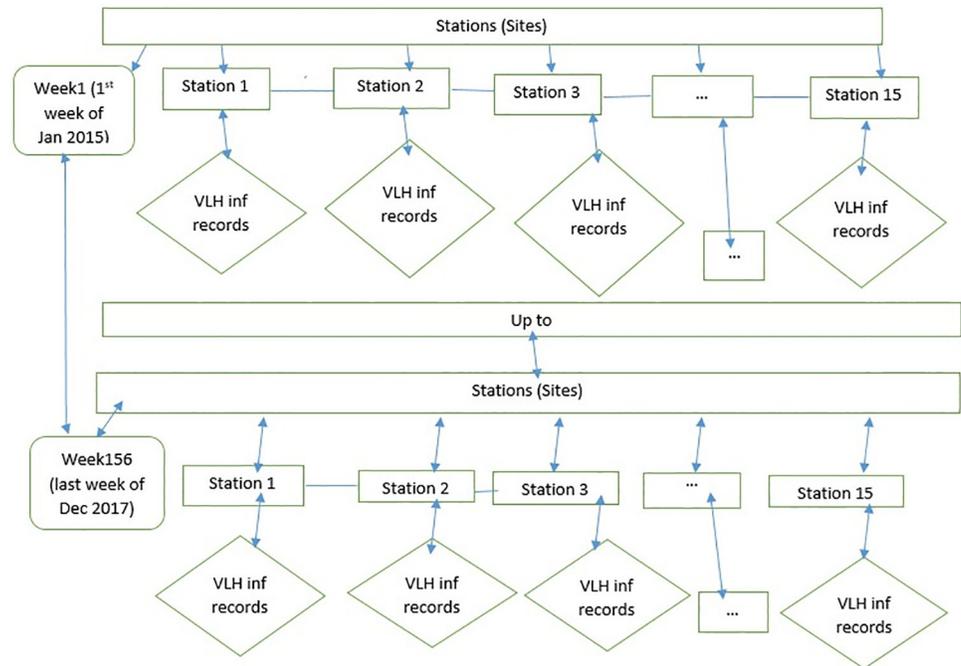
Fig 2. The flowchart shows that weekly visceral leishmaniasis in human (VLH) infection at every 15 sites (stations).

during admission. Visceral Leishmaniasis should always be suspected when an individual presents with prolonged fever from the endemic areas. The test of VLH infection has been depending on clinical, serological, parasitological and molecular tests to detect the infection based on the species of leishmaniasis. We collected data at every 15 VLH stations in the endemic regions, from VLH infection registration book weekly spatiotemporal data from January 2015 to December 2017 and we calculated the weekly infection rate based on the total number of infected population and the total number of populations resident in the region. Thus, ministry heath has to access infected rate of VLH in order to take appropriate and operative actions in order to refine the situation of the most VLH infection regions and to get the map of the infection rate of VLH. As a result, we propose a hierarchical spatiotemporal model able to catch the complex spatiotemporal dynamics of infection the rate of VLH, including the meteorological, sociodemographic and geographical variables as external covariates. We consider Amhara Regional State, Ethiopia. We analyzed a weekly average infected rate of visceral leishmaniasis in human (VLH) particularly in South and North Gondar zone of Amhara region by considering infection areas from the first week of January 2015 to last week of December 2017 with a total of T = 156 weeks. In particular, we consider a total of m = 10 stations for spatiotemporal estimation as shown in Fig 3 below the red hexagon indicates the location of 10 stations and in Fig 4 below the yellow hexagon indicates the location of 5 validation stations. In addition, we have used external covariates such as meteorological, sociodemographic and geographical variables. Moreover, we consider altitude (A, m) and spatial geographic coordinates (X and Y, in km) in addition to weekly temporal dynamics. The data for this study is provided by the University of Gondar leishmaniasis research and treatment center, Ethiopian Central Statistics Agency(CSA) and Amhara Regional Health Bureau.
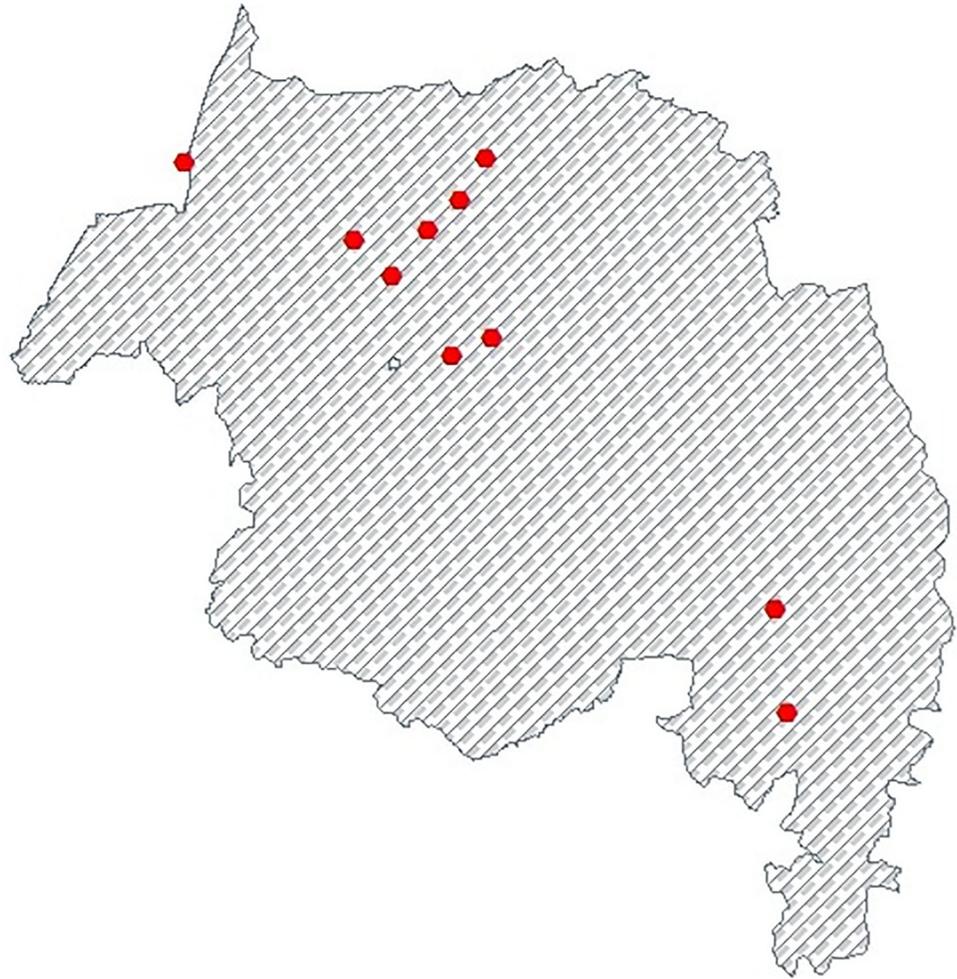
Fig 3. The map shows 10 locations of VLH infection rate stations.
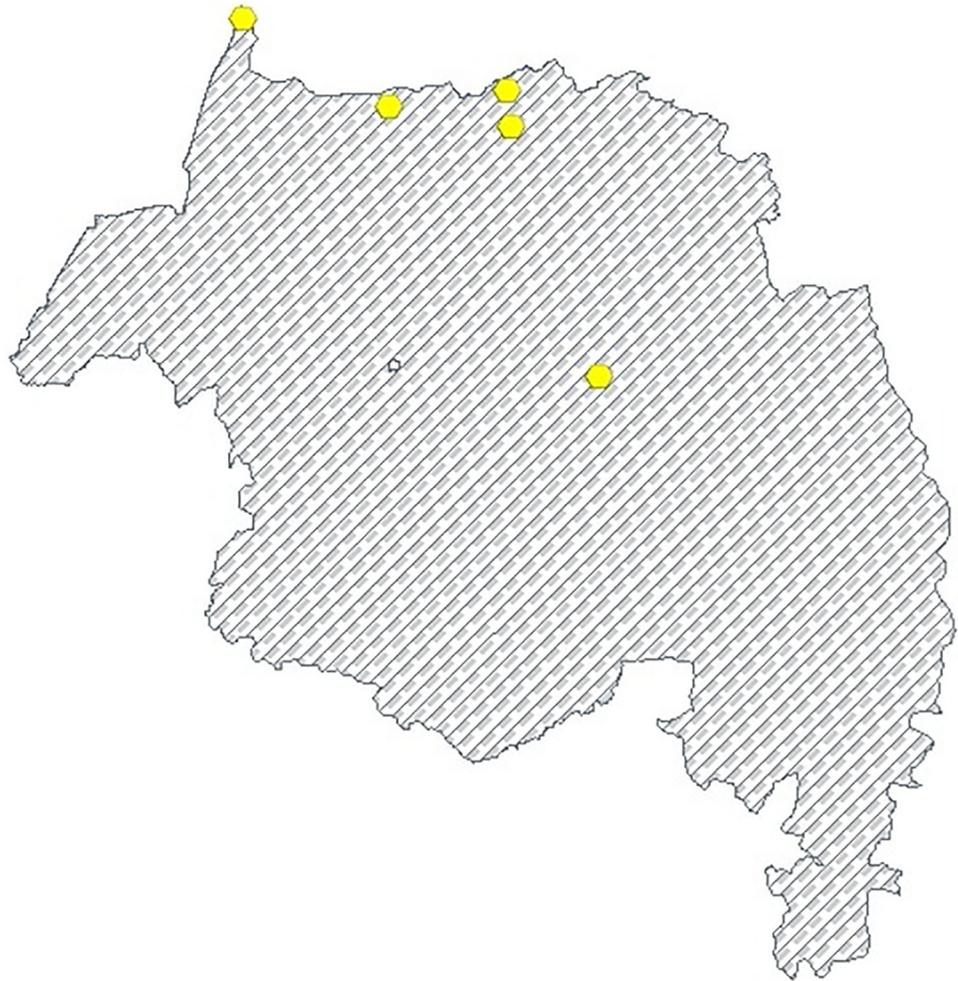
Spatiotemporal models

We assumed that $Y(s; t)$ be the realization of the spatiotemporal process $Y(,;,)$ that indicates the weekly average infected rate of Visceral leishmaniasis in human(VLH) at locations $i = 1, 2, \ldots, m$ located at s_i stations and week $t = 1, 2, \ldots, T$, We assumed the following spatiotemporal model;

$$Y(s_i; t) = \beta Z(s_i, t) + \Phi(s_i, t) + \epsilon(s_i, t) \tag{4}$$

Where; $Z(s_i, t) = Z_1(s_i, t), \ldots, Z_p(s_i, t)$ denote the vector of covariates and $\beta = (\beta_1, \beta_2, \ldots, \beta_p)'$ is a coefficient of the vector, Moreover; $\epsilon(s_i, t) \sim N(0, \sigma_\epsilon^2)$ is spatiotemporal Gaussian white noise process both serially and spatially uncorrelated. $\epsilon(s_i, t)$ is the realization of state process of the unobserved rate of infection of VLH. We assumed the spatiotemporal Gaussian field process that changes in time with first-order autoregressive dynamics with a coefficient of λ and colored innovations is given by;

$$\Phi(s_i, t) = \lambda \Phi(s_i, t - 1) + \gamma(s_i, t) \tag{5}$$

For $t = 2, 3, \ldots, T$ where $|\lambda| < 1$ and $\Phi(s_i, 1)$ is derived from a stationary Gaussian distribution

Fig 4. The map shows 5 locations of VLH infection rate validation stations.

$\sim N\left(0, \frac{\sigma_\gamma^2}{1-\lambda^2}\right)$ Moreover, $\gamma(s_i, t)$ has a zero-mean Gaussian distribution, is assumed to be temporally independent and characterized by spatiotemporal covariance functions;

$$Cov\left(\gamma(s_i, t), \gamma(s_i', t')\right) = \begin{cases} 0, if & t \neq t' \\ \sigma_\gamma^2(r) & if\, t = t' \end{cases}$$

for $i \neq j$ the purely spatial correlation function $C(r)$ is depending on the location s_i and s_j only through the euclidian spatial distance $r = ||s_i - s_j|| \in$ R. Thus, the process is second order stationary and isotropic [10]. It follows that $Var(\gamma(s_i, t)) = \sigma_\gamma^2$ for each s_i and t. The Spatial correlation function is defined as a Matern function;

$$C(r) = \frac{1}{\Gamma(v)2^{v-1}}(kr)^v \kappa_v(kr)$$

With κ_v denoting the modified Bessel function of the second kind and order $v > 0$, The parameter v, which usually kept stable measures of the degree of smoothness of the process and its integer value decides the mean square differentiability of the process, $k > 0$ is the scaling parameter

related to the range ρ we use empirically derived definition $\rho = \frac{\sqrt{8v}}{k}$ with ρ corresponding the distance where the spatial correlation is close to 0.1 [8], [20]. From Eqs (4) and (5) by collecting observation in vector t we have:

$$Y_t = (y(s_1, t), y(s_2, t), \ldots, y(s_m, t))'$$
$$Z_t = (z(s_1, t), z(s_2, t), \ldots, z(s_m, t))'$$
$$\Phi_t = (\Phi(s_1, t), \Phi(s_2, t), \ldots, \Phi(s_m, t))'$$
$$\epsilon_t = (\epsilon(s_1, t), \epsilon(s_2, t), \ldots, \epsilon(s_m, t))'$$

It follows that;

$$Y_t = \beta Z_t + \Phi_t + \epsilon_t \tag{6}$$

$\epsilon_t \sim N(0, \sigma_\epsilon^2 I_m)$ and also;

$$\Phi_t = \lambda \Phi_{t-1} + \gamma_t \tag{7}$$

$\gamma_t \sim N(0, \Sigma = \sigma_\epsilon^2 \tilde{\Sigma})$, we assume is coming from the stationary distribution of AR(1) process with, $\Phi_1 \sim N\left(0, \frac{\Sigma}{1-\lambda^2}\right) =\sim N\left(0, \frac{\sigma_\epsilon^2 \tilde{\Sigma}}{1-\lambda^2}\right)$ Moreover, the dense correlation matrix of dimension m with an element $C(\|s_i - s_j\|)$, Where $C(r)$ is a Matern covariance function;

$$C(r) = \frac{1}{\Gamma(v)2^{v-1}}(kr)^v \kappa_v(kr)$$

and parameterized by κ and v. Let $\Theta = (\beta, \sigma_\epsilon^2, \lambda, \sigma_\gamma^2, \kappa)$ denote the parameter of the vector to be estimated, thus, the joint posterior distribution is given by;

$$P(\Theta, \Phi|Y) \propto P(Y|\Theta, \Phi)P(\Phi|\Theta)P(\Theta) \tag{8}$$

Where $P(.)$ denotes the probability density functions, $Y = (Y_t)$ and $\Phi = (\Phi_t)$ with $t = 2, 3, \ldots, T$ and assume the hyper parameters are independent $P(\Theta) = \prod_{i=1}^{dim(\Theta)} P(\Theta_i)$, The left-hand side of Eq (8) above is the joint posterior distribution and the right-hand side of Eq (8), $P(Y|\Theta, \Phi)$ is the data model (likelihood of the data), $P(\Phi|\Theta)$ is the latent (unobserved) process model and $P(\Theta)$ is the parameter model. From Eq (6) above we have

$$Y_t = \beta Z_t + \Phi_t + \epsilon_t$$

$\epsilon_t \sim N(0, \sigma_\epsilon^2 I_m)$ Thus; the data model (likelihood of the data) can be written as

$$P(Y_t|\Phi, \Theta) \sim N(\beta Z_t + \Phi_t, \sigma_\epsilon^2 I_m)$$
$$= \left(\frac{1}{\sqrt{2\pi}\sigma_\epsilon}\right)^{mT} exp^{\frac{-1}{2}} \sum_{i=1}^{T} \left(\frac{\epsilon_t \epsilon_t'}{\sigma_\epsilon^2}\right)$$

but $\epsilon_t = Y_t - \beta Z_t - \Phi_t$

$$P(Y_t|\Phi, \Theta) \propto (\sigma_\epsilon^2)^{\frac{-mT}{2}} exp^{\frac{-1}{2\sigma_\epsilon^2}} \left[\sum_{i=1}^{T} \left((Y_t - \beta Z_t - \Phi_t)(Y_t - \beta Z_t - \Phi_t)'\right)\right] \tag{9}$$

The joint distribution of spatiotemporal process model is can be factorized as with the total law of probability and Markovian process;

$$P(\Phi_1, \Phi_2, \ldots, \Phi_T|\Theta) = P(\Phi_T|\Phi_{T-1}, \Phi_{T-2}, \ldots, \Phi_1, \Theta) \times \ldots, P(\Phi_2|\Phi_1, \Theta) \times P(\Phi_1|\Theta)$$

using a Markovian process assumption we have;

$$P(\Phi_T|\Phi_{T-1}, \Phi_{T-2}, \ldots, \Phi_1, \Theta) = P(\Phi_T|\Phi_{T-1}, \Theta)$$

We have the joint distribution of the process model;

$$P(\Phi_1, \Phi_2, \ldots, \Phi_T|\Theta) = P(\Phi_1|\Theta)\prod_{t=2}^{T} P(\Phi_t|\Phi_{t-1}, \Theta) \tag{10}$$

We assume the hyperparameter are independent, the parameter model is given by;

$$P(\Theta) = \prod_{i=1}^{dim(\Theta)} P(\Theta_i) \tag{11}$$

The posterior distribution is given by the product of the data model, process model, and parameter model [21]. we assume Φ_1 is AR(1) stationary process, $\Phi_1|\Theta \sim N\left(0, \frac{\Sigma}{1-\lambda^2}\right)$, $\Phi_1|\Theta$ also called the initial distribution [21].

$$P(\Phi_1|\Theta) = \left(\frac{1}{\sqrt{2\pi}\left|\frac{\Sigma}{1-\lambda^2}\right|^{\frac{1}{2}}}\right)^m exp^{\frac{-1}{2}}\left(\Phi_1'\left(\frac{\Sigma}{1-\lambda^2}\right)^{-1}\Phi_1\right)$$

From above we define and assume $\Sigma = \sigma_\gamma^2\tilde{\Sigma}$

$$P(\Phi_1|\Theta) \propto \left(\frac{\sigma_\gamma^2}{1-\lambda^2}\right)^{\frac{-m}{2}}|\tilde{\Sigma}|^{\frac{-1}{2}}exp^{\left(-\frac{1-\lambda^2}{2\sigma_\gamma^2}\Phi_1'\tilde{\Sigma}^{-1}\Phi_1\right)} \tag{12}$$

Also the evolution distribution; $P(\Phi_t|\Phi_{t-1}, \Theta)$ since follows an AR(1) process, $\Phi_t = \lambda\Phi_{t-1} + \gamma_t$ and $\gamma_t \sim N(0, \Sigma = \sigma_\epsilon^2\tilde{\Sigma})$, The joint eveloution distribution is given as;

$$P(\Phi_t|\Phi_{t-1}, \Theta) = \left(\frac{1}{\sqrt{2\pi\sigma_\gamma^2}}\right)^{m(T-1)}|\tilde{\Sigma}|^{-\left(\frac{T-1}{2}\right)}exp^{\left(-\frac{1}{2\sigma_\gamma^2}\sum_{i=1}^{T}(\Phi_t-\lambda\Phi_{t-1})\tilde{\Sigma}^{-1}(\Phi_t-\lambda\Phi_{t-1})'\right)}$$

$$P(\Phi_t|\Phi_{t-1}, \Theta) \propto (\sigma_\gamma^2)^{\frac{-m(T-1)}{2}}|\tilde{\Sigma}|^{-\left(\frac{T-1}{2}\right)}exp^{\left(-\frac{1}{2\sigma_\gamma^2}\sum_{i=1}^{T}(\Phi_t-\lambda\Phi_{t-1})\tilde{\Sigma}^{-1}(\Phi_t-\lambda\Phi_{t-1})'\right)} \tag{13}$$

the joint process model is expressed as the product of the joint initial distribution and the joint evolution distribution [21], Thus from Eqs (12) and (13), we have the process model the joint distribution given by;

$$P(\Phi_1, \Phi_2, \ldots, \Phi_T|\Theta) = P(\Phi_1|\Theta)\prod_{t=2}^{T} P(\Phi_t|\Phi_{t-1}, \Theta)$$

$$\propto \left(\frac{\sigma_\gamma^2}{1-\lambda^2}\right)^{\frac{-m}{2}}|\tilde{\Sigma}|^{\frac{-1}{2}}exp^{\left(-\frac{1-\lambda^2}{2\sigma_\gamma^2}\Phi_1'\tilde{\Sigma}^{-1}\Phi_1\right)} \times$$

$$(\sigma_\gamma^2)^{\frac{-m(T-1)}{2}}|\tilde{\Sigma}|^{-\left(\frac{T-1}{2}\right)} \times \tag{14}$$

$$exp^{\left(-\frac{1}{2\sigma_\gamma^2}\sum_{i=1}^{T}(\Phi_t-\lambda\Phi_{t-1})\tilde{\Sigma}^{-1}(\Phi_t-\lambda\Phi_{t-1})'\right)}$$

The parameter model also from Eq (11) above;

$$P(\Theta) = \prod_{i=1}^{dim(\Theta)} P(\Theta_i)$$

Thus, the joint posterior model is the product of the data model, process model and parameter model [21] give us;

$$A = (\sigma_\epsilon^2)^{\frac{-mT}{2}} exp^{\frac{-1}{2\sigma_\epsilon^2}} \left[\sum_{i=1}^{T} \left((Y_t - \beta Z_t - \Phi_t)(Y_t - \beta Z_t - \Phi_t)' \right) \right] \tag{15}$$

$$B = \left(\frac{\sigma_\gamma^2}{1-\lambda^2} \right)^{\frac{-m}{2}} |\tilde{\Sigma}|^{\frac{-1}{2}} exp^{\left(-\frac{1-\lambda^2}{2\sigma_\gamma^2} \Phi_1' \tilde{\Sigma}^{-1} \Phi_1 \right)} \times$$

$$(\sigma_\gamma^2)^{\frac{-m(T-1)}{2}} |\tilde{\Sigma}|^{-\left(\frac{T-1}{2} \right)} \times \tag{16}$$

$$exp^{\left(-\frac{1}{2\sigma_\gamma^2} \sum_{i=1}^{T} (\Phi_t - \lambda \Phi_{t-1}) \tilde{\Sigma}^{-1} (\Phi_t - \lambda \Phi_{t-1})' \right)}$$

$$C = \prod_{i=1}^{dim(\Theta)} P(\Theta_i) \tag{17}$$

The joint posterior distribution;

$$P(\Theta, \Phi | Y) \propto P(Y|\Theta, \Phi) P(\Phi|\Theta) P(\Theta)$$

From Eqs (15) to (17) the joint posterior distribution is approximated by;

$$P(\Theta, \Phi | Y) \propto A \times B \times C \tag{18}$$

where $|\tilde{\Sigma}|$ is the determinant of the dense m-dimensional correlation matrix $\tilde{\Sigma}$.

GMRF and Stochastics Partial Differential Equation (SPDE) approach

Gaussian Markov random field (GMRF). GRMF is a zero mean multivariate Gaussian distribution with a sparse precision matrix [8], [20].

$$P(X) = \left(\frac{1}{\sqrt{2\pi}} \right)^n |Q|^{\frac{1}{2}} exp^{\frac{-1}{2} \left(X' QX \right)}$$

Where Q is the precision matrix it just the inverse of the covariance matrix.

Theorem: $Q_{ij} = 0 \Leftrightarrow X_i$ is conditionally independent X_j given all other variables, *i.e* $(X_i \perp X_j | X_{-ij})$. A collection of a random variable X is a Gaussian Markov random field (GMRF) with respect to a graph with vertex V and edge G, *i.e* $G = (V, E)$ with mean μ and a precision matrix Q with its probability distribution of a multivariate Gaussian, if the elements of the precision matrix is non zero ($Q_{ij} \neq 0$) implies that $(i, j) \in E \, \forall i \neq j$.

$$P(X) = \left(\frac{1}{\sqrt{2\pi}} \right)^n |Q|^{\frac{1}{2}} exp^{\frac{-1}{2} \left((X-\mu)' Q(X-\mu) \right)} \tag{19}$$

A GMRF is a spatial process that used to model the spatial dependence of data observed on areal units, for instance, regular grid, lattice structure or geographical regions [14]. A Gaussian

Markov random field X can be specified over the conditional distributions for every part given all the others. Furthermore, the Markovian property is connected to the definition of a neighborhood structure, in that the full conditional distribution of X_i for $i = 1, 2, 3 \ldots, n$ depends only on a few of the components of X. This set of components is denoted by ζ_i, which constitutes the set of neighbors of unit i, and

$$P(X_i|X_{-i}) = P(X_i|X_{\zeta_i}) \tag{20}$$

where the notation X_{-i} denotes all elements in X except for X_i. This implies that given the neighborhood ζ_i, the terms X_i and $X_{-(i,\zeta_i)}$ are independent. The conditional independence relation can be written as;

$$X_i \perp X_{-(i,\zeta_i)}|X_{\zeta_i}$$

for $i = 1, \ldots, n$. The main aims of this conditional independence property are rigorously related to the precision matrix Q, Generally for i and j with $i \neq j$ it holds that;

$$X_i \perp X_j|X_{-}(i,j) \Leftrightarrow Q_{ij} = 0$$

which indicates that the nonzero pattern of Q is given by the neighborhood structure of the process. Thus, $Q_{ij} \neq 0$ if $j \in (i, \zeta_i)$.

The computational advantage of making inference with a GMRF stalks directly from the sparsity of the precision matrix Q. The computational properties of GMRFs are enhanced by utilizing Integrated Nested Laplace Approximations (INLA) for Bayesian inference [9], [15]. is a computationally capable strategy that produces quick and exact approximations to posterior distributions.

Stochastics Partial Differential Equation (SPDE) approach

Let $X(s) \equiv (x(s), s \in D_s \subseteq R^2)$ denote a Matern field, this implies that the second order stationary and isotropic Gaussian field with Materns covariance functions

$$C(r) = \frac{1}{\Gamma(v)2^{v-1}}(kr)^v \kappa_v(kr)$$

and depending on the scale κ and smoothness parameter v. Suppose we observe a realization of the spatial process $X(s_i)$ at m spatial location $s_1, s_2, \ldots, s_m$. The primary points of the SPDE approach are to discover a GMRF, with a nearby neighborhood and a sparse precision matrix Q, that profoundly speaks to the Matern field. Accordingly, it is conceivable to make inference utilizing the GMRF with the best computational expenses. This makes it conceivable to maintain a computational problem-related to the large dimension of the matrix [8] that emerges when working with the dense covariance matrix of a GF.

On a very basic level, the SPDE approach utilizes a finite element representation to define the Matrn field as a linear combination of basis functions defined on a triangulation of the domain D_s. This comprises of subdividing D_s into a set of non-crossing triangles meeting in at most a typical edge or corner. To begin with, the triangle initial vertices are put at the areas s_1, $s_2, \ldots, s_m$ and after that extra vertices are included request to get an appropriate triangulation valuable for spatial forecast purposes. Matern covariance functions show up in different fields [8] in any case, essential acknowledgment we will make utilization of its that GF $X(s)$ with Matern covariance is a solution to the linear fractional SPDE

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}}X(s) = \omega(s) \tag{21}$$

$s \in \mathrm{R}^m$, $\alpha = v + \frac{m}{2}$, $\kappa > 0$, $v > 0$, where $(\kappa^2 - \Delta)^{\frac{\alpha}{2}}$ is a pseudodifferential operator. The innovation process ω is spatial Gaussian white noise with unit variance and Δ is the Laplacian

$$\Delta = \sum_{i=1}^{m} \frac{\partial^2}{\partial x_i^2}$$

and the marginal variance

$$\sigma^2 = \frac{\Gamma(v)}{\Gamma\left(v + \frac{m}{2}\right) + 4\pi^{\frac{m}{2}}\kappa^{2v}}$$

We shall name any solution to Eq (21) is a Matern field [8]. The Matern field is the main stationary solution for SPDE. Ordinarily, the triangulation has amplified the base inside the triangle edge, purported Delaunay triangulations, which guarantees that the advances among little and substantial triangles are smooth. The additional vertices are included heuristically [8], [20] to endeavor to limit the aggregate number of triangles that are expected to satisfy the size and shape constraints. To build a GMRF portrayal of Matern's field on the triangulated grid, we begin the stochastics week solution of SPDE in Eq (21). We define the inner product;

$$(f, g) = \int f(s)g(s)ds$$

where the integral is over the region of interest. The stochastic weak solution of the SPDE is found by requiring that

$$\left(\langle \eta_j, (\kappa^2 - \Delta)^{\frac{\alpha}{2}}x \rangle, j = 1, 2, 3 \ldots, q \right) \underline{Dist} \left(\langle \eta_j, \omega \rangle, j = 1, 2, 3 \ldots, q \right)$$

For every appropriate finite set of the test function (η_j, $j = 1, 2, 3\ldots, q$). Given the triangulation, the basis function representation of the Matérn field, The finite element representation the solution to SPDE [8], [22].

$$X(S) = \sum_{h=1}^{k} \psi_h(S)\omega_h \tag{22}$$

for some chosen basis functions ψ_h and Gaussian-distributed weights ω_h. Here, n is the number of vertices in the triangulation. We choose to use functions ψ_h that are piecewise linear in each triangle, defined such that ψ_h is 1 at vertex h and 0 at all other vertices. The key purpose of the SPDE approach is the finite element representation in Eq (22) that sets up the connection between the GF $X(s)$ and the GMRF characterized by the Gaussian weights (ω_h) to which a Markovian structure can be given.

Specifically, the precision matrix Q of the GMRF $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ is a function of κ^2 [8], for $\alpha = 1, 2, 3, \ldots$ and $v = 0, 1, 2, \ldots$ with $\alpha = v + 1$. This characterizes an explicit mapping from the parameters of the GF covariance function (κ) and (v) to the elements of the precision matrix Q of the GMRF ω, with a computational expense of $O(n)$ for any triangulation [20]. We describe how to actualize the spatiotemporal model characterized in Eqs (4) and (5) by utilizing the SPDE approach. We focus on reclassifying the model make use of the connection between GF and GMRF and estimation of the posterior parameters.

For each time point $t = 1, 2, 3\ldots, T$, the Matern field γ_t in Eq (7) is characterized through GMRF, $\tilde{\gamma}_t \sim N(0, Q_s^{-1})$ where, Q_s is the precision matrix originates from the SPDE approach. The matrix Q_s does not change in time because of the serial independence assumption of Eq (5) and its dimension n is given by the number of vertices of the domain triangulation. In this

manner Eq (7) can modify as;

$$\Phi_t = \lambda \Phi_{t-1} + \tilde{\gamma}_t \tag{23}$$

$$\tilde{\gamma}_t \sim N(0, Q_s^{-1})$$

for $t = 1, 2, 3\ldots, T$ with $\Phi_1 \sim N\left(0, \frac{Q_s^{-1}}{1-\lambda^2}\right)$ follows that the joint distribution of Tn dimensional GMRF, $\Phi = (\Phi_1', \Phi_2', \ldots, \Phi_T')$ is given by;

$$\Phi \sim N(0, Q^{-1}) \tag{24}$$

Where, $Q = Q_s \otimes Q_T$ and Q_T is T- dimensional precession matrix of the autoregressive process of order 1 in Eq (23). Equation Eq (6) we can rewrite as also;

$$Y_t = \beta Z_t + H\Phi_t + \epsilon_t, \epsilon_t \sim N(0, \sigma_\epsilon^2 I_m) \tag{25}$$

where the $(m \times n)$ dimensional matrix H selects the value of the GMRF Φ_t for each observation vector Y_t. In particular, H is a sparse matrix with only one unit element for each row and such that;

$$Y(s_i, t) = \beta Z(s_i, t) + \sum_{i=1}^{n} H_{ij}\Phi_t + \epsilon(s_i, t)$$

Where $H_{ij} = 1$ if the triangulation vector j is placed at location s_i and 0 otherwise.

Parameter estimation

The hierarchical model characterized by Eqs (23) and (25) belongs to the class of latent Gaussian models and can be evaluated using the INLA procedure proposed by [15]. INLA is a computational methodology for Bayesian inference and is an option to MCMC for getting the approximated posterior marginals for the latent variable and additionally for the hyper parameters.

Let $X = (\Phi, \beta)$ mean the basic latent field with a priori independent parts. We dole out a vague Gaussian with known precession to β and the Gaussian Markov Random Field (GMRF) distribution to Φ in Eq (24). Consequently the density $P(X|\Theta)$ is a Gaussian distribution with mean zero and precession matrix Q with hyperparameter vector $(\sigma_\gamma^2, \lambda, \kappa)$. Also, we have the observation $Y = (Y_t)$ is normally distributed and conditionally independent given X and σ_ϵ^2. Thus, $\Theta = (\sigma_\gamma^2, \sigma_\epsilon^2, \lambda, \kappa)$ is the hyper parameter vector, The joint posterior distribution is given by;

$$P(X, \Theta|Y) = P(\Theta)P(X|\Theta)\prod_{t=1}^{T} P(Y_t|\Theta, X) \tag{26}$$

Where $P(Y_t|\Theta, X) \sim N(\beta Z_t + H\Phi_t, \sigma_\epsilon^2 I_m)$ is the conditional distribution of the infected rate of VLH observation at time t defined by Eq (25). Thus, the posterior marginal distribution of the

latent field and the hyper paramater is given by;

$$P(X_i|Y) = \int P(X_i|\Theta, Y)P(\Theta|Y)d\Theta \tag{27}$$

$$P(\Theta_j|Y) = \int P(\Theta|Y)d\Theta_{-j} \tag{28}$$

for $i = 1, 2, 3, \ldots, T + p$ and $j = 1, 2, \ldots, 4$.

It merits nothing that for the specific model we are managing, described by Gaussian observations, we have that $\tilde{P}(X_i|Y)$ is correct and Gaussian and the main approximation is the numerical integration required for computing $\tilde{P}(\Theta|Y)$.

Results and discussion

Results

We considered the average weekly infected rate of VLH from the first week of January 2015 to last week of December 2017 with $T = 156$ weeks, at every m = 10 locations and 5 validation sites in Amhara Regional State, Ethiopia, including geographical, socioeconomic, metrological covariates, spatial coordinates and temporal dynamics. We had also average infected rate data and average infected rate validation data with 1560 rows and 780 rows respectively, one row for each week, 15 columns (Station ID, time, the infected rate of VLH, Average number of Household (ANH), Average number of Persons Per housing Unit (ANPPHU), Population Density per square kilometer (PDPSK), Health Facility Coverage in Percentage (HFCIP), Education coverage in percentage (ECIP), Net Migration rate (NMR), Average weekly temperature (AT), Mortality rate (MR), Sex ratio (SR), Clean water coverage in percentage (CWC) and Spatial coordinates (X, Y, Elevation)).

Since the covariates are very extraordinary we applied standardization technique for covariates, Additionally, so as to stabilize the variances, which increment with the mean values, and to make the distribution of the infected rate of VLH data approximately normal, we utilized a logarithmic transformation. The data for this study is provided by the University of Gondar Leishmaniasis research center, Ethiopian Central Statistical Agency (CSA) and Amhara Regional Health Bureau. Fig 5 below demonstrates that the mesh construction and triangulation of Amhara regional state, the red specks showing the area of 10 stations and blue specks shows the validation of 5 stations. We sought a triangulation based on initial vertices at the $m = 10$ station locations and further vertices are included to fulfill the triangulation constraints. We created an SPDE model object for a Matérn-like spatial covariance function with a parameter $\alpha = 2$ implies that the smoothness parameter of Matérn covariance function $\nu = 1$.

Appreciations to R-INLA, so as to evade having to keep track of vertex indexing, we utilized an R-INLA feature that enables the observation equation to be written in matrix form, $Y = A\tau + \epsilon$ where, Y are the observations, τ is a linear predictor, ϵ is the observation noise and A is an observation (projection) matrix. The function inla.stack(.) is used to build the necessary data structures, combining simple model building blocks into large complicated models [9] Using a helper function, we built an observation matrix that extracts the values of the spatiotemporal field at the measurement locations and time points used for the parameter estimation. Table 1 above shows that The posterior summary statistics(Mean, Standard Deviation and 95% credible interval) of the fixed effect (covariate coefficient) β, Specifically, the posterior mean of the intercept is 0.00247 on the log scale, which corresponds to an average infected rate of visceral leishmaniasis (VLH) is 0.0025, after alteration for covariates.
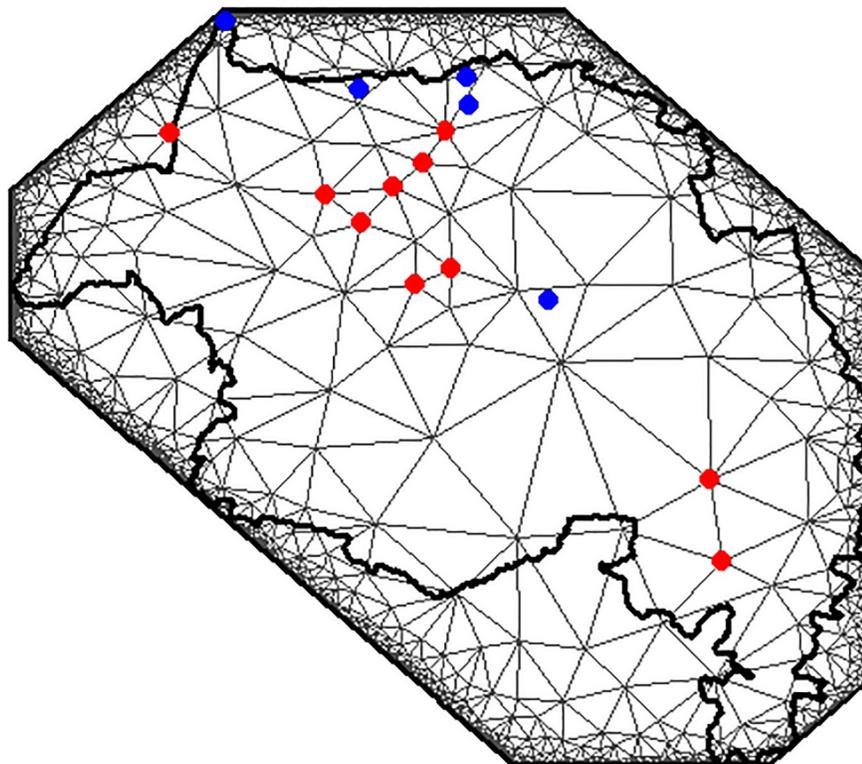
Constrained refined Delaunay triangulation

Fig 5. Mesh construction and triangulation of Amhara regional state, for 10 locations and 5 validation sites.

Table 1. Posterior estimates (mean, standard deviation and 95% credible interval) of the covariate coefficient vector β.

Covariates	Mean	St.Dev	Quantiles (0.025)	Quantiles (0.5)	Quantiles(0.975)
Intercept	0.00274	3.8869	0.007445	0.002781	0.00293
ANH	0.0019	0.234	-0.00357	0.0019	0.0037
ANPPHU	0.0014	0.176	-0.00127	0.0014	0.0027
PDPSK	-0.0005	0.0700	-0.1379	-0.0005	0.1368
HFCIP	0.0023	0.0051	-0.0099	0.0023	0.0100
ECIP	0.0127	0.0129	-0.0253	0.0127	0.0253
MR	0.0003	0.0401	-0.0785	0.0003	0.0790
AT	0.001	0.0749	0.00015	0.001	0.1469
NMR	0.0012	0.0005	-0.0019	0.0012	0.00329
CWC	0.0005	0.0358	-0.0697	0.0005	0.0707
SR	0.0001	0.0806	-0.1584	-0.0001	0.1580
X	-0.00322	0.034744	-0.0068345	-0.00322	0.0068
Y	-0.00191	0.0291	-0.00572	-0.00191	0.00571
Elevation	-0.00331	0.028319	-0.055689	-0.0041095	0.00556

Expectedly, a significant and positive relationship is perceived between the Average number of Household (ANH) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of ANH is 0.0019 and it lies in 95% credible interval. This indicates that an Average number of household (ANH) is one of the determinates of the infected rate of VLH as upheld by the posterior summaries in Table 1 above. Also, a significant and positive relationship is observed between the Average number of Persons Per housing Unit (ANPPHU) and the Infected Rate of Visceral Leishmaniasis (VLH) as the posterior mean of ANPPHU is positive (0.0014) and it lies at 95% credible interval as depicted in posterior summaries above. Additionally, a significant and positive relationship is observed between the Health Facility Coverage in Percentage (HFCIP) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of HFCIP is positive (0.0023) and it lies at 95% credible interval, indicates that HFCIP is one of the determinants of the Infected Rate of Visceral Leishmaniasis (VLH). Likewise, there is a significant and positive relationship perceived between Education coverage in percentage (ECIP) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of ECIP is positive (0.0127) and it lies at 95% credible interval. As a result, ECIP is also the determinants of the Infected Rate of Visceral Leishmaniasis (VLH) in the region. Similarly, there is a significant and positive relationship observed between Mortality rate (MR) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of MR is positive (0.0003) and it lies at 95% credible interval. Correspondingly, there is also a significant and positive relationship observed between Clean water coverage in percentage (CWC) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of CWC is positive (0.0005) and it lies at 95% credible interval. As a result, CWC is also another determinant of the Infected Rate of Visceral Leishmaniasis (VLH) in the region. Similarly, there is also a significant and positive relationship observed between Net Migration Rate (NMR) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of NMR is positive (0.0012) and it lies at 95% credible interval. Consequently, NMR is also another determinant of the Infected Rate of Visceral Leishmaniasis (VLH) in the region. Additionally, there is also a significant and positive relationship observed between Sex ratio (SR) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of SR is positive (0.0001) and it lies at 95% credible interval. Moreover, there is also a significant and positive relationship observed between Average Temperature (AT) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of AT is positive (0.001) and it lies at 95% credible interval. Consequently, AT is also another determinant of the Infected Rate of Visceral Leishmaniasis (VLH) in the region and confirms that the importance of meteorological variables on the Infected rate of Visceral Leishmaniasis (VLH). Besides, there is also a significant and negative relationship observed between Population Density per square kilometer (PDPSK) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of PDPSK is negative (-0.0005) and it lies at 95% credible interval. Subsequently, PDPSK is also another determinant of the Infected Rate of Visceral Leishmaniasis (VLH) in the region. Finally, there is also a significant and negative relationship observed between, Spatial coordinates (X, Y, Elevation) and the Infected Rate of Visceral Leishmaniasis (VLH) since the posterior mean of (X, Y, Elevation) is negative (-0.00322,-0.00191,-0.00331) respectively and it lies at 95% credible interval. Therefore, Spatial coordinates is also another determinant of the Infected Rate of Visceral Leishmaniasis (VLH) in the region and confirm that the importance of spatial effects on the Infected rate of Visceral Leishmaniasis (VLH).

Table 2 above shows, the summary statistics of the posterior distribution of the AR(1) coefficient λ, the R-INLA function provides us with the mean, quantiles and standard deviation of the Gaussian observation precision parameters $\frac{1}{\sigma_\epsilon^2}$. As we are apprehensive in the variance σ_ϵ^2,

Table 2. Posterior estimates (mean, standard deviation and 95% credible interval) of the Parametrs $\Theta = (\sigma_\gamma^2, \sigma_\epsilon^2, \lambda, \rho)$.

Parametrs (Θ)	Mean	St.Dev	Quantiles (0.025)	Quantiles (0.5)	Quantiles(0.975)
σ_ϵ^2	0.0176	0.0081	0.0161	0.0176	0.0193
σ_γ^2	0.276	0.0035	0.211	0.287	0.321
ρ	526	26.1	518	531.6	582
λ	0.7263	0.0273	0.6983	0.7421	0.7916

We transformed the marginal density of the precision $\frac{1}{\sigma_\epsilon^2}$ to variance σ_ϵ^2. The parameter estimates for the spatial SPDE model we obtained using R-INLA which extracts all relevant information from the model summary, also transforms the results from internal parameter scales, giving posterior distributions for nominal variance and nominal range in addition to the internal $\theta_1 = log(\tau)$ and $\theta_2 = log(\kappa)$ where κ is a scaling parameter and τ is parameter that rescales the field. We get a value of 526 km for the empirically derived correlation range $\rho = \frac{\sqrt{8\nu}}{\kappa}$. As this is the distance at which the correlation is close to 0.1952, we can conclude that the data are pigeonholed by a strong spatial correlation which diminishes gradually with distance [9]. As depicted by [8], [9] and [23], The Marginal variance, σ_γ^2 can be approximated with,

$$\sigma_\gamma^2 = \frac{1}{4\pi\kappa^2\tau}$$

From Table 2 above, σ_γ^2 was 0.276. We perceived that more variation is explained by the spatial term rather than by the error. Moreover, the high value of the AR(1) temporal correlation coefficient (0.7263) confirms that the short-term persistence of the Infected rate of Visceral leishmaniasis in human(VLH).

Discussion

In this research, we had applied the Stochastics Partial Differential Equation (SPDE) approach for a spatiotemporal hierarchical model for Visceral leishmaniasis in human (VLH) that encompasses a GF and a state process is accompanying with an autoregressive order one temporal dynamics and the spatially correlated error term, together with the effect of land-living cover, metrological, demographic, sociodemographic and geographical covariates in endemic area of Amhara regional state, Ethiopia.

We used a constrained refined Delaunay triangulation (CRDT) for SPDE approximation and mesh construction as a result of the CRDT part is the least relevant as compared to the unconstrained type, it just means that one can specify that certain polygons/line must be part of the triangulation edges. The only constrained edges are usually the boundary edges. The quality of the spde approximation depends mostly on the "refined" aspect. Meshes with small and well-formed triangles (short edges, and no small acute angles) provide closer approximations. Meshes with large and/or sharp angled triangles exhibit approximation artifacts; the fields are always conditionally deterministic, given the values at the mesh vertices. The posterior distributions of covariate effect sizes demonstrate that the mean weekly temperature had a positive association with the infected rate of VLH, This study is in accordance with, the study by [16] in a similar region. Indicates that the outbreak and infection of Visceral leishmaniasis in human (VLH) disease are distributed in the arid and semi-arid parts of the country, high temperatures increase VLH infection rate, this result is in line with the study by [24], [25]. The Average number of Household (ANH) and Average number of Persons Per housing Unit (ANPPHU) had a positive association with the infected the Infected Rate of Visceral

Leishmaniasis (VLH). This related to the source of reservoir hosts and human infection anthroponotic leishmaniases, in which the reservoir host is human, the infection and disease transmitted from human to human [1]. Net Migration Rate (NMR) and Mortality rate (MR) had a significant positive association with the Infected Rate of Visceral Leishmaniasis (VLH), this agrees with the study by [6] indicates that migration of laborers and farmers to and from endemic areas increases the infected rate of VLH. Additionally, Health Facility Coverage in Percentage (HFCIP), Education coverage in percentage (ECIP) and Clean water coverage in percentage (CWC) had a significant positive association with the Infected Rate of Visceral Leishmaniasis (VLH), lack of awareness about the disease and symptom, shortage of drug expansions and health facilities and shortage of clean water coverage in the endemic area of Visceral Leishmaniasis (VL) as a result there will be a significant increment of the infection of Visceral Leishmaniasis. Well organized and planned drug expansions in the infected and endemic area, disease controlling mechanisms, better hygiene and sufficient clean water coverage and medical treatments to the infected persons decrease the death and infection rate of VLH. Moreover, Population density per square kilometer (PDPSK) had a significant negative association with the Infected Rate of Visceral Leishmaniasis in humans (VLH), the infection and disease transmitted from human to human, in which the reservoir host is human if the population is very dense and increase the infection of the disease. Sex ratio (SR) also had a significant positive association with on the infected rate of visceral leishmaniasis in human (VLH), Males are more exposed to develop the infection and disease as they are usually engaged in farms, trades and other agricultural activities, which will make them progressively available to the sandfly chomp [6], [26]. From random effects, we get a value of 526 km for the empirically derived correlation range $\rho = \frac{\sqrt{8\nu}}{\kappa}$. This indicates that the distance at which the correlation is close to 0.1952, we can say that the data are characterized by a strong spatial correlation which diminishes gradually with distance, This result is in line with the study by [20], [27], [28]. σ_γ^2 was 0.276 and σ_ϵ^2 was 0.0176. We designated that more variation is explained by the spatial term rather than by the error, This result is in line with the study by [20], [29], [30]. Additionally, the high value of the AR(1) temporal correlation coefficient confirms the short-term persistence of Infected rate of Visceral leishmaniasis. There are different possibilities for modifying this study. This study developed dynamic spatiotemporal modeling of the Infected Rate of Visceral Leishmaniasis in humans (VLH) it examined the spatial and temporal effects in addition to the covariate effects through SPDE approach using hierarchical Bayesian modeling. The future studies can modify this study by incorporating the interaction effects of time and space. Also, this study only considered infection of only Infected Rate of Visceral Leishmaniasis in humans (VLH) future studies can extend the study by considering other additional Leishmaniasis infection cases like cutaneous, mucosal or post kala-azar dermal Leishmaniasis (PKDL) in the same study area. We used too many metrological, demographic, sociodemographic and geographic covariates in addition to spatial and temporal effects which lead us model complexity and posterior approximation took too much time. Moreover, future studies can incorporate seasonal dummies to identify the seasonal effects on the infection of Visceral Leishmaniasis in humans (VLH). Lastly, Visceral Leishmaniasis in humans (VLH) case data from stations were very poorly recorded and required very cautious integration. Future studies should modify this study by considering the above limitations in the endemic area of the region.

Even though the above limitations our study identify the spatiotemporal covariates associated with Infected Rate of Visceral Leishmaniasis in humans (VLH) in the region using INLA method. We believe that our modeling result could be used as information and motivation for

other studies for identifying spatiotemporal determinants of VLH infection in the endemic areas of Amhara regional state, Ethiopia.

Conclusions

In this study, we investigated the dynamic spatiotemporal modeling of the infected rate of visceral leishmaniasis (VLH) through SPDE approach. The model involves a Gaussian field (GF) its state process is an autoregressive order one of temporal dynamics and the spatially correlated innovations. Our study had confirmed that both metrological, demographic, sociodemographic and geographic covariates had a significant association with the infected rate of visceral leishmaniasis (VLH) in the region.

Furthermore, Integrated nested Laplace approximations (INLA) is a computationally proficient strategy for incorporating both spatial and temporal effects into spatiotemporal general mixed effect models. Spatiotemporal model fitting is commonly exceptionally complex to execute and requires a powerful computing machine, a long running time, or both. INLA package runs utterly within the commonly utilized R statistical software and is relatively simple to implement with intermediate levels of programming expertise. In ongoing research, we are working on by incorporating the use of expert knowledge of spatiotemporal prediction it just an extension of this work. We also inspire future research in spatiotemporal disease ecology and rare events prediction to consider the INLA SPDE approach for spatiotemporal mixed modeling.

Supporting information

S1 File. S1_File.rar.
(RAR)

S2 File. R code for posterior estimation.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Anteneh Asmare Godana.

Data curation: Anteneh Asmare Godana.

Formal analysis: Anteneh Asmare Godana.

Investigation: George Otieno Orwa.

Methodology: Anteneh Asmare Godana, Samuel Musili Mwalili.

Software: Anteneh Asmare Godana, George Otieno Orwa.

Supervision: Samuel Musili Mwalili.

Validation: Anteneh Asmare Godana.

Visualization: Anteneh Asmare Godana, George Otieno Orwa.

Writing – original draft: Anteneh Asmare Godana, Samuel Musili Mwalili.

Writing – review & editing: Anteneh Asmare Godana.

References

1. Alvar Jorge and Velez Ivan D and Bern Caryn and Herrero Merce and Desjeux Philippe and Cano Jorge and Jannin Jean and den Boer Margriet and WHO Leishmaniasis Control Team and others Leishmaniasis worldwide and global estimates of its incidence. PloS one. 2012 May; 7(5): e35671. https://doi.org/10.1371/journal.pone.0035671 PMID: 22693548

2. Ohno S. Nackers Fabienne and Mueller Yolanda Kathrin and Salih Niven and Elhag Mousab Siddig and Elbadawi Mobarak Elnour and Hammam Omer and Mumina Ann and Atia Atia Abdalla and Etard Jean-François and Ritmeijer Koert and others. Determinants of visceral leishmaniasis: a case-control study in Gedaref State, Sudan. PLoS neglected tropical diseases. 2015 Nov 6; 9(11):e0004187. https://doi.org/10.1371/journal.pntd.0004187

3. Druzian Angelita F and de Souza Albert S and de Campos Diogo N and Croda Julio and Higa Minoru G Jr and Dorval Maria Elizabeth C and Pompilio Mauricio A and de Oliveira Polliana A and Paniago Anamaria MM. Risk factors for death from visceral leishmaniasis in an urban area of Brazil. PLoS neglected tropical diseases. 2015 Aug 14; 9(8):e0003982. https://doi.org/10.1371/journal.pntd.0003982 PMID: 26274916

4. Das A, Karthick M, Dwivedi S, Banerjee I, Mahapatra T, Srikantiah S, Chaudhuri I. Epidemiologic correlates of mortality among symptomatic visceral leishmaniasis cases: findings from situation assessment in high endemic foci in India. PLoS neglected tropical diseases. 2016 Nov 21; 10(11):e0005150. https://doi.org/10.1371/journal.pntd.0005150 PMID: 27870870

5. Elnaiem DE, Hassan HK, Osman OF, Maingon RD, Killick-Kendrick R, Ward RD. A possible role for Phlebotomus (Anaphlebotomus) rodhaini (Parrot, 1930) in transmission of Leishmania donovani. Parasites & vectors. 2011 Dec; 4(1):238. https://doi.org/10.1186/1756-3305-4-238

6. Leta S, Dao TH, Mesele F, Alemayehu G. Visceral leishmaniasis in Ethiopia: an evolving disease. PLoS neglected tropical diseases. 2014 Sep 4; 8(9):e3131. https://doi.org/10.1371/journal.pntd.0003131 PMID: 25188253

7. Desjeux P. The increase in risk factors for leishmaniasis worldwide. Transactions of the royal society of tropical medicine and hygiene. 2001 May 1; 95(3):239–43. https://doi.org/10.1016/S0035-9203(01)90223-8 PMID: 11490989

8. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2011 Sep 1; 73(4):423–98. https://doi.org/10.1111/j.1467-9868.2011.00777.x

9. Blangiardo M, Cameletti M. Spatial and spatio-temporal Bayesian models with R-INLA. John Wiley & Sons; 2015 Jun 2.

10. Cressie N. Statistics for spatial data. Terra Nova. 1992 Sep; 4(5):613–7. https://doi.org/10.1111/j.1365-3121.1992.tb00605.x

11. Ribeiro Jr PJ, Diggle PJ. Analysis of geostatistical data. The geoR package, version. 2006:1-6.

12. Cressie N, Johannesson G. Fixed rank kriging for very large spatial data sets. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2008 Feb 1; 70(1):209–26. https://doi.org/10.1111/j.1467-9868.2007.00633.x

13. Carlin BP, Gelfand AE, Banerjee S. Hierarchical modeling and analysis for spatial data. Chapman and Hall/CRC; 2014 Sep 12.

14. Rue H, Held L. Gaussian Markov random fields: theory and applications. CRC press; 2005 Feb 18.

15. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the royal statistical society: Series b (statistical methodology). 2009 Apr 1; 71(2):319–92. https://doi.org/10.1111/j.1467-9868.2008.00700.x

16. Gadisa E, Tsegaw T, Abera A, Elnaiem DE, den Boer M, Aseffa A, Jorge A. Eco-epidemiology of visceral leishmaniasis in Ethiopia. Parasites & vectors. 2015 Dec; 8(1):381. https://doi.org/10.1186/s13071-015-0987-y

17. Utzinger J, Becker SL, Knopp S, Blum J, Neumayr AL, Keiser J, Hatz CF. Neglected tropical diseases: diagnosis, clinical management, treatment and control. Swiss medical weekly: official journal of the Swiss Society of Infectious Diseases, the Swiss Society of Internal Medicine, the Swiss Society of Pneumology. 2012; 142.

18. Tsegaw T, Gadisa E, Seid A, Abera A, Teshome A, Mulugeta A, Herrero M, Argaw D, Jorge A, Aseffa A. Identification of environmental parameters and risk mapping of visceral leishmaniasis in Ethiopia by

using geographical information systems and a statistical approach. Geospatial health. 2013; 7(2):299–308. https://doi.org/10.4081/gh.2013.88 PMID: 23733292

19. Murray HW, Berman JD, Davies CR, Saravia NG. Advances in leishmaniasis. The Lancet. 2005 Oct 29; 366(9496):1561–77. https://doi.org/10.1016/S0140-6736(05)67629-5

20. Cameletti M, Lindgren F, Simpson D, Rue H. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Advances in Statistical Analysis. 2013 Apr 1; 97(2):109–31. https://doi.org/10.1007/s10182-012-0196-3

21. Cressie N, Wikle CK. Statistics for spatio-temporal data. John Wiley & Sons; 2015 Nov 2.

22. Brenner S, Scott R. The mathematical theory of finite element methods. Springer Science & Business Media; 2007 Dec 14.

23. Hefley TJ, Hooten MB, Hanks EM, Russell RE, Walsh DP. Dynamic spatio-temporal models for spatial data. Spatial statistics. 2017 May 31; 20:206–20. https://doi.org/10.1016/j.spasta.2017.02.005

24. Berrozpe PE, Lamattina D, Santini MS, Araujo AV, Torrusio SE, Salomón OD. Spatiotemporal dynamics of Lutzomyia longipalpis and macro habitat characterization using satellite images in a leishmaniasis endemic city in Argentina. Medical and veterinary entomology. 2018 Sep 9.

25. Abdullah AY, Dewan A, Shogib MR, Rahman MM, Hossain MF Environmental factors associated with the distribution of visceral leishmaniasis in endemic areas of Bangladesh: modeling the ecological niche. Tropical medicine and health. 2017 Dec; 45(1):13. https://doi.org/10.1186/s41182-017-0054-9 PMID: 28515660

26. Malaria Consortium. Leishmaniasis control in eastern Africa: Past and present efforts and future needs. Situation and gap analysis. 2010 Nov;86.

27. Bermudi PM, Guirado MM, Rodas LA, Dibo MR, Chiaravalloti-Neto F. Spatio-temporal analysis of the occurrence of human visceral leishmaniasis in Araçatuba, State of São Paulo, Brazil. Revista da Sociedade Brasileira de Medicina Tropical. 2018 Aug; 51(4):452–60. https://doi.org/10.1590/0037-8682-0505-2017 PMID: 30133627

28. Buckingham-Jeffery E, Hill EM, Datta S, Dilger E, Courtenay O. Spatio-temporal modelling of Leishmania infantum infection among domestic dogs: a simulation study and sensitivity analysis applied to rural Brazil. bioRxiv. 2018 Jan 1:418558.

29. Dewan A, Abdullah AY, Shogib MR, Karim R, Rahman MM. Exploring spatial and temporal patterns of visceral leishmaniasis in endemic areas of Bangladesh. Tropical medicine and health. 2017 Dec; 45 (1):29. https://doi.org/10.1186/s41182-017-0069-2 PMID: 29167626

30. Myer MH, Campbell SR, Johnston JM. Spatiotemporal modeling of ecological and sociological predictors of West Nile virus in Suffolk County, NY, mosquitoes. Ecosphere. 2017 Jun; 8(6):e01854. https://doi.org/10.1002/ecs2.1854 PMID: 30147987