Spatiotemporal mapping and detection of mortality cluster due to cardiovascular disease with Bayesian hierarchical framework using integrated nested Laplace approximation: A discussion of suitable statistic applications in Kersa, Oromia, Ethiopia

Melkamu Dedefo,[1] Henry Mwambi,[1] Sileshi Fanta,[1] Nega Assefa[2,3]

[1]School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa; [2]Kersa Health and Demographic Surveillance System (Kersa HDSS), Harar, Ethiopia; [3]College of Health and Medical Sciences, Haramaya University, Harar, Ethiopia

Correspondence: Melkamu Dedefo, School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Private Bag X01, 3201 Pietermaritzburg, South Africa.
E-mail: melkyed@gmail.com

Acknowledgements: the authors would like to acknowledge Haramaya University, Kersa HDSS for providing data for the study. The authors are also indebted to the study communities for their unreserved cooperation and willingness to participate in the health and demographic surveillance tracking of data.

Key words: Cardiovascular diseases; Integrated nested Laplace approximations; Second order random walk; Bayesian hierarchical framework; Spatio-temporal models; Kersa HDSS

Contributions: MD, HM, SF and NA conceived the concept. MD analysed the data and wrote the first draft. MD, HM, SF and NA contributed to the writing of the paper and approved the final manuscript.

Conflict of interest: the authors declare no potential conflict of interest.

Funding: this study was supported through the DELTAS Africa Initiative SSACAB (Grant No. 107754/Z/15/Z). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant No. 107754/Z/15/Z) and the UK government. We would like to thank Kersa HDSS of INDEPTH network and Haramaya University for providing the data. The views expressed in this publication are those of the authors and not necessarily those parties mentioned above.

Received for publication: 3 March 2018.
Revision received: 14 October 2018.
Accepted for publication: 15 October 2018.

Geospatial Health 2018; 13:681
doi:10.4081/gh.2018.681

Abstract

Cardiovascular diseases (CVDs) are the leading cause of death globally and the number one cause of death globally. Over 75% of CVD deaths take place in low- and middle-income countries. Hence, comprehensive information about the spatio-temporal distribution of mortality due to cardio vascular disease is of interest. We fitted different spatio-temporal models within Bayesian hierarchical framework allowing different space-time interaction for mortality mapping with integrated nested Laplace approximations to analyze mortality data extracted from the health and demographic surveillance system in Kersa District in Hararege, Oromia Region, Ethiopia. The result indicates that non-parametric time trends models perform better than linear models. Among proposed models, one with non-parametric trend, type II interaction and second order random walk but without unstructured time effect was found to perform best according to our experience and. simulation study. An application based on real data revealed that, mortality due to CVD increased during the study period, while administrative regions in northern and south-eastern part of the study area showed a significantly elevated risk. The study highlighted distinct spatiotemporal clusters of mortality due to CVD within the study area. The study is a preliminary assessment step in prioritizing areas for further and more comprehensive research raising questions to be addressed by detailed investigation. Underlying contributing factors need to be identified and accurately quantified.

Introduction

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) are the leading cause of death globally, currently accounting for 17.9 million deaths per year and projected to increase to more than 23.6 million by 2030. More than three quarters of CVD deaths take place in low- and middle-income countries (Mathers and Loncar, 2006; Laslett *et al.*, 2012; WHO, 2018). It would be a catastrophe for developing countries to have this additional burden, as they are already faced with a multitude of other challenges, such as poor socio-economic indices, high prevalence of infectious diseases and a trend towards high-caloric nutrition and sedentary lifestyles (Popkin *et al.*, 2012). Even

worse, a greater proportion of younger people are also affected - more than 80 per cent of deaths resulting from CVD in developing countries occur between the ages of 30 and 70 years (Bloom *et al.*, 2012; Leeder *et al.*, 2012; WHO, 2018).

Reliable evidence on CVD mortality, its causes and trends should aid in developing appropriate interventions are such data are highly needed when evaluating the global and regional health impact of the CVDs. Monitoring trends and distribution of mortality due to CVD have substantial value towards progress with respect to health. In the absence of routine mortality statistics, data from the health and demographic surveillance system (HDSS) provide a valuable source for estimating trends and distribution of mortality on a longitudinal basis (Hammer *et al.*, 2006; Assefa *et al.*, 2016; Dedefo *et al.*, 2016a; Dedefo *et al.*, 2016b).

It is necessary to limit high risk areas where certain adverse health effects are most likely to occur as intervention at a wider population level is too expensive to implement. There is thus a need to identify more affected areas where adverse health outcomes seem to aggregate and to develop specific health strategies targeting these regions (Benzler and Sauerborn, 1998; Sankoh *et al.*, 2001). The key benefit of such mapping is that it allows public health officials to identify clusters of areal units that exhibit elevated disease risks, which in turn enables interventions to be appropriately targeted at the communities with greatest needs. Such interventions can for example include a vaccination program or a public awareness campaign about potential risk factors. Thus, in addition to the obvious public health benefit, the identification of high-risk clusters through the use of disease maps can help to reduce health-care costs (Anderson *et al.*, 2014). In addition, cluster detection is an important part of spatial epidemiology because it may help suggesting potential factors associated with disease and therefore guide investigations of the nature of the disease and its aetiology (Torabi and Rosychuk, 2011).

Spatio-temporal disease mapping models are useful in describing the temporal evolution of geographic patterns of mortality and disease rates. The results from such analyses would not only help decision-makers and investigators to formulate hypotheses regarding the aetiology of a disease, to look for risk factors and to allocate resources efficiently, but also to design intervention programmes in advance. Among the benefits of using space-time models in disease mapping is to borrow strength from spatial and temporal neighbours to reduce the high variability, which is characteristic of classical disease/mortality risk estimators, such as the standardized mortality ratio (SMR); particularly, while dealing with rare diseases or areas with small populations (Ugarte *et al.*, 2014).

Use of spatiotemporal analysis has increasingly been applied in epidemiological research in recent years (Kulldorff and Nagarwalla, 1995; Gangnon and Clayton, 1998; Elliot *et al.*, 2000; Gangnon and Clayton, 2000; Gómez-Rubio *et al.*, 2005; Gómez-Rubio *et al.*, 2006; Gómez-Rubio *et al.*, 2009; Bilancia and Demarinis, 2014). Developments in data accessibility and advanced analytic approaches have created new opportunities to observe variations in disease occurrence rates at the small-area level (Elliot *et al.*, 2000). Many methods have been proposed for the detection of disease clusters, most of them based on moving windows, such as Kulldorff's Spatial Scan Statistics (https://www.satscan.org/). Among the most important exploratory methods for cluster detection are those that identify significant clusters in space and/or time (Hjalmars *et al.*, 1994; Boyle *et al.*, 1996; Hjalmars *et al.*, 1996; Kulldorff, 1998, Hjalmars *et al.*, 1999; Torabi and Rosychuk, 2011; Gómez-Rubio *et al.*, 2006).

In this paper we fitted different spatiotemporal models within the Bayesian hierarchical framework allowing different and flexible space time interactions for mortality mapping based on standard spatial mapping models using integrated nested Laplace approximations (INLA) (http://www.r-inla.org/). The approach proposed is the use of a Bayesian inference tool in latent Gaussian models (Rue and Held, 2005). Our objectives were to detect significant risk clusters of mortality within a population-based surveillance site in eastern Ethiopia applying models to analyse mortality data extracted from Kersa District HDSS for the period of 2007-2016.

Spatio-temporal disease mapping

Disease mapping can be used to assess the spatial pattern of a disease, to estimate a measure of spatially observed health outcomes and to identify clusters (Pascutto *et al.*, 2000; Rezaeian *et al.*, 2007; Everitt and Dunn, 2011). The data are obviously discrete in nature, as they are counts of diseases or deaths in specific area. Public health data are often aggregated over small administrative areas due to issues of confidentiality. Still, household- and individual-level data are often available for modelling, both at the individual level and aggregated, by generalized linear mixed models (GLMM) (McCullagh and Nelder, 1989) to identify mortality and disease clustering. A wide range of spatio-temporal models for disease mapping have been proposed in the literature, most of them based on conditional autoregressive (CAR) models extending the well-known BYM model (Besag *et al.*, 1991). The key to applying spatio-temporal models in disease-mapping studies is to borrow strength from spatial and temporal neighbours to reduce the high variability inherent in classical risk estimators, such as the SMR; in particular, when studying rare diseases or areas with low populations. Models used in spatio-temporal disease mapping are usually GLMMs dealing with counts assuming a Poisson distribution. These models are formulated within a hierarchical Bayesian framework. Let the region in the study area be divided into non-overlapping n small administrative areal units (kebeles) labelled as $i = 1,\ldots, n$. Data are available for each area i and time t, $t = 1,\ldots, T$. A response y_{it} is observed in each areal unit and time. Here, it is important to account for differences in population demographics across the study region, since some sub-regions are likely to contain a larger at-risk population. For example, areas which have a higher percentage of elderly people are likely to have higher rates of heart disease than those with a younger population, but this does not necessarily mean that there is any underlying difference in disease risk rate between the regions. We can account for these demographic differences by constructing a set of expected disease counts e_{it}, where e_{it} is the expected number of disease cases in area i at time t.

Conditional on the relative risk (ρ), the observed number of counts y_{it} is assumed to be Poisson distributed with mean $\mu_{it} = e_{it}\rho_{it}$. (Eq. 1):

$$y_{it}|\rho_{it} \sim Poisson(e_{it}\,\rho_{it})$$

$$\text{from } \mu_{it}=e_{it}\,\rho_{it},\ log\,\mu_{it} = loge_{it} + \log\rho_{it} \qquad \text{Eq. 1}$$

Most of the time the interest is in modelling the relative risk usually using log link function and depending on that different models can be specified, *e.g.* Eq. 2:

$$\log \rho_{it} = a_0 + \upsilon_i + \nu_i + Temporal_t \quad \text{Eq. 2}$$

where α_0 is the intercept quantifying the logarithm of average risk for all areas, υ_i spatially structured residual or random effects, and ν_i a spatially unstructured random effect or residual. A temporal component $Temporal_t$ with $t = 1,\ldots, T$ can be specified with some parametric or non-parametric structure.

Parametric time trend models

We consider here a parametric Bayesian model with a linear time trend based on Bernardinelli *et al.* (1995). The model is an extension of the BYM model (Besag *et al.*, 1991), which includes two spatial effects: the unstructured random effect with a Gaussian exchangeable prior; and the spatially structured random effect of an intrinsic conditional autoregressive prior (iCAR) with an additional linear time trend and a differential time trend for each small area, *i.e.* Eq. 3:

$$\log \rho_{it} = a_0 + \upsilon_i + \nu_i + (\beta + \delta_i) * t \quad \text{Eq. 3}$$

where α_0 is the intercept quantifying the logarithm of average risk for all areas, υ_i the spatially structured residual or random effect, ν_i the spatially unstructured residual or random effect, β an overall linear time trend, and δ_i the differential trend which identifies the interaction between linear time and space.

Non-parametric dynamic trend models

In parametric trend models, a linearity constraint is imposed on the differential temporal trend. However, this may not be the case in practice, where there is some non-linearity in temporal trends due to a day-to-day development in treatments, intervention programmes and advancement of research findings in general. Thus, it is important to relax the linearity assumption imposed. In this paper, we considered different nonparametric models extending the work of Knorr-Held and Rasser (2000). Thus, the log risk is modelled as Eq. 4:

$$\log \rho_{it} = a_0 + \upsilon_i + \nu_i + \gamma_t + \phi_t \quad \text{Eq. 4}$$

where α_0 is the intercept quantifying the logarithm of average risk for all areas, υ_i spatially structured residual or random effect, ν_i spatially unstructured residual or random effect, γ_t the temporally structured effect, and ϕ_t the unstructured temporal effect.

Space-time interaction models

To better explain differences in the time trend of diseases for different administrative regions, it would be necessary to expand the previous non-parametric dynamic model to allow for a flexible interaction between space and time where accordingly the log risk is modelled as Eq. 5:

$$\log \rho_{it} = a_0 + \upsilon_i + \nu_i + \gamma_t + \phi_t + \delta_{it} \quad \text{Eq. 5}$$

where α_0 is the intercept quantifying the logarithm of average risk for all areas, υ_i spatially structured residual or random effect, ν_i spatially unstructured residual or random effect, γ_t the temporally structured effect, ϕ_t the unstructured temporal effect, and δ_{it} the differential trend which identifies the interaction between time and space.

Modelling the spatial dependency structure

The most commonly used model for modelling the spatial dependence is the BYM model (Besag *et al.*, 1991), but her there is a problem in identifying the structured and unstructured dependence. In this respect Leroux *et al.* (2000) proposed an alternative model formulation to make the compromise between unstructured and structured variation more explicit. Here, $\xi = u + v$ is assumed to follow a normal distribution with mean zero and covariance matrix of $(\sigma^2_\xi(\lambda_\xi R_\xi + (1-\lambda_\xi) I_n^{-1})(\sigma^2_\xi, \lambda_\xi)$.

For the vector of spatial effect $\xi = (\xi_1,\ldots,\xi_n)'$, $\xi \sim N(0, \sigma^2_\xi(\lambda_\xi R_\xi + (1-\lambda_\xi) I_n^{-1})(\sigma^2_\xi, \lambda_\xi)$ where $\lambda_\xi \in [0,1]$ denotes a mixing or smoothing parameter. I_n is an identity matrix of dimension $n \times n$ and R_ξ the spatial neighborhood matrix.

The Leroux CAR prior accounts for spatially unstructured random effect with an exchangeable prior $\xi \sim N(0, \sigma^2_\xi I_n)$ when $\lambda_\xi = 0$ and consider spatially structured effect with intrinsic CAR prior when $\lambda_\xi = 1$, which finally gives the log risk as Eq. 6:

$$\log \rho_{it} = a_0 + \xi_i + \gamma_t + \phi_t + \delta_{it} \quad \text{Eq. 6}$$

Here, the unstructured temporal random effects ϕ_t are modelled as independent and identically distributed normal random variables with mean 0 and variance σ^2_ϕ as $\phi \sim N(0, \sigma^2_\phi I_T)$, where $\phi = (\phi_1,\ldots \phi_T)'$ and I_T constitute an identity matrix of dimension $T \times T$. For structured temporal effects $\gamma = (\gamma_1,\ldots \gamma_T)'$ a random walk of first and second order can be considered and its distribution given by $\gamma \sim N(0, \sigma^2_\gamma R_t^-)$, where R_t denotes the structure matrix of first and second order random walk where the symbol "-" denotes the Moore-Penrose generalized inverse. The interaction term $\delta = (\delta_{11},\ldots \delta_{nT})'$, also assumed to be distributed normally as $\delta \sim N(0, \sigma^2_\delta R_\delta)$, where σ^2_δ is the variance parameter and R_δ the matrix structure of main effects identifying the type of temporal and/or spatial dependence between the elements of δ and defined according to the Kronecker product (Gilks and Roberts, 1996). In this work the four types of interactions in Knorr-Held and Rasser (2000) based on the structure matrix are considered for R_δ.

Type I interaction

Type I assumes that the two unstructured effects interact but result in no structure in space and /or time. In type I interaction, all δ_{nT}'s are *a priori* independent and the structure matrix or precision matrix can be written as Eq. 7:

$$R_\delta = R_\nu \otimes R_\phi = I \otimes I = I, \text{ consequently } \delta_{nt} \sim N\left(0, 1/\sigma^2_\delta\right) \quad \text{Eq. 7}$$

The rank of R_δ for both first order and second order random walk of γ is $n \cdot T$.

Type II interaction

Type II interactions combine unstructured spatial effects with structured temporal effect and the structure matrix can be given as Eq. 8:

$$R_\delta = R_\nu \otimes R_\gamma \quad \text{Eq. 8}$$

where $R_\nu = I$ and R_γ is the neighbourhood structure which may be specified through a first or second order random walk which means that each δ_i. $i = 1,\ldots,n, (\delta_{i1},\ldots \delta_{iT})$ follows a first order or second order random walk independently of all other regions. This

type of interaction would be better suited if the temporal trend is independent from one another, but do not have any structure in space. Here the matrix R_δ has the rank of $n \cdot (T\text{-}1)$ for first order random walk of γ and $n \cdot (T\text{-}2)$ for second order random walk of γ.

Type III interaction

Type III interaction combines structured spatial main effects and unstructured temporal effects and the associated structure matrix given as Eq. 9:

$$R_\delta = R_v \otimes R_\phi \quad \text{Eq. 9}$$

where R_ϕ = I and R_v can be seen as different spatial trends for each year without any temporal structure and each δ_t, $i = 1,\ldots,T$, $(\delta_{t1},\ldots \delta_{tn})$ follows an (independent) intrinsic conditional autoregression. The matrix R_δ has the rank of $(n - 1) \cdot T$ for first order random walk of γ and $(n - 1) \cdot T$ for second order random walk of γ.

Type IV interaction

Type IV is the most complex type of interaction between spatially and temporally structured effects, where δ_{it}'s are completely dependent over space and time. This type of interaction would be suitable if temporal trends are different from region to region, but are more likely to be similar for adjacent regions.

The structure matrix will be given as Eq. 10:

$$R_\delta = R_v \otimes R_\gamma \quad \text{Eq. 10}$$

This matrix has the rank of $(n - 1) \cdot (T - 1)$ for first order random walk of γ and $(n - 1) \cdot (T - 2)$ for second order random walk of γ.

The models

With a combination of the four types of interaction given above and by choosing different priors for the structured time effect we proposed nineteen different models for both the parametric and non-parametric cases. Three parametric models were considered based on different parametric form of the trend. The log risk model is specified as Eq. 11:

$$\log \rho_{it} = a_0 + \xi_i + (\beta + \delta_i) * t \quad \text{Eq. 11}$$

Model 1

The Leroux CAR prior is considered for the spatial effects. For model 1 we assumed exchangeable distribution for the differential effect, that is δ_i, $i = 1,\ldots, I$ are independently and identically distributed normal random variables $\delta \sim N(0,\sigma_\delta^2)$.

Model 2

The Leroux CAR prior is considered for the spatial effects. For model 2, iCAR prior is assumed for the differential effect, that is (Eq. 12):

$$\delta_i \,|\, \delta_{j \neq i} = \left(\frac{1}{m_i}\sum_{i \sim j} \delta_i, \frac{\sigma^2}{m_i}\right) \quad \text{Eq. 12}$$

where $i \sim j$ indicates that area i and j are neighbours, m_i the number of neighbours of area i, and σ^2 is the variance component. Thus, the joint distribution of the random effects can be written as (Eq. 13):

$$\delta \sim N(0, \sigma^2 R_s^-), \text{ where } \delta = (\delta_1,\ldots \delta_n)' \quad \text{Eq. 13}$$

The R_s matrix is determined by the spatial neighborhood structure with non-diagonal elements (Eq. 14):

$$(R_s)_{ij} = \begin{cases} -1 & \textit{if } i \textit{ and } j \textit{ are neighbours} (\text{share a common border}) \\ 0 & \textit{otherwise} \end{cases} \quad \text{Eq. 14}$$

The diagonal element $(R_s)_{ii}$ gives the number of neighbours (m_i) of area and the symbol – indicates the Moore-Penrose generalized inverse.

Model 3

The Leroux CAR prior is considered for both the spatial effects and for the differential trend.

Non-parametric models

Models 1 and 2

Here, we considered a non-parametric trend model without the interaction term but allow for first order and second order random walk respectively for the structured time effect, *i.e.* (Eq. 15):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t + \phi_t \quad \text{Eq. 15}$$

Models 3 and 4

Here, we considered a non-parametric trend model with type I interaction with first order and second order random walk respectively for the structured time effect, *i.e.* (Eq. 16):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t + \phi_t + \delta_{it} \text{ with } R_\delta = R_v \otimes R_\phi = \text{I} \otimes \text{I} = \text{I} \quad \text{Eq. 16}$$

Models 5 and 6

Here, we considered non-parametric trend with type II interaction with first order and second order random walk respectively for the structured time effect, *i.e.* (Eq. 17):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t + \phi_t + \delta_{it} \text{ with } R_\delta = R_v \otimes R_\gamma \quad \text{Eq. 17}$$

Models 7 and 8

Here, we considered non-parametric trend with type III interaction with first order and second order random walk respectively for the structured time effect, *i.e.* (Eq. 18):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t + \phi_t + \delta_{it} \text{ with } R_\delta = R_v \otimes R_\phi \quad \text{Eq. 18}$$

Models 9 and 10

Here, we considered non-parametric trend with type IV interaction with first order and second order random walk respectively for the structured time effect, *i.e.* (Eq. 19):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t + \phi_t + \delta_{it} \text{ with } R_\delta = R_v \otimes R_\gamma \quad \text{Eq. 19}$$

Models 11 and 12

Here, we considered non-parametric trend without unstructured time effect, without interaction effect with first order and second order random walk, *i.e.* (Eq. 20):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t \quad \text{Eq. 20}$$

Models 13 and 14

For model 13 and 14, we considered non-parametric trend without unstructured time effect, with type II interaction with first order and second order random walk respectively, *i.e.* (Eq. 21):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t + \delta_{it} \text{ with } R_\delta = R_v \otimes R_\gamma \quad \text{Eq. 21}$$

Models 15 and 16

Here, we considered non-parametric trend without unstructured time effect, with type IV interaction with first order and second order random walk respectively, *i.e.* (Eq. 22):

$$\log \rho_{it} = \alpha_0 + \xi_i + \gamma_t + \delta_{it} \quad \text{with} \quad R_\delta = R_\nu \otimes R_\gamma \qquad \text{Eq. 22}$$

Priors

Based on recommendation from different literature sources, more suitable user-defined hyperpriors have been given using suitable expression in INLA. Specifically, the non-informative uniform prior distributions $\sigma \sim U(0, \infty)$ and $\lambda \sim U(0, 1)$ have been defined for hyperparameters of random effects in the modelling process.

Approximation of Bayesian inference using integrated nested Laplace approximations

Here, a full Bayesian approximation was implemented to fit the proposed models with Bayesian hierarchical framework in over three stages. The first stage was the observational model and the second stage the latent Gaussian Markov random field (GMRF) with precision matrix R which in turn is controlled by hyperparameters which are not necessarily Gaussian (third stage).

INLAs were proposed for estimation in Bayesian inference with latent Gaussian field according to Rue *et al.* (2009). The methodology would be particularly attractive, if the latent Gaussian model were a GMRF us used by Rue and Held (2005) with precision matrix controlled by a hyperparameter τ (details about implementation of INLA method can be found in Rue *et al*., 2009). Note that estimation using the standard Markov chain Monte Carlo (MCMC) has a higher computational cost in terms of time so INLA is faster. In addition, parameter samples may have high correlation resulting in a large Monte Carlo error during estimation. Particularly the application of MCMC in spatiotemporal analysis is difficult because of the strong posterior dependence between components of the latent spatial or spatiotemporal fields, while INLA provides very accurate approximations to the posterior marginals in a relatively short computational time. All components in the models proposed in this work can be modelled using GMRF and INLA can be run on Windows, Mac and Linux in the software environment R using the R-package r-inla (Rue *et al.*, 2017). In our case we used version 17.06.20 released 2017-06-20.

Mortality analysis due to centers for disease control and prevention in Kersa health and demographic surveillance system 2007-2016

Kersa is a district in the eastern part of Ethiopia. According to the Ethiopian Government census of 2007, it has a total population of 172,626; out of which, 6.87 % are urban dwellers (Central Statistical Agency, 2010). The Kersa HDSS, established in 2007 with 10 522 households and 50,830 population to track demographic and health changes in the community, is a member of the INDEPTH Network, a network of HDSSs in Africa, South America and Asia.

Kersa HDSS is currently a platform for various health related research by the college of Health and Medical Sciences in Haramaya University with a broader vision to become a centre of excellence for health science research in Eastern Africa. After the first census, a continuous registration system for demographic and health related events have been operational in the whole of the HDSS area (Assefa *et al.*, 2016; Dedefo *et al.*, 2016a; Dedefo *et al*., 2016b). Data are entered into the HRS-2 relational database. The sex ratio and average number of persons per household was 1.0 and 5.1, respectively. At the end of 2016 the population was 130,358. Until the end of 2016, 20,935 births and 5,195 deaths were registered. Over 85% of births and deaths occurred at home. The annual net population growth ranged from -0,1 to 1.6. Meanwhile, the population growth rate at the national level ranged from 1.63 to 2.94. The majority of the population in Kersa is out of work; hence the dependency ratio in most of the years is below 1 and ranged from 0.88 to 0.98. The young population dependency ration was the highest (0.88) as compared with the old dependency ratio (0.05). A reduction in neonatal, infant and under five mortalities was observed. For all deaths, verbal autopsies were done. Non-communicable diseases (NCDs) were the second leading cause of death among adults and the trend indicates that NCDs may surpass infectious diseases as a leading cause of death in the near future, while malnutrition is the leading cause of death among children under five years. For the past ten years, the Kersa HDSS has been supported with regard to advancement of research undertakings, health science education, generating evidence for improving planning and the delivery of health service. Currently the network has three centres at neighbourhood locations (Kersa, Harar and Haramaya) with a total population nearing 200,000. For this study, the cause of death due to CVDs was extracted from a ten-year adult mortality database available at Kersa HDSS and triangulated with their corresponding verbal autopsy records. Cases were merged in a two-year interval to attain sufficiency.

Results

Model choice

The different models defined in the previous section were fitted to mortality data from Kersa HDSS. To select the best model, we considered the deviance information criterion (DIC) among several other quantities for model choice and model calibration available in INLA. The DIC is the sum of the posterior mean of the deviance $\overline{D}$ (a measure of goodness of fit) and the number of effective parameters p_D (a measure of model complexity). The DIC is a well-known Bayesian model choice criterion to decide which model provides the best trade-off between model fit and complexity (Spiegelhalter *et al.*, 2002). Models with the smallest DIC value provide the best trade-off between model fit and complexity.

After fitting all the proposed models, models with non-parametric time trend performed better than the parametric ones with respect to the trade-off between model fit and complexity based on DIC. From Table 1 we could easily observe that the parametric models exhibit relatively low values of the effective number of parameters (p_D) with the highest values of posterior deviance ($\overline{D}$) with the largest DIC values. Overall, models with type II interactions together with a RW2 prior for the structured temporal random effect were the one showing lower DIC values. Furthermore, models without unstructured temporal components seemed to be better. Finally, Model 14 with a DIC value of 264.03 turned out to be the best model in terms of a trade-off between model fit and complexity (the smallest DIC value). This model included the spatial effect with a Leroux CAR prior, a structured temporal random effect with a RW2 prior and a type II interaction effect. The estimated log-relative risk $(\log\widehat{\rho_{it}})$, where $(\log\widehat{\rho_{it}}) = (\hat{\alpha}) + (\hat{\xi}) + (\hat{\gamma}) + (\hat{\delta})$ obtained

from the selected model can be split up into its separate components: a global risk estimate (); an estimate for the spatial risk due to location (ξ) that may be attributed to factors associated to a particular administrative region; an estimate for the temporal risk trend (γ) for all areas that may be attributed to changes in distribution of the disease, associated diagnostics, related policies affecting the region and an estimate for the area specific temporal risk trend (δ) that may reflect particular effects of each administrative region for the observed difference in each administrative region (kebele).

Figure 1 shows the spatial patterns of mortality, specifically the mortality risk ($exp\,(\xi)$) due to CVD at each administrative unit of Kersa HDSS and Figure 2 displays the posterior probability that the spatial risk ($exp\,(\xi)$) is greater than 1. Most literatures set posterior probability of spatial risk above 0.8 as a cutting point towards high risk administrative regions and more detail about the thresholds and cut-off probabilities can be seen from Richardson *et al.* and Ugarte *et al.* (Richardson *et al.*, 2004; Ugarte *et al.*, 2009). From both Figures it can be observed that, administrative regions in the eastern part and far south-west areas are those with high risk.

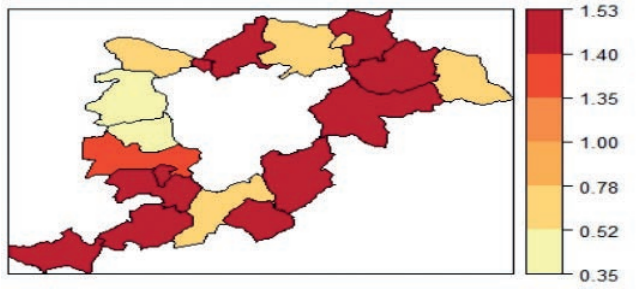

Figure 1. A map showing the spatial pattern of mortality ($exp\,(\xi)$) due to cardiovascular diseases in Kersa health and demographic surveillance system.

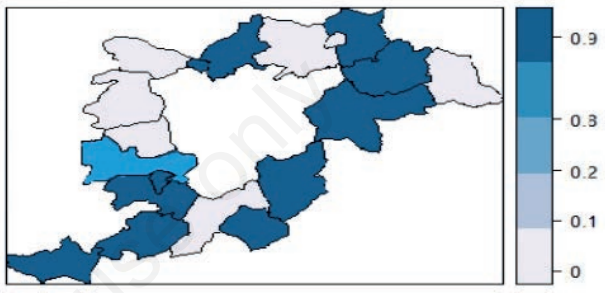

Figure 2. A map showing the posterior probability that the spatial risk($exp\,(\xi)$) due to cardiovascular diseases in Kersa health and demographic surveillance system is greater than 1.

Table 1. Summary of the posterior mean of the deviance ($\overline{D}$), the number of effective parameters pD and the deviance information criterion (DIC) as a measure of trade-off between model fit and complexity for all the models.

PARAMETRIC MODELS Model		$\overline{D}$	p_D	DIC
Model 1		296.93	20.37	317.30
Model 2		296.59	19.61	316.20
Model 3		297.44	20.52	317.96
NON- PARAMETRIC MODELS **Model**	**SP interaction**	$\overline{D}$	p_D	**DIC**
Model 1	Additive (RW1)	249.74	20.60	270.34
Model 2	Additive (RW2)	249.61	20.53	270.14
Model 3	Type I (RW1)	246.77	25.51	272.28
Model 4	Type I (RW2)	246.72	25.36	272.08
Model 5	Type II (RW1)	250.41	19.10	269.51
Model 6	Type II (RW2)	247.07	16.93	264.96
Model 7	Type III (RW1)	246.05	25.99	272.04
Model 8	Type III (RW2)	245.95	25.92	271.87
Model 9	Type IV (RW1)	248.26	24.00	272.26
Model 10	Type IV (RW2)	248.17	23.86	272.03
Model 11	Additive (RW1)	249.77	20.41	270.18
Model 12	Additive (RW2)	249.45	20.26	269.71
Model 13	Type II (RW1)	251.56	19.30	270.86
Model 14	Type II (RW2)	247.04	16.63	264.03
Model 15	Type IV (RW1)	248.38	23.95	272.33
Model 16	Type IV (RW2)	248.06	23.89	271.95

In Figure 3, the top map panel displayed the spatio-temporal trend of mortality due to CVD risks for each administrative region with comparison to the whole region during the study period. The bottom panel in Figure 3 displays the posterior probabilities that the relative risks are greater than 1. The risk scale was originally constructed in the logarithmic scale expressing the magnitudes of excess risk and default risk with respect to the whole region and was then back-transformed to facilitate the display maps and interpretation. In the map, the value 1.47 means 47% excess risk with comparison to the whole region during the study period. Overall, it can be seen from both maps that the mortality risk due to CVD is on the rise both in space and time, specifically the years 2011-2012 were identified as having the highest risk, while administrative regions in the eastern and far south-western part of the region exhibited a consistent high risk.

Simulation study

A simulation study was conducted by generating spatio-temporal data based on setting of mortality data of Kersa HDSS. We generated 300 data sets where the random effects for each of them were generated from the following multivariate normal distributions:

$$\xi \sim N(0, \widehat{\sigma_S^2}(\widehat{\lambda_S} R_S + (1-\widehat{\lambda_S})\mathrm{I_n})^{-1}), \gamma \sim N(0, \widehat{\sigma_\gamma^2} R_t^-) \text{ and } \delta \sim N(0, \widehat{\sigma_\delta^2} R_\delta)$$

where each of the estimates are the average of the models in the study to generate the log risk, $\log\rho_{it}$. For each case, counts were generated from a Poisson distribution with the mean $\mu_{it} = e_{it}$

Table 2. Summary of the DIC for the simulated data.

LINEAR/PARAMETRIC MODELS	
Model	**DIC**
Model 1	290.56
Model 2	290.32
Model 3	288.03
NON-PARAMETRIC MODELS	
Model	**DIC**
Model 1	248.55
Model 2	247.99
Model 3	251.11
Model 4	248.24
Model 5	243.07
Model 6	239.57
Model 7	247.98
Model 8	247.55
Model 9	248.54
Model 10	248.01
Model 11	245.11
Model 12	244.51
Model 13	246.12
Model 14	240.44
Model 15	249.99
Model 16	247.02

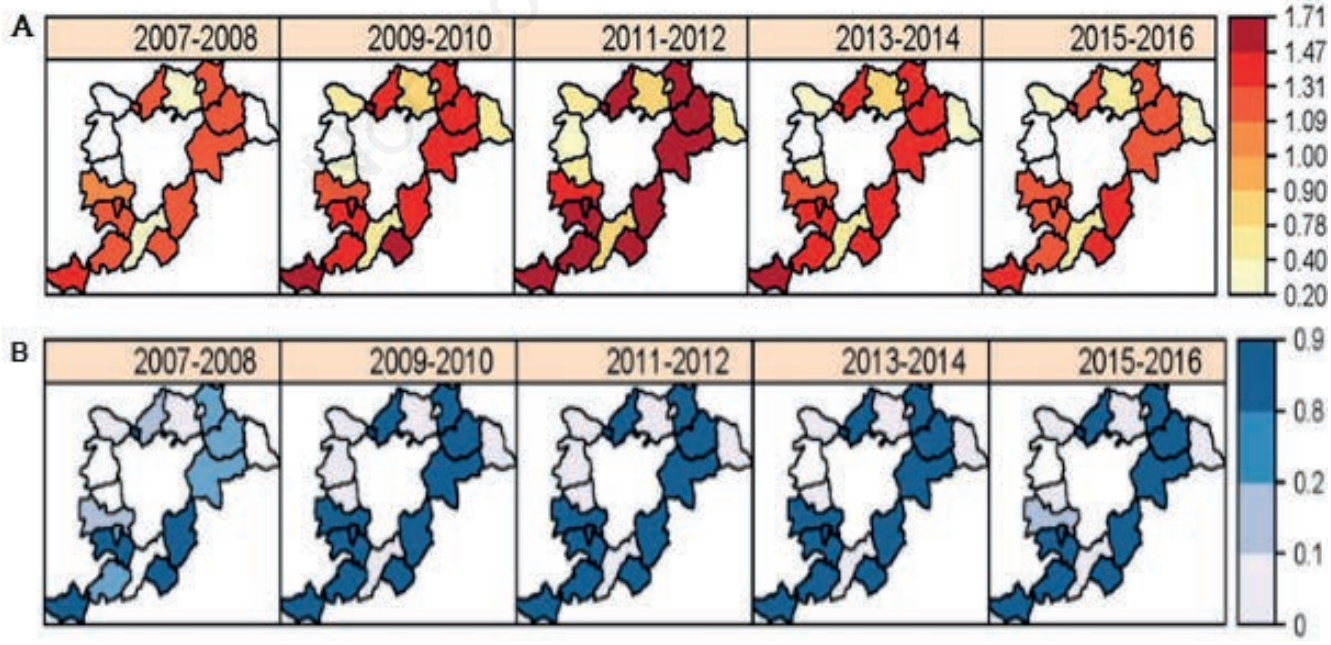

Figure 3. The distribution of relative mortality risk due to cardiovascular diseases in Kersa health and demographic surveillance system (A) and posterior probability distribution mortality for each administrator region during the time 2007-2016 (B).

exp(logρ_{it}). This was repeated three times by multiplying the expected count by 2^n for $n = -1,1,2$ to account for the effect of the population. Finally, the same hyperprior used in the analysis were used to generate these data. After fitting the data generated by using the models proposed in this study, a model with non-parametric time trend performed much better than the models with linear time trends. As shown in Table 2, the best model selected when fitting the model with real data still performed well in the simulation but now it was the second best model with very little difference from the first best model from the simulated data. Overall, with small discrepancies the proposed models performed equally well with the simulated data.

Conclusions

This work focuses on modifying and extending the existing structural models in spatio-temporal data analysis for disease mapping to present a flexible model to analyze aerial data for mortality clustering. Different models with parametric and non-parametric components were proposed and fitted using a fully Bayesian framework. Model fitting was carried out using INLAs and DIC was used to choose the best model among the proposed candidates. All the models were applied to mortality data collected from Kersa HDSS during the period 2007-2016. Overall nonparametric models performed much better than parametric models. A model with non-parametric trend, without unstructured time effect, with type II interaction and second order random walk stood as the best among all the proposed models. The simulation study confirmed the same with little discrepancy of results among the non-parametric models. Importantly, our analysis shows that the trend of mortality due to CVD is increasing over time and it is obvious that the administrative regions in the eastern and south-western regions need considerable attention. The results from this study highlights areas requiring more targeted health interventions, which in turn should lead to more detailed inquiries regarding the mortality due to CVD in space and time as well as the associated risk factors that account for these patterns.

References

Anderson C, Lee D, Dean N, 2014. Identifying clusters in Bayesian disease mapping. Biostatistics 15:457-69.

Assefa N, Oljira L, Baraki N, Demena M, Zelalem D, Ashenafi W, Dedefo M, 2016. HDSS profile: the kersa health and demographic surveillance system. Int J Epidemiol 45:94-101.

Benzler J, Sauerborn R, 1998. Rapid risk household screening by neonatal arm circumference: results from a cohort study in rural Burkina Faso. Trop Med Int Health 3:962-74.

Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M, 1995. Bayesian analysis of space-time variation in disease risk. Statist Med 14:2433-43.

Besag J, York J, Mollié A, 1991. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Statist Mathemat 43:1-20.

Bilancia M, Demarinis G, 2014. Bayesian scanning of spatial disease rates with integrated nested Laplace approximation (INLA). Statist Methods Appl 23:71-94.

Bloom D, Cafiero E, Jané-Llopis E, Abrahams-Gessel S, Bloom L, Fathima S, Feigl A, Gaziano T, Hamandi A, Mowafi M, 2012. The global economic burden of noncommunicable diseases. Program on the Global Demography of Aging. Geneva: World Economic Forum.

Boyle P, Walker A, Alexander F, 1996. Methods for investigating localized clustering of disease. Historical aspects of leukaemia clusters. IARC Sci Publ 135:1-20.

Central Statistical Agency (CSA), 2010. Population and housing census country 2007. Ethiopia: Central Statistical Agency (CSA).

Dedefo M, Oljira L, Assefa N, 2016a. Small area clustering of under-five children's mortality and associated factors using geo-additive Bayesian discrete-time survival model in Kersa HDSS, Ethiopia. Spatial Spatio-Temporal Epidemiol 16:43-9.

Dedefo M, Zelalem D, Eskinder B, Assefa N, Ashenafi W, Baraki N, Tesfatsion MD, Oljira L, Haile A, 2016b. Causes of death among children aged 5 to 14 years old from 2008 to 2013 in Kersa Health and Demographic Surveillance System (Kersa HDSS), Ethiopia. Plos One 11:e0151929.

Elliot P, Wakefield JC, Best NG, Briggs D, 2000. Spatial epidemiology: methods and applications. Oxford, UK: Oxford University Press.

Everitt B, Dunn G, 2011. Applied multivariate analysis. London, UK: Arnold.

Gangnon RE, Clayton MK, 1998. Detecting and modeling spatial disease clustering: A Bayesian approach. Technical Report 131. Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison WI. Available from: https://www.biostat.wisc.edu/content/detecting-and-modeling-spatial-disease-clustering-bayesian-approach

Gangnon RE, Clayton MK, 2000. Bayesian detection and modeling of spatial disease clustering. Biometrics 56:922-35.

Gilks WR, Roberts GO, 1996. Strategies for improving MCMC. In: Gilks WR, Richardson S, Spiegelhalter DJ, eds. Markov Chain Monte Carlo in Practice. London: Chapman and Hall. pp 89-114.

Gómez-Rubio V, Ferrándiz-Ferragud J, López-Quílez A, 2005. Detecting clusters of disease with R. J Geogr Syst 7:189-206.

Gómez-Rubio V, Moraga P, Molitor J, 2009. Fast Bayesian classification for disease mapping and the detection of disease clusters. Technical report. Universidad de Castilla-La Mancha, Spain. Available from: https://www.researchgate.net/publication/313369582_Fast_Bayesian_classification_for_disease_mapping_and_the_detection_of_disease_clusters

Gómez-Rubio V, Moraga P, Molitor J, Rowlingson B, 2006. DClusterm: model-based detection of disease clusters. J Stat Soft 55:1-26.

Hammer GP, Some F, Muller O, Kynast-Wolf G, Kouyate B, Becher H, 2006. Pattern of cause-specific childhood mortality in a malaria endemic area of Burkina Faso. Malar J 5:47.

Hjalmars U, Kulldorff M, Gustafsson G, 1994. Risk of acute childhood leukaemia in Sweden after the Chernobyl reactor accident. Swedish Child Leukaemia Group. BMJ 30:154-7.

Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N, 1996. Childhood leukaemia in Sweden: using GIS and a spatial scan-statistic for cluster detection. Statist Med 15:707-15.

Hjalmars U, Kulldorff M, Wahlqvist Y, Lannering B, 1999. Increased incidence rates but no space–time clustering of childhood astrocytoma in Sweden, 1973-1992. Cancer 85:2077-90.

Knorr-Held L, Rasser G, 2000. Bayesian detection of clusters and

discontinuities in disease maps. Biometrics 56:13-21.
Kulldorff M, 1998. Statistical methods for spatial epidemiology: tests for randomness. In: Löytönen M, Gatrell A, eds. GIS and Health. Abingdon, UK: Taylor & Francis. pp 49-62.
Kulldorff M, Nagarwalla N, 1995. Spatial disease clusters: detection and inference. Stat Med 14:799-810.
Laslett LJ, Alagona P Jr, Clark BA 3rd, Drozda JP Jr, Saldivar F, Wilson SR, Poe C, Hart M, 2012. The worldwide environment of cardiovascular disease: prevalence, diagnosis, therapy, and policy issues: a report from the American College of Cardiology. J Am Coll Cardiol 60:S1-49.
Leeder S, Raymond S, Greenberg H, Liu H, Esson K, 2012. A race against time: the challenge of cardiovascular disease in developing economies. New York, NY: Trustees of Columbia University.
Leroux BG, Lei X, Breslow N, 2000. Estimation of disease rates in small areas: A new mixed model for spatial dependence. In: Halloran ME, Berry D, eds. Statistical models in epidemiology, the environment, and clinical trials. The IMA Volumes in Mathematics and its Applications, vol 116. New York, NY: Springer. pp 116:179.
Mathers CD, Loncar D, 2006. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 3:p.e442.
McCullagh P, Nelder JA, 1989. Generalized linear models. Monograph on statistics and applied probability 37. London, UK: Chapman & Hall.
Pascutto C, Wakefield J, Best N, Richardson S, Bernardinelli L, Staines A, Elliott P, 2000. Statistical issues in the analysis of disease mapping data. Statist Med 19:2493-519.
Popkin BM, Adair LS, Ng SW, 2012. Global nutrition transition and the pandemic of obesity in developing countries. Nutr Rev 70:3-21.
Rezaeian M, Dunn G, St Leger S, Appleby L, 2007. Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. J Epidemiol Community Health 61:98-102.
Richardson S, Thomson A, Best N, Elliott P, 2004. Interpreting posterior relative risk estimates in disease-mapping studies. Environ Health Perspect 112:1016-25.
Rue H, Held L, 2005. Gaussian Markov random fields: theory and applications. London, UK: Chapman & Hall/CRC.
Rue H, Martino S, Chopin N, 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J Royal Statist Soc Ser B-Statist Methodol 71:319-92.
Rue H, Riebler A, Sorbye SH, Illian JB, Simpson DP, Lindgren FK, 2017. Bayesian Computing with INLA: a review. Annu Rev Statist Appl 44:395-421.
Sankoh OA, Ye Y, Sauerborn R, Muller O, Becher H, 2001. Clustering of childhood mortality in rural Burkina Faso. Int J Epidemiol 30:485-92.
Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A, 2002. Bayesian measures of model complexity and fit. J Royal Statist Soc Ser B-Statist Methodol 64:583-616.
Torabi M, Rosychuk RJ, 2011. An examination of five spatial disease clustering methodologies for the identification of childhood cancer clusters in Alberta, Canada. Spat Spatiotemporal Epidemiol 2:321-30.
Ugarte MD, Adin A, Goicoa T, Militino AF, 2014. On fitting spatio-temporal disease mapping models using approximate Bayesian inference. Statist Methods Med Res 23:507-30.
Ugarte MD, Goicoa T, Militino AF, 2009. Empirical Bayes and fully Bayes procedures to detect high-risk areas in disease mapping. Computation Statist & Data Anal 53:2938-49.
World Health Organization (WHO), 2018. Fact sheet on CVD. Available from: https://www.who.int/cardiovascular_diseases/about_cvd/en/ Accessed: 14 October 2018.