Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Bayesian disease mapping: Past, present, and future
Ying C. MacNab
School of Population and Public Health, University of British Columbia, Vancouver, Canada

A R T I C L E  I N F O
Article history:
Received 2 November 2021
Received in revised form 6 January 2022
Accepted 7 January 2022
Available online 19 January 2022

Keywords:
Bayesian hierarchical models
(Adaptive) Conditional autoregressive models
Empirical Bayes
Linear coregionalization
(Multidimensional) Gaussian Markov random fields
Bayesian disease mapping

A B S T R A C T
On the occasion of the Spatial Statistics’ 10th Anniversary, I reflect on the past and present of Bayesian disease mapping and look into its future. I focus on some key developments of models, and on recent evolution of multivariate and adaptive Gaussian Markov random fields and their impact and importance in disease mapping. I reflect on Bayesian disease mapping as a subject of spatial statistics that has advanced to date, and continues to grow, in scope and complexity alongside increasing needs of analytic tools for contemporary health science research, such as spatial epidemiology, population and public health, and medicine. I illustrate (potential) utility and impact of some of the disease mapping models and methods for analysing and monitoring communicable disease such as the COVID-19 infection risks during an ongoing pandemic.

1. Introduction
On the occasion of the Spatial Statistics’ 10th Anniversary, I reflect on the past and present of Bayesian disease mapping, now part of an established branch of spatial statistics for (areal) lattice data.

Disease mapping has many connotations. For some, typically in the contemporary contexts of geography of disease and of population and public health, it concerns with small area mapping of disease incidence and mortality rates or relative risks, commonly for a population under study and as a part of research into geographical distribution of disease. For Haviland (1871, 1888), for example, mapping observed disease and cancer incidence and mortality rates for the counties of England served to display geographic distribution of a disease and to illustrate possible associations...
between disease and the environment; see Arnold-Forster (2020, 2021) for recent research into Haviland’s pioneering works on disease and cancer mapping.

Maps of rare diseases (such as cancer) based on crude rates or relative risks typically lack stability in the illustrated “geographic distribution of disease”. The early efforts made to statistically ascertain localities (such as counties) of elevated or lowered disease incidence and mortality rates and risks gave rise to a new topic in spatial statistics for lattice data, known as disease mapping, a topic of my main focus here. Crude rates of a rare disease calculated for (small) geographic areas such as counties and health administrative regions are typically subject to high chance variations. Earlier statistical research of disease mapping focused on the development of empirical Bayes (EB) models and related estimation procedures for stabilizing (i.e. smoothing) maps of disease rates or relative risks (Tsutakawa et al., 1985; Clayton and Kaldor, 1987; Tsutakawa, 1988; Manton et al., 1989; Cressie, 1993, to name a few). What was notably absent during the earlier years of EB disease mapping is the statistical inference of relative risks, inference in the sense that elevated or lowered disease risks are statistically ascertained.

The seminal works of Besag et al. (1991), Clayton and Bernardinelli (1992), Bernardinelli and Montomoli (1992), and Clayton et al. (1993) influenced and shaped a paradigm shift from the frequentist to Bayesian approach to disease mapping. Known as Bayesian disease mapping, the paradigm has Bayesian hierarchical models at its core and Markov chain Monte Carlo (MCMC) simulation methods provide the computational tools for Bayesian estimation, learning, and inference of all unknown quantities and parameters. Since the 1980s, the research agenda has broadened from providing tools for producing smoothed rate or risk maps to establishing a methodological paradigm with progressive, and systemically developed ideas, models, and methods for studying spatial and spatiotemporal variations, patterns, and clusters of disease and health outcome risks in populations of geographical regions, and for identification of (risk and protective) factors that might explain the variations and patterns and clusters. The body of Bayesian disease mapping models and related methods for estimation, learning, and inference has grown in remarkable proportion from the early development of Bayesian hierarchical models for univariate disease mapping to multidimensional disease mapping based on data of space, time, multivariate, and multi-array (e.g. data of space, time, multiple diseases, age group and/or gender stratified).

The richness of the disease mapping models and their wide range of health science applications are well documented in books (Lawson et al., 1999; Banerjee et al., 2004; Waller and Gotway, 2004; Cressie and Wikle, 2015; Lawson, 2018; Martinez-Beneito and Botella-Rocamora, 2019) and review articles (Richardson et al., 2004; Best et al., 2005; Wakefield, 2007; Ugarte et al., 2009; Lee, 2011; MacNab, 2011, 2018), in addition to a large quantity of research papers. Here, I review key developments of Bayesian disease mapping, with a focus on recent evolution of multivariate and adaptive Gaussian Markov random fields and their impact and importance in Bayesian disease mapping. I illustrate (potential) utility and impact of some of the disease mapping models and methods for analysing communicable disease such as the COVID-19 infection risks in the context of pandemic intervention and monitoring. I reflect on Bayesian disease mapping as a subject of spatial statistics that has advanced to date, and continues to grow, in scope and complexity alongside increasing needs of analytic tools for contemporary health science research, such as spatial epidemiology, population and public health, and medicine.

2. Empirical Bayes disease mapping since the 1980s

While mapping disease and cancer rates or risks began century into the past, disease mapping gained attentions of the statistics community during the 1980s when cancer registry made population-level data more readily available for small geographical (e.g. administrative) areas of a region or country. A common goal of the researchers was to develop model-based procedures for stabilized cancer risk maps, from which environmental determinants of different types of cancer and their possible etiologic factors may be hypothesized for further investigations (Tsutakawa et al., 1985; Clayton and Kaldor, 1987; Tsutakawa, 1988; Manton et al., 1989, to name a few).

The 1980s also marked the rise of popularity of EB approaches to formulate and estimate models of multi-parameters, models in which the number and dimension of unknown parameters may equal or exceed those of the data (see Carlin and Louis, 2000a,b and references therein). Disease mapping models of the past and present are models of this nature.
2.1. An example of disease mapping model: Poisson model for count data

To facilitate further discussions, let us consider a typical disease mapping setting where we study areal-level cancer mortality risks through modelling of observed area-specific death counts, denoted \( O = (O_1, O_2, \ldots, O_n) \), where \( n \) is the number of geographic areas, such as wards, counties, local health administrative areas of a region or a country under study. For brevity, and without essential loss of generality, we will consider the simplest situation where the death counts are compared to a set of expected counts, denoted \( E = (E_1, E_2, \ldots, E_n) \), which is often derived for a set of area- and age-specific at risk population and the corresponding age-specific reference rates, known as age-standardization (Clayton and Kaldor, 1987; Manton et al., 1989). The resulting relative risks, often defined as \( RR_i = O_i/E_i \), are the standardized mortality ratios (SMR). This is a well known method used in studies of disease epidemiology and in cancer research (Breslow and Day, 1980). In disease mapping, this approach is often taken when areal-level age-specific data are not available or extremely small.

Working with rare events such as cancer mortality, the observed death counts may be viewed as Poisson events and modelled as follows:

\[
O_i \sim \text{Poisson}(\phi_i), \quad \phi_i = E_i \varphi_i, \quad \forall i,
\]

where \( \varphi_i, \forall i \), is the Poisson intensity, and \( \varphi_i \) is the unknown relative risk.

In frequentist disease mapping, the unknown relative risks in (1) are fixed parameters to be estimated from the data. One often assumes that the mortality counts are randomly varying (i.e. statistically independent) Poisson events over geographic areas. Under this assumption, the maximum likelihood estimates (MLEs) of the relative risks, denoted \( \hat{\phi}_i, \forall i \), are the previously mentioned SMRs:

\[
\hat{\phi}_i = O_i/E_i, \quad \text{with associated standard errors } se(\hat{\phi}_i) = \sqrt{\hat{\phi}_i/E_i}, \quad \forall i.
\]

It is well known that the MLEs (2) may lack stability and efficiency when small counts are modelled. The independence assumption is often invalid due to possible spatial correlation, e.g., the tendency that areas of close proximity (say, neighbouring areas) have similar mortality risks.

In addition, Poisson distribution has the unusual characteristic that the intensity \( \phi_i \) quantifies both the mean and the variance of the distribution. In real-life applications, this variance-equals-mean assumption may be invalid. For example, disease mapping data may exhibit extra-Poisson variation, i.e. \( \text{Var}(O_i) > E(O_i) \), also known as Poisson over-dispersion (Breslow, 1984; Clayton and Kaldor, 1987; Besag et al., 1991; Breslow and Clayton, 1993).

2.2. Empirical Bayes disease mapping

Model (1) is a model of multi-parameters, where the dimension of parameters equals the dimension of the data. Taking an EB approach to this problem, one would place a prior on \( \varphi \), denoted \( f(\varphi | \theta) \), where \( \theta \) is a (vector of) hyper-parameter(s) to be estimated from data. This approach can facilitate borrowing information and risk smoothing and, with certain priors, account for Poisson over-dispersion. Many EB methods for disease mapping were proposed since the 1980s (Tsutakawa et al., 1985; Clayton and Kaldor, 1987; Tsutakawa, 1988; Manton et al., 1989; Cressie, 1993; Leroux et al., 1999, to name a few). Here, the basic ideas are outlined in the following two illustrative examples, both are taken from, and presented in detail in, Clayton and Kaldor (1987).

Since relative risks are non-negative, one prior option for \( \varphi \) is the gamma distribution \( \varphi_i \sim \text{gamma}(u, v) \), \( \forall i \), where \( u \) and \( v \) are the respective scale and shape parameters. This is a non-spatial prior, where the Poisson–gamma mixing leads to negative binomial distribution for \( O_i, \forall i \), with expectation and variance functions:

\[
E(O_i) = E_i v/u, \quad \text{Var}(O_i) = E_i v/u + E_i^2 v/u^2, \quad \forall i, u > 0, v > 0.
\]

Negative binomial distribution is well known for its use to model Poisson over-dispersion (McCullagh and Nelder, 1983), as illustrated in expression (3) where the variance of \( O_i \) is greater than its mean. Once the hyper-parameters \( u \) and \( v \) are estimated from the data, say, the MLEs of the
negative binomial likelihood, denoted \( \hat{u} \) and \( \hat{v} \), the EB estimate of \( \psi_i \) is readily derived from its posterior expectation, \( E(\psi_i|O_i) = \frac{O_i + \hat{v}}{E_i + \hat{u}} \). \( \forall \ i \): Replacing the \( u \) and \( v \) in \( E(\psi_i|O_i) \) by \( \hat{u} \) and \( \hat{v} \) leads to the EB estimate

\[
\hat{\psi}_i = \frac{O_i + \hat{v}}{E_i + \hat{u}}
\]  

(4)

In addition to modelling Poisson over-dispersion, another feature of this method is that the EB estimate \( \hat{\psi}_i \) is a compromise between the SMR \( O_i/E_i \) and the estimated mean of \( \psi_i \), \( \hat{v} / \hat{u} \). This method leads to “global smoothing” of relative risks (see Clayton and Kaldor, 1987 for additional details).

However, global smoothing may be suboptimal when data are spatially distributed and correlated. An alternative approach is to formulate models that allow for spatial correlation and enable spatial smoothing. In Clayton and Kaldor (1987), a conditionally formulated Gaussian Markov random field (GMRF, Besag, 1974), named CAR for conditional autoregression, was proposed as a spatial prior for the log relative risks ensemble \( \psi = (\psi_1, \ldots, \psi_n) \):

\[
\psi \sim \text{GMRF}(\mu, \Omega), \Omega = \sigma^{-2}(I_n - \lambda W),
\]  

(5)

where \( \Omega \) is the precision matrix, \( \psi_i = \log(\psi_i) \), \( \forall i \), \( \sigma \) is a scale parameter, \( \lambda \) is a spatial dependence parameter, and \( W \) is the well known adjacency matrix for a given lattice (of map) and its neighbourhood system (e.g. two areas are neighbours if they share common border(s)).

This CAR model postulates symmetric spatial dependence, accounts for spatial correlation, facilitates spatial smoothing, and enables borrowing strength for risks prediction and inference. The key features of this conditionally formulated GMRF can be understood and explained via the conditional mean and precision functions of its full conditionals:

\[
E(\psi_i|\psi_{-i}) = \lambda \sum_{j \neq i} \psi_j, \quad \text{Prec}(\psi_i|\psi_{-i}) = \tau
\]  

(6)

where \( \tau = \sigma^{-2}, \psi_{-i} = (\psi_1, \ldots, \psi_{i-1}, \psi_{i+1}, \ldots, \psi_n) \), \( j \sim i \) stands for the area \( j \) and \( i \) are neighbours. The area-specific conditional risk expectation of \( \psi_i \) given \( \psi_{-i} \), \( \forall i \), is proportional to the total risks of its neighbouring areas, regulated by a positive spatial dependence parameter \( \lambda \). The conditional risk prediction variance is regulated by the scale parameter \( \sigma \).

Unlike the Poisson–gamma mixing, the Poisson–log-normal mixing does not lead to an analytically known distribution for \( O \). In Clayton and Kaldor (1987), an EM algorithm (Dempster et al., 1977) was used for a two-stage estimation of \( \psi \) (and \( \phi \)) and its hyper-parameters.

The EB estimation is computationally straightforward for the first example and more complex and evolved for the second one. In an influential paper, Breslow and Clayton (1993) popularized several EB methods for inference in generalized linear mixed effects (GLMM) models, a model class that contains the majority of the disease mapping models developed to date.

The main limitation of the EB estimation of \( \phi \) is that while the point estimates are nearly unbiased, the EB standard errors typically understate the associated estimation uncertainties. This can be readily observed from the “plug-in” EB estimation of the posterior (log) relative risks in the first example, where the uncertainties associated with the estimation of the hyper-parameters may not be adequately accounted for. This is the well known “naive” EB estimation issue for nearly all EB approaches to estimation of \( \psi \) (Carlin and Gelfand, 1990, 1991; Carlin and Louis, 2000a; MacNab et al., 2004; Ainsworth and Dean, 2006; MacNab and Lin, 2009).

Into the 1990s and 2000s, some progresses were made to improve EB posterior estimation of relative risks, mainly to provide better standard error estimates by bootstrap methods (Carlin and Gelfand, 1990; MacNab et al., 2004) or via analytic adjustment (Ainsworth and Dean, 2006; MacNab and Lin, 2009). With adequately calibrated standard errors that quantify risk uncertainties, EB methods for posterior inference of relative risks may be used to statistically ascertain elevated or lowered disease risks.

3. Bayesian disease mapping from the 1990s: A paradigm shift

Over the broad landscape of statistical science, the 1990s saw the rise of Bayesian statistics and a new trend of Bayesian hierarchical approaches to formulating complex models, with hierarchical
Bayesian methods for posterior estimation and inference of all unknown quantities and parameters, typically implemented by Markov chain Monte Carlo (MCMC) simulations (Gelfand and Smith, 1990; Smith and Gelfand, 1992; Gelman et al., 1995; Gilks et al., 1998, to name a few). Amid the MCMC movement into the 1990s and the new 21st century, a shift from EB to fully Bayesian (FB) approach to disease mapping was taking place (Bernardinelli and Montomoli, 1992; Gilks et al., 1998; Lawson et al., 1999; Gelman et al., 1995; Congdon, 2001).

The move from EB to FB disease mapping was an analytically natural transition. The two are analytically comparable in the sense that the latter produces Bayesian posterior estimates for the prior parameters, but doing so by placing non-informative or weakly informative priors on the prior parameters, an approach facilitates “data-driven” posterior estimation and inference of the prior parameters (Carlin and Louis, 2000a; MacNab et al., 2004).

Analytic and conceptual transitions take place when the fully Bayesian approach constructs models hierarchically (Cressie and Wikle, 2015) and within a Bayesian hierarchical inferential framework. Within this framework, the prior distribution, such as the previously mentioned gamma or CAR prior, can be postulated as an underlying process model for the underlying risks \( \psi \) that give rise to the observed disease incidence and mortality counts. Statistical inference utilizes the Bayes’s rule to compute (approximate) posterior distributions for all unknown quantities and parameters. Conceptually, one can explore meaningful and intuitively appealing prior and hyper-prior options within a Bayesian hierarchical inferential framework for model formulation and Bayesian posterior estimation, learning, and inference of relative risks and parameters of interests and importance. This approach provides a cohesive and principled way to quantify risk uncertainties and uncertainties of all unknown parameters.

With the arrival of freely available BUGS/WinBUGS software (Spiegelhalter et al., 2003) for Bayesian hierarchical analysis and later the GeoBUGS add-on module (Thomas et al., 2004) to WinBUGS for Bayesian analysis of spatial data, in addition to various publicly available R packages, e.g. the R-INLA (Rue et al., 2017) and the CARBayes (Lee, 2013), for (approximate) Bayesian analysis of spatial areal data, fully Bayesian approach to disease mapping model formulation, estimation, and inference became mainstreamed and stayed so to this day.

4. Multidimensional disease mapping

Multidimensional disease mapping is a broad topic that concerns with spatiotemporal disease mapping (Held and Besag, 1998; Held, 2000; MacNab and Gustafson, 2007; Ugarte et al., 2017), multivariate disease mapping (Kim et al., 2001; Jin et al., 2007; Greco and Trivisano, 2009; Martinez-Beneito, 2013; MacNab, 2018), and multi-array disease mapping, say, when space, time, age (and gender, etc.) stratifications are all considered in a disease mapping model (Held and Besag, 1998; MacNab, 2003a; Martinez-Beneito et al., 2017; Martinez-Beneito and Botella-Rocamora, 2019). Literature on this topic is quite rich (see Lawson, 2018; Martinez-Beneito and Botella-Rocamora, 2019, and references therein). Here, I briefly reflect on common ideas and models that characterize cross-dimensional dependencies and interactions.

Without essential loss of generality, let us begin with adding another dimension to the Poisson data model (1), in which one has a matrix of log relative risks \( \psi = (\psi_{ij}) \), where \( j = 1, 2, \ldots, J \) for mapping of \( J \) diseases.

Spatiotemporal disease mapping models presented in the literature focused mainly on characterizing the presence of space and time risk interactions, in addition to model components of risks in space and time, respectively (Waller et al., 1997; Held and Besag, 1998; Held, 2000). These models are often motivated to facilitate borrowing information and risk smoothing within and between space and time. They are also natural extensions of the spatial risk models.

Some of the models are richly parameterized to hypothesize space and time risk interactions, in addition to decompose the \( \psi \) into (parameterized) components of space and time. Typical examples are the models of Waller et al. (1997) and Held (2000), two highly cited and influential papers. For example, Held (2000) proposed a general model formulation

\[
\psi_{it} = \alpha_t + \varphi_i + \phi_i + \eta_i + \delta_{it}, \quad i = 1, 2, \ldots, n, \quad t = 1, 2, \ldots, T,
\]
and considered types of space–time interactions and associated prior options for \( \{ \delta_{it}, \forall i, t \} \). These spatiotemporal (interaction) models may be highly parameterized and, in this situation, additional identification constraints may be imposed (Held, 2000).

Alternatively, one can take a parsimonious way to parameterize spatiotemporal models. For example, the following model formulation was considered in the literature (MacNab and Dean, 2001):

\[
\psi_{it} = \alpha_t + \theta_i + \delta_{it}, \forall i, t,
\]

(8)
in which functional analysis may be used to hierarchically model space \( \times \) time interactions. For instance, illustrated in MacNab and Dean (2001), MacNab and Gustafson (2007), and MacNab (2007), space–time interactions \( \delta_{it} \) can be modelled via spatially varying splines (e.g. B-splines, P-splines, and smoothing splines), in addition to modelling the time components \( \alpha_t \) by a fixed spline for non-linear risk trend. In Ugarte et al. (2010), Lee and Durbán (2011), Ugarte et al. (2017), and Feng (2021), space–time interactions are modelled by tensor-based 2d or 3d P-splines or thin plate spline. Lawson (2018) and MacNab (2018) also considered parsimonious spatiotemporal model parameterizations. They illustrate modelling linear risk trends \( \alpha_t = \beta t \) and \( \delta_{it} = \delta_i t_j \), which can be viewed as a special (degenerated) case of the previously mentioned spline-based approaches for modelling non-linear risk trends.

Modelling spatially varying non-linear risk trends via spatially varying random walk (RW) or auto-regressive (AR) or CAR processes are also viable options (Held, 2000; Ugarte et al., 2017; MacNab, 2018; Sahu and Böhning, 2021); see Section 9 for an example. Spatiotemporal disease mapping models were mainly developed for posterior prediction of the unobservable relative risks, although a few have been developed for short-term forecasts of disease rates (Etxeberria et al., 2015; Sahu and Böhning, 2021).

Without sufficient difficulty, the spatiotemporal models (7) and (8) can be extended further for multi-array disease mapping. Adding age-effects is one example. Age-effects could be included in (7) or (8) as additional components (Held and Besag, 1998; Goicoa et al., 2016; Martinez-Beneito et al., 2017) or modelled by a spline for non-linear age-effects (MacNab, 2003a). Multivariate spatiotemporal disease mapping is another direction (Tzala and Best, 2008; Jack et al., 2019; Vicente et al., 2020, 2021).

Multivariate disease mapping has been an active area of research since the early 2000s. The main motivation of the efforts is to develop methods for jointly modelling diseases that share common risk factors, enabling more effective borrowing-information, and facilitating spatial and cross-spatial risk smoothing. Two main approaches to multivariate disease mapping are considered in the disease mapping literature. One is to model \( \psi_{ij} \) as a sum of two components:

\[
\psi_{ij} = \theta_i + \delta_{ij}, \forall i, j,
\]

(9)
where \( \theta_i \) is a component shared by all diseases. This is the shared component model (SCM) proposed by Held and Best (2001). SCMs typically hypothesize positive cross-variable local dependencies and correlations (MacNab, 2010). The other more flexible approach is to place a multivariate GMRF prior on \( \text{vec}(\psi) \) or \( \text{vec}(\psi^\top) \), which I discuss in Section 7.

5. Disease mapping with areal-level covariates: Ecological regression

In the context of disease mapping, including a regression part in a disease mapping model can serve several important and practical purposes. One important purpose of a spatial regression model is to quantify disease risk variations explained and unexplained by the included risk (and/or protective) factors (Clayton et al., 1993; MacNab, 2003b,c, 2004).

Known as ecological modelling and analysis, ecological regression can serve to identify areal-level characteristics (e.g. environmental exposure, socio-economic deprivation, contextual characters) that influence the disease risks. Most of the ecological models developed in the disease mapping literature were aimed for this purpose. They are commonly motivated and applied to research into the (social and/or physical) environmental determinants of health (Clayton et al., 1993;
Ecological regression faces several methodological challenges. One of them is the long-standing ecological bias, referring to the situation when the degree of association between risk exposure and disease at the areal-level is different from that of the individual level, due to within-area variability in exposures and confounders (Clayton et al., 1993; Wakefield and Salway, 2001; Wakefield, 2007; Lawson, 2018).

Another analytic challenge is spatial confounding, commonly discussed as collinearity between fixed effects and random effects in a spatial GLMM, and in the presence of unmeasured confounding. Literature of ecological studies offered two main approaches to the issue. One is to deconfound the fixed and random effects (Reich et al., 2006; Hughes and Haran, 2013; Ugarte et al., 2017; MacNab, 2018; Zimmerman and Hoef, 2021). Another is to formulate transformed GMRF or MGMRF (Prates et al., 2015; Hughes, 2015; Prates et al., 2021) so that the fixed and random effects can be modelled independently. Named TGMRF or TMGMRF, the basic idea of this approach is to model the relative risks $\psi$, and its spatiotemporal or multivariate extensions, via a copula-inspired model structure that allows fixed and random components to be estimated independently: The fixed effects are modelled on the univariate margins, whereas the random effects spatial dependencies are modelled via a GMRF or MGMRF spatial dependence matrix. One limitation of this approach is that the regression part in the model is designed to explain marginal variations, not spatial dependence. In traditional spatial regression GLMMs, the included covariates may (partially) explain risk variation or spatial risk dependence or both; see MacNab (2003b,c, 2004, 2009, 2010), and Section 9, for illustrative examples.

Of the deconfounding approach, earlier proposals are known as the restricted spatial regression (RSR) method (Reich et al., 2006; Hughes and Haran, 2013; Ugarte et al., 2017; Zimmerman and Hoef, 2021). The key idea of the RSR approach is to re-parameterize a $n$-dimensional spatial prior and project the original Gaussian spatial prior onto a Gaussian spatial prior of its sub-space orthogonal to the column space of the fixed effects model matrix (Zimmerman and Hoef, 2021). Very recently, Zimmerman and Hoef (2021) researched into the RSR approach, named it deconfounding spatial confounding, and put the method into question. In the context of spatial linear mixed model and its frequentist inference, they showed that the RSR approach can lead to underestimation of the uncertainties associated with the estimation of the fixed effects and is generally inferior to that of the original spatial regression model.

An alternative to the RSR method is to formulate spatially misaligned ecological model, misaligned in a sense that the spatial prior of the random effects and the included covariates are varying at different spatial scales and/or configurations (Lee and Sarran, 2015; Prates et al., 2019; Azevedo et al., 2021). This approach to deconfounding is also consistent with the Paciorek (2010) work on spatial confounding, which offered insight into the importance of scale for spatial-confounding bias and precision of spatial regression estimators. In Paciorek (2010), analytic and simulation results suggested evidence that bias can be reduced when fitting a spatial model in which the variation of the covariate is at a scale smaller than the scale of the unmeasured confounding.

Errors-in-covariate is another important issue in ecological regression, where areal-level covariates are commonly measured imprecisely, or because surrogates or proxies are only available and used. In the Bayesian disease mapping literature, the main approach is to formulate measurement models for the covariates and incorporate them into a hierarchically formulated Bayesian spatial regression model (Bernardinelli et al., 1998; Wakefield and Morris, 1999; MacNab, 2009, 2010; Lawson, 2018).

6. Conditionally formulated GMRFs

Conditionally formulated GMRFs have played important roles in disease mapping and ecological regression (Clayton and Kaldor, 1987; Besag et al., 1991). They are formulated to model spatial dependencies and spatially structured correlation and covariance functions (MacNab, 2011, 2016a,b, 2018). Formulating GMRF via full conditionals is conceptually appealing for articulating the idea of spatial smoothing, as illustrated earlier for CAR (5) expressed by its full conditionals (the Expression
And the CAR(10) conditional precision is proportional to \( w \) of neighbourhood risks, rather than the sum of the neighbourhood risks as defined in CAR(6).

Typically defined on an irregular lattice of areal map, the pCAR, with its conditionalexpectation \( E(\psi_i | \psi_{-i}) = \lambda \sum_j \psi_j / w_{ij+} \), conditional covariance \( \text{Cov}(\psi_i, \psi_j | \psi_{-i}) = \tau w_{ij+} \), \( \tau = \sigma^{-2} \), \( i \neq j \), is a slight modification to the CAR (6), where \( w_{ij+} \) is the (predicted) risk. Consequently, when \( \lambda \in (0, 1) \), the pCAR conditional expectation \( E(\psi_i | \psi_{-i}) \), \( \forall i \), is proportional to an average of neighbourhood risks, rather than the sum of the neighbourhood risks as is defined in CAR (6). And the pCAR (10) conditional precision is proportional to \( w_{ij+} \), which implies that the greater the number of neighbours, the higher the conditional precision. That is, borrowing information from a higher number of neighbours could lead to higher precision of posterior risk prediction. This could be a plausible assumption in the context of disease mapping, where an areal map under study typically constitutes an irregular lattice with varying sizes of neighbourhoods over the map.

Another feature of the pCAR, which is not often discussed in the literature, is that it postulates asymmetric conditional dependency (i.e. direct influence) of \( \psi_j \) on \( \psi_i \), quantified by \( \lambda / w_{ij+} \) versus \( \psi_i \) on \( \psi_j \), defined by \( \lambda / w_{ij+} \), provided \( w_{ij+} \neq w_{ji+}, i \sim j \), and \( 0 < \lambda < 1 \). That is, the direct influence of the area \( j \) on its neighbouring area \( i \) is inversely proportional to the neighbourhood size of the area \( i \). Put it another way, an area with a higher number of neighbours is less influenced by its neighbour who has a lower number of neighbours. This could also be an intuitively plausible assumption, consistent with the conditional precision function \( \text{Prec}(\psi_i | \psi_{-i}) \) in (10); one might expect that an area of higher precision of risk prediction should be less influenced by an area with a lower precision of (predicted) risk. The spatial dependence parameter \( \lambda \) in pCAR is often called a spatial smoothing parameter; it regulates smoothly varying relative risks over the map.

### Table 1

| Model          | \( E(\psi_i | \psi_{-i}) \) | \( \text{Var}(\psi_i | \psi_{-i}) \) | \( \Omega \) |
|----------------|-----------------------------|---------------------------------|--------|
| iCAR(\( \sigma^2 \)) (Besag et al., 1991) | \( \frac{\lambda \sum_j \psi_j}{w_{ij+}} \) | \( \frac{\sigma^2}{w_{ij+}} \) | \( \sigma^{-2}(D_w - W) \) |
| pCAR(\( \lambda, \sigma^2 \)) (Cressie, 1993) | \( \frac{\lambda \sum_j \psi_j}{w_{ij+}} \) | \( \frac{\sigma^2}{w_{ij+}} \) | \( \sigma^{-2}(D_w - \lambda W) \) |
| LCAR(\( \lambda, \sigma^2 \)) (Leroux et al., 1999) | \( \frac{\lambda \sum_j \psi_j}{1 - \lambda \sum w_{ij+}} \) | \( \frac{\sigma^2}{1 - \lambda \sum w_{ij+}} \) | \( \sigma^{-2}(\lambda(D_w - W) + (1 - \lambda)I_n) \) |
| BYM(\( \sigma_1, \sigma_0 \)) (Besag et al., 1991) | \( \psi = \psi^s + \psi^h, \psi^s \perp \psi^h \) | \( \sigma^2(D_w - W)^{-1} + \sigma_0^2 I_n \) |
| MBYM(\( \psi, \sigma \)) (MacNab, 2011) | \( \psi = \sqrt{\lambda} \psi^s + \sqrt{1 - \lambda} \psi^h, \psi^s \perp \psi^h \) | \( \sigma^2(\lambda(D_w - W)^{-1} + (1 - \lambda)I_n) \) |

Analytically, the GMRF full conditionals also facilitate coding the powerful Gibbs sampler as a computational tool for posterior estimation of the GMRFs via MCMC simulations; see Besag et al. (1991) for an accessible explanation of Gibbs sampler.

Several CARs have been proposed since the CAR (6). Three stood the time and are commonly considered prior options for disease mapping. They are the well known intrinsic CAR (iCAR, Besag et al., 1991), proper CAR (pCAR, Cressie, 1993; Sun et al., 1999), and Leroux et al. CAR (LCAR, Leroux et al., 1999); see Table 1 for their full conditionals and associated precision matrices.

The pCAR(\( \lambda, \sigma \)) and LCAR(\( \lambda, \sigma \)) are two-parameter and full rank GMRFs, where \( \lambda \) and \( \sigma \) are the respective spatial and non-spatial parameters; \( \lambda \) is a spatial dependence or weight or smoothing parameter and \( \sigma \) is a (non-spatial) Gaussian scale parameter (Cressie, 1993; Leroux et al., 1999).

Typically defined on an irregular lattice of areal map, the pCAR, with its conditional expectation and precision functions

\[
E(\psi_i | \psi_{-i}) = \lambda \sum_j \psi_j / w_{ij+}, \quad \text{Prec}(\psi_i | \psi_{-i}) = \tau w_{ij+}, \quad \tau = \sigma^{-2}, \quad \forall i,
\]

\(^{(10)}\)}
To facilitate discussion, I name the coefficients of the conditional autoregression in (10) the coefficient of influence functions, or simply influence functions, denoted

\[
\text{Influence}(j, i)_{\text{ICAR}} = \lambda / w_{ij+}, \forall j \sim i.
\]

The LCAR conditionals also allow asymmetric spatial dependencies, provided \(\text{Influence}(j, i) \neq \text{Influence}(i, j), i \sim j\), where

\[
\text{Influence}(j, i)_{\text{LCAR}} = \lambda / (1 - \lambda + \lambda w_{ij+}), \forall j \sim i.
\]

That is, the LCAR influence function is a non-linear but increasing function of \(\lambda\). The spatial dependence parameter \(\lambda\) has triple roles and is also a spatial weight or smoothing parameter; it weights a precision matrix of the iCAR(\(\sigma\)) and a precision matrix of \(n\) IIDN(\(\sigma\)) (normal) variates, a mixing of purely local smoothing and global smoothing (Congdon, 2008). The LCAR conditional variances are non-linear decreasing functions of \(\lambda\) (see Table 1); that is, \(\lambda\) also partially controls the risk prediction variance of \(\psi_i\) conditional on \(\psi_j, \forall j \sim i\) (MacNab, 2011). Noted in MacNab (2018), the LCAR parameterization is an “entangled” spatial and non-spatial parameterization, and, as a result, LCAR has limited options for multivariate generalization (see Section 7).

The iCAR(\(\sigma\)) is a GMRF with a singular precision matrix of rank \(n-1\). It is the limiting distribution of the pCAR(\(\lambda, \sigma\)) and LCAR \((\lambda, \sigma)\) (when \(\lambda \to 1\)). The iCAR(\(\sigma\)) typically considered as a spatial prior for modelling spatially structured risk variation or clustered heterogeneity. It is commonly used when the \(n\)-vector of log relative risks \(\psi\) is modelled as additive components of \(\psi \sim \text{iCAR}(\sigma^2_i)\), for modelling spatially structured heterogeneity or effects of omitted covariates that are spatially varying, and \(\psi \sim \text{IIDN}(\sigma^2_i)\), for modelling extra-Poisson variation or effects of omitted covariates that are randomly varying. This is the well known Besag, York, and Mollie (BYM) model (Besag et al., 1991), widely used in Bayesian disease mapping and ecological regression. This model is highly parameterized. The BYM is typically used as a prior in Bayesian disease mapping, where estimation and inference of the relative risks may be implemented with (weakly) informative priors for (functions of) the scale parameters \(\sigma^2_i\) and/or \(\sigma^2_i\) (Thomas et al., 2004; MacNab, 2011).

To improve computational efficiency and parameter identification, MacNab (2011) proposed a modification to the BYM, named MBYM. In the MBYM, the spatial dependence parameter \(\lambda\) has triple roles and is also a spatial weight or smoothing parameter; it weights a precision matrix of the iCAR(\(\sigma\)) and a precision matrix of \(n\) IIDN(\(\sigma\)) (normal) variates, a mixing of purely local smoothing and global smoothing (Congdon, 2008). The LCAR conditional variances are non-linear decreasing functions of \(\lambda\) (see Table 1); that is, \(\lambda\) also partially controls the risk prediction variance of \(\psi_i\) conditional on \(\psi_j, \forall j \sim i\) (MacNab, 2011). Noted in MacNab (2018), the LCAR parameterization is an “entangled” spatial and non-spatial parameterization, and, as a result, LCAR has limited options for multivariate generalization (see Section 7).

The LCAR has gained popularity due in part to the analytic result that when \(\lambda = 0\), the LCAR reduces to independent and identical Gaussian priors for the log relative risks, whereas the pCAR, when \(\lambda = 0\), reduces to \(n\) independent Gaussian priors with prior precisions proportional to \(w_{ij+}, \forall i\) (MacNab, 2003c, 2011; Lee, 2011).

However, compared to LCAR, pCAR has its own advantages. Its spatial and non-spatial parameters play separate and different roles, one regulates spatial dependency, the other controls non-spatial variance. With its rich options of multivariate generalization, the pCAR, and some of its adaptive and multivariate extensions, have theoretical and practical appeals for modelling and interpreting spatial dependencies (MacNab, 2018, 2020, 2021a,b, see Sections 7 and 8).

Of the five CAR models, none has shown to outperform the others in all disease mapping situations. This is consistent with the analytic results that the iCAR, pCAR, and LCAR have different spatial dependence and correction functions (see details in MacNab, 2011, 2014) and the five CAR models may play nuanced roles in Bayesian disease mapping applications. Bayesian sensitivity analysis under various prior and hyper-prior options for Bayesian estimation, learning, and inference is an important part of Bayesian disease mapping application (MacNab, 2011); and, with goodness-of-fit and complexity assessment, say, based on deviance information criterion (the devance, pD, DIC scores, Spiegelhalter et al. (2002), also see Section 9), it remains a viable and commonly taken approach to model evaluation, comparison, and selection.

7. Multivariate GMRFs for multidimensional disease mapping

In the broad context of multidimensional disease mapping discussed earlier, multivariate GMRFs play important roles in modelling multidimensional spatial dependencies and cross-covariances, enabling borrow-information over multi-dimensions, and facilitating spatial and cross-spatial smoothing. The five commonly used disease mapping models in Table 1 have their multivariate extensions
Table 2
A selected multivariate CARs. $\psi$ is a $n \times K$ matrix, for $K$-variate CARs.

| Model | $\psi$ | $\Sigma_M$ |
|-------|--------|------------|
| MiCAR($\Sigma$) (G+V 2003) | $\text{vec}(\psi^T)$ | $(D_m - \mathbf{W})^{-1} \otimes \Sigma$ |
| MpCAR($c$, $\Sigma$) (G+V 2003) | $\text{vec}(\psi^T)$ | $(D_m - c \mathbf{W})^{-1} \otimes \Sigma$ |
| MLCAR($c$, $\Sigma$) (M+G 2007) | $\text{vec}(\psi^T)$ | $(c(D_m - \mathbf{W}) + (1 - c) J_n)^{-1} \otimes \Sigma$ |
| MLCAR($c$, $\Sigma$) (MacNab, 2018) | $\text{vec}(\psi^T)$ | $(I_n \otimes A) \text{bdia}[(c_1(D_m - \mathbf{W}) + (1 - c_k) J_n)^{-1}] (I_n \otimes A)$ |
| MpCAR($C_j$, $\Sigma$) (Jin et al., 2007) | $\text{vec}(\psi)$ | $(A \otimes I_n) S(C_j)^{-1} (A^T \otimes I_n)$ |
| MpCAR($C$, $\Sigma$) (G+T 2009) | $\text{vec}(\psi)$ | $(A \otimes I_n) S(C)^{-1} (A^T \otimes I_n)$ |
| MpCAR($C$, $\rho$, $\sigma$) (Sain et al., 2010) | $\text{vec}(\psi)$ | $(\sigma^2 \otimes I_n) S(\rho, C)^{-1} (\sigma^2 \otimes I_n)$ |
| MultBYM($\Sigma_s$, $\Sigma_n$) | $\text{vec}(\psi) = \text{vec}(\psi^s) + \text{vec}(\psi^h)$ | $(D_m - \mathbf{W})^{-1} \otimes \Sigma_s + I_n \otimes \Sigma_h$ |

1: Gelfand and Vounatsou (2003). 2: MacNab and Gustafson (2007). 3: Greco and Trivisano (2009). $\Sigma$ is a $K \times K$ covariance matrix. Bdiag$[(c_1(D_m - \mathbf{W}) + (1 - c_1) J_n)^{-1}] = \text{Bdiag}[(c_1(D_m - \mathbf{W}) + (1 - c_1) J_n)^{-1}], \ldots, (c_k(D_m - \mathbf{W}) + (1 - c_k) J_n)^{-1}]$.

$S(C)$, $D_m$, $C$, $\Sigma_s$, $\Sigma_n$, $\Psi^s$, $\Psi^h$, $\Sigma_s$, $\Sigma_h$ are (assumed) statistically independent; $\Psi^s \sim \text{MiCAR}(\Sigma)$, $\Psi^h \sim \text{IIDN}(\Sigma)$; $\Psi_i^s$ and $\Psi_j^h$, $\forall k$, are (assumed) statistically independent.

presented in the literature. A list of illustrative models are presented in Table 2, where their joint covariance matrices are given.

For mapping two or more diseases, for example, conditionally formulated multivariate GMRFs, such as the 2-fold CAR of Kim et al. (2001), the proper multivariate CARs of Gelfand and Vounatsou (2003), and the coregionalized MpCARs of Jin et al. (2007), were proposed for modelling multivariate disease risks that are correlated at co-locations as well as within- and cross-neighbourhood (i.e. cross-correlation between diseases at neighbouring locations). It is worth mentioning that the three groups of authors took three approaches to arrive their MCARS, all are multivariate generalizations of the pCAR. Working with a matrix of $\psi$, one has a variety of ways to construct MGMRFs; see MacNab (2018) for a discussion and references therein.

Since the influential works of Besag (1974) and Mardia (1988), MGMRFs are often formulated via full conditionals (e.g. Kim et al., 2001; Gelfand and Vounatsou, 2003; Sain and Cressie, 2007; Sain et al., 2011). However, as shown in MacNab (2018), formulation and implementation of MGMRFs face several challenges, most notably the entanglement of spatial and non-spatial parameterizations, compliance with symmetry requirement, and enforcement for positivity. With the exception of separable MCARS, neither the Besag (1974) nor the Mardia (1988) framework readily enables multivariate generalizations of all CARs.

One example is the LCAR, for which non-separable multivariate generalization of full conditionals is not readily derived, due in part to its entangled spatial and non-spatial parameterizations. The MLCAR($c$, $\Sigma$) presented in Table 2 was derived via linear transformation of independent LCARs (MacNab, 2016a), a method known as coregionalization (Jin et al., 2007).

Linear transformation is a well known and powerful idea for multivariate generalizations of univariate distributions and models. In a series of recent papers (MacNab, 2016a,b, 2018, 2021a,b), and built on the works of Gelfand et al. (2004), Jin et al. (2007), Greco and Trivisano (2009), MacNab develops a coregionalization framework that connects and unifies the various lines of MGMRF development. The framework enables a systematic development, categorization, and implementation of a broad range of multivariate models via linear or spatially varying coregionalization, with
extensions to locally adaptive models. The coregionalization methods readily facilitate multivariate generalizations of (any) GMRF and contain the major MCARs proposed in the literature.

To briefly explain the key ideas, consider a \( p \)-variate latent GMRF construction \( \text{vec}(\eta^\top) \sim \text{MCAR}(\theta, S(C)) \) with the following precision matrix (Greco and Trivisano, 2009; MacNab, 2018)

\[
S(C) = D_m \otimes I_p - (W_U \otimes C + W_U^\top \otimes C^\top), \quad C \neq C^\top,
\]

(11) provided \( S(C) > 0 \), where \( C \) is a \( p \times p \) full matrix of spatial dependence parameters, \( D_m = \text{diag}(m_1, m_2, \ldots, m_n) \), \( m_2, \forall i \) are area-specific scaling factors, \( W_U \) is the upper triangular part of \( W, W \) is the “neighbourhood” connectivity or weight matrix. \( S(C) \) is a spatial dependence matrix for locally dependent GMRF constructions. Expression (11) is simplified to \( S(C_s) = D_m \otimes I_p - (W \otimes C_s) \) when \( C_s \) is a symmetric matrix or a diagonal matrix of spatial dependence parameters. It also reduces to a separable model when \( C = CI_p \).

Without the “entanglement” challenges, GMRFs of construction (11) can be readily formulated within the Besag (1974) or Mardia (1988) framework (MacNab, 2018).

Linear transformation of \( \text{vec}(\eta^\top) \), denoted

\[
\text{vec}(\psi^\top) = (I_n \otimes A)\text{vec}(\eta^\top),
\]

(12) leads to a coregionalization GMRF (cGMRF) with covariance matrix (Jin et al., 2007; Greco and Trivisano, 2009; MacNab, 2016a,b, 2018):

\[
\Sigma_{\text{cGMRF}}(C, A) = (I_n \otimes A)S(C)^{-1}(I_n \otimes A),
\]

(13)

where \( A \) is a \( K \times K \) coregionalization coefficients matrix, \( AA^\top = \Sigma > 0, \Sigma \) is a covariance matrix.

Notice that (12) and (13) exemplify one way to decompose and mix spatial and non-spatial components that characterizes and parameterizes cross-spatial interactions. Alternatively, let the latent fields be \( \text{vec}(\eta^\top) \sim \text{MVN}(0, I_n \otimes \Sigma) \), a spatially varying (non-stationary) coregionalization (SVC) construction may be formulated via

\[
\text{vec}(\psi^\top) = H(C)\text{vec}(\eta^\top),
\]

(14) to have the following covariance matrix (MacNab, 2018):

\[
\Sigma_{\text{vec}(\psi^\top)}(C, \Sigma) = H(C)(I_n \otimes \Sigma)H(C)^\top,
\]

(15)

where \( H(C)H(C)^\top = S(C)^{-1} \). In SVC (15) the latent variables are only correlated at same locations. SVC constructions with latent (M)GMRF can also be readily formulated.

The decomposition and mixing schemes of (12) and (14) are typically used to produce valid covariance and cross-covariance functions for cGMRFs. The precision matrices of the cGMRFs lose their appeal of interpreting cross-spatial dependencies (MacNab, 2018, 2021a,b).

As an alternative to (12) and (14), a cGMRF can be formulated via

\[
\text{vec}(\psi^\top) = (I_n \otimes \sigma)\text{vec}(\eta^\top)
\]

(16) to have the following precision matrix:

\[
\Omega_{\text{vec}(\psi^\top)}(C, \rho, \sigma) = (I_n \otimes \sigma^{-1})S(\rho, C)(I_n \otimes \sigma^{-1}),
\]

(17)

where \( S(\rho, C) = I_n \otimes \rho + S(C) \), \( \rho \) is a \( p \) by \( p \) non-spatial (with-in location) partial correlation matrix, \( \sigma \) is a diagonal matrix of \( p \) scale parameters (MacNab, 2018, 2020, 2021a). Notice that \( S(\rho, C) \) is a precision (dependence) matrix of a latent GMRF; its full conditionals are readily available (Sain et al., 2011; MacNab, 2018, 2020, 2021a). It models spatial dependencies by \( S(C) \) and non-spatial dependencies by \( \rho \). A main feature of this approach to separating and mixing the cGMRF spatial and non-spatial components, with its spatial and non-spatial parameterizations, is that the cGMRF preserves the appeal of interpreting cross-spatial dependencies in the latent GMRF with precision matrix \( S(\rho, C) \); also see MacNab (2018, 2021a) for connections between the cGMRFs (13) and (17).

With options of parameterization for \( S(C), \rho, A, \Sigma \), the three constructions contain all major proposals of MGMRFs to date and much more; they are rich options of multivariate generalizations
of pCAR. These constructions may be parameterized to build multivariate spatial and spatiotemporal smoothers, as well as multivariate spatial and spatiotemporal covariance models. cMGMRFs of constructions (13) and (15) could be parameterized for latent spatial components and factor analysis, include, but not limited to, principal spatial components analysis and dimensionality reduction (see MacNab, 2018, 2021a, for illustrative examples).

Of the cMGMRF (17) construction, options of parameterizations for the spatial and non-spatial components may be considered for their intuitively plausible characterizations and interpretations of conditional spatial dependencies within and between variables, as well as local interactions in space and time (Sain et al., 2011; MacNab, 2018, 2020, 2021a,b; Prates et al., 2021).

A $p$-variate GMRF that allows for asymmetric cross-covariance functions can be formulated by linear coregionalization of latent MGMRF with a full matrix $C \neq C^\top$ of spatial and cross-spatial dependence parameters or by spatially varying coregionalization of $p$ independence GMRFs; the latter is a non-stationary adaptive MGMRF with adaptive non-spatial parameterization (Martinez-Beneito, 2020; MacNab, 2021a). In addition, as I show in Section 8, non-stationary GMRFs that allow for asymmetric pair-wise spatial dependencies can be formulated via adaptive parameterization.

8. Adaptive CARs

GMRFs are undirected graphical models, commonly defined for a lattice system of nodes and edges (Rue and Held, 2005). Of the GMRFs commonly used in disease mapping (say, those in Table 1), the edges are defined by the $W$ matrix of a given map and its neighbourhood definition. These GMRFs often have single spatial and/or scale parameter. In the context of disease mapping, they postulate pair-wise interactions of neighbouring risks. They typically lead to smoothly varying posterior relative risk prediction and lack the flexibility to allow within- and between-neighbourhood risk heterogeneities that likely happen in parts of the map (MacNab et al., 2006; Brewer and Nolan, 2007).

To mitigate the mentioned limitations, adaptive CARs have been proposed with applications to analysis of imaging data (Brezger et al., 2007), analysis of correlated data (Reich and Hodges, 2008), and disease mapping (Lu and Carlin, 2005; MacNab et al., 2006; MacNab, 2018; Lu et al., 2007; Brewer and Nolan, 2007; Congdon, 2008; Lee and Mitchell, 2012; Rushworth et al., 2017; Gao and Bradley, 2019; Corpas-Burgos and Martinez-Beneito, 2020, among others).

Adaptive CARs proposed to date can be broadly grouped under two motivating directions. One is to hierarchically formulate adaptive CAR(GMRF) with unknown adjacency matrix to be modelled and estimated from data. This is the direction taking by Lu and Carlin (2005), Lu et al. (2007), Lee and Mitchell (2012), Rushworth et al. (2017), and Gao and Bradley (2019), among others. With the exception of Gao and Bradley (2019), these authors proposed adaptive iCAR, CAR (5), and LCAR in which the non-zero elements of $W = (w_{ij}, \forall i \sim j)$ were modelled as Bernoulli random variates or random variates of unit interval (i.e. $w_{ij} \in (0,1), \forall i \sim j$, Rushworth et al., 2017). In Gao and Bradley (2019), an adaptive iCAR was formulated for unknown adjacency matrix $W$, where $w_{ij}, \forall i \neq j$, were modelled as Bernoulli random variates. This line of research primarily concerned with adjacency modelling for boundary analysis and detection (Lu et al., 2007; Lee and Mitchell, 2012; Gao and Bradley, 2019) and/or analysis of spatial discontinuity in disease risk maps (Lee et al., 2014; Rushworth et al., 2017; Sahu and Böhning, 2021).

An adaptive LCAR was also proposed in Baptista et al. (2016), in which the $W$ matrix is replaced by a similarity matrix $S = (S_{ij})$, where the elements $S_{ij}, \forall i \neq j$, are functions of covariates that characterize risk similarities over a map.

Formulating adaptively parameterized CARs over a given $W$ is the second approach. Under this approach, adaptive CARs have been proposed for each of the CARs presented in Table 1; the key proposals are presented in Table 3. This group of adaptive CARs may be used to model locally varying (asymmetric) spatial dependencies, interactions, and heterogeneities. These adaptive CARs might capture micro-level phenomenons such as locally varying (asymmetric) risk dependencies and influences and heterogeneities that typically manifest macro-level phenomenons such as the spatial risk heterogeneities and discontinuity. They can be broadly sub-grouped into three categories. One contains CARs with adaptive spatial (dependence) parameters; they were typically motivated to
The higher the $c_i$, the greater the $\psi_i$ prediction is directly influenced by its neighbouring $\{\psi_j, \forall j \sim i\}$. When $c_i$ is near-zero, prediction of $\psi_i$ is not directly influenced by its neighbouring $\{\psi_j, \forall j \sim i\}$.

The pCAR(b) is a variant of the adaptive pCAR proposed in MacNab (2018). Here, the pCAR(b) has an additional spatial dependence parameter $\lambda$ that characterizes neighbourhood risk dependence under the new weight matrix $W^c$ (see Table 3). Similar to its non-adaptive counterpart, its coefficient of influence functions imply asymmetric pair-wise local dependencies and influences, provided $w_{ij+} \neq w_{i+j}$: An area with a higher number of neighbours is less influenced by its neighbour who has a lower number of neighbours. Two neighbours exert same influence on each other if and only if they have the same number of neighbours.

### 8.1. Adaptive CARs of the category I

There are important differences between the adaptive pCARs and LCARs of this category. A common feature of the adaptive pCARs is that the adaptive parameters $\{c_i, \forall i\}$ control spatial dependencies in the conditional mean functions, whereas its scale parameter $\sigma$ regulates (log relative) risk prediction variances in the conditional variance functions (see Table 3). The notable differences between the pCARs are their influence functions, as illustrated here for the pCAR(a) and pCAR(b):

$$\text{Influence}(j, i)_{\text{pCAR}(a)} = \frac{c_i}{w_{ij+}} \quad \text{Influence}(j, i)_{\text{pCAR}(b)} = \frac{\lambda c_i c_j}{w_{ij+}}. \quad (18)$$

In pCAR(a), $c_i$ quantify how each area is directly influenced by its neighbours: The higher the $c_i$, the greater the $\psi_i$ prediction is directly influenced by its neighbouring $\{\psi_j, \forall j \sim i\}$. When $c_i$ is near-zero, prediction of $\psi_i$ is not directly influenced by its neighbouring $\{\psi_j, \forall j \sim i\}$.
Similar to the LCAR(a), the adaptive pCAR(a) full conditionals do not lead to a GMRF due to noncompliance with the GMRF symmetry requirement. Nevertheless, the adaptive pCAR(a) or LCAR(a) can be used as a prior in a Bayesian hierarchical model for characterizing locally varying spatial dependencies and heterogeneities (which I illustrate in Section 9).

In the LCARs, the adaptive spatial parameters serve dual roles of regulating locally varying spatial dependencies and heterogeneities (see Table 3). For example, the Influence\(\psi_{ij}\) of \(i\) on \(j\) in \(i\)CAR \(= c_i/(1 + c_i(w_{ij} + 1))\), \(\forall j \sim i\), is a non-linear increasing function of \(c_i\), the corresponding conditional variance function is a non-linear decreasing function of \(c_i\).

Illustrated in Table 3, the LCAR(b) and LCAR(c) are comparable GMRFs. While the LCAR(c) reduces to its non-adaptive LCAR, the LCAR(b) does not. Compared to the LCAR(a), the two have more complex influence functions for interpreting locally varying spatial dependencies and influences (given in Table 3). An alternative interpretation of the two is that, as adaptive GMRFs, their adaptive parameters are mixing parameters that allows adaptive mixing of spatial and non-spatial smoothing (Congdon, 2008).

8.2. Adaptive CARs of the category II

Adaptive CARs of this category began with proposals of adaptive iCAR, primarily motivated for modelling spatial discontinuity and heterogeneity. The basic idea is to replace the connectivity matrix \(W\) in iCAR by adaptive parameterization of a weight matrix of non-negative elements, denoted \(W = (w_{ij}, \forall i \sim j)\) in Table 3, where \(w_{ij}\), \(\forall i \sim j\), models locally weighted pair-wise interaction between \(\psi_i\) and \(\psi_j\). In Brezger et al. (2007), for example, the elements \(\{w_{ij}, \forall i \sim j\}\) were estimated via a gamma prior. Brewer and Nolan (2007), Reich and Hodges (2008), and Corpas-Burgos and Martinez-Beneito (2020) proposed more parsimonious parameterizations to \(W\): \(w_{ij} = \tau_1\tau_2/(\tau_1 + \tau_2)\) in Brewer and Nolan (2007), \(w_{ij} = (\sigma_i\sigma_j)^{-1}w_{ij}\) in Reich and Hodges (2008), and \(w_{ij} = (c_i\sigma_i)^{1/2}w_{ij}\) in Corpas-Burgos and Martinez-Beneito (2020). The Corpas-Burgos and Martinez-Beneito (2020) adaptive iCAR can be viewed as a re-parameterization of the Reich and Hodges (2008) adaptive iCAR, by letting \(\sigma_i = \sigma c_i^{1/2}\), where \(\sigma\) is an additional scale parameter. The iCAR(a) in Table 3 is the Corpas-Burgos and Martinez-Beneito (2020) adaptive iCAR equivalent, with \(c_i^{1/2}\) being replaced by \(c_i\).

With an adaptive iCAR, one can derive its extensions of adaptive LCAR or pCAR. For example, Corpas-Burgos and Martinez-Beneito (2020) extended their adaptive iCAR to two adaptive LCARS; the LCAR(d) and LCAR(e) in Table 3 (as extensions of the iCAR(a)) are the equivalents. The pCAR(c) in Table 3 is a pCAR extension of the iCAR(a).

Similar adaptive iCAR, pCAR and LCAR of this category can be formulated via linear transformation of non-adaptive CARs. For example, let

\[ \Psi = \tilde{\sigma} \Psi, \quad \tilde{\Psi} \sim \text{CAR}, \]

where \(\tilde{\sigma} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)\), and \(\tilde{\Psi} \sim \text{iCAR}(1)\) or \(\tilde{\Psi} \sim \text{LCAR}(\lambda)\) or \(\tilde{\Psi} \sim \text{pCAR}(\lambda)\). Linear transformation (19), named linear (model of) coregionalization in MacNab (2018), leads to adaptive iCAR or pCAR or LCAR. The adaptive iCAR(b), pCAR(d), and LCAR(f) in Table 3 are examples.

A common feature of this category of CARs is that the adaptive parameters serve dual roles: They regulate both the spatial risk dependencies in the conditional mean functions and risk heterogeneities in the conditional variance functions (see Table 3). A noteworthy difference among them is their influence functions, as illustrated here for the iCAR(a) and iCAR(b):

\[ \text{Influence}(j, i)_{\text{iCAR}(a)} = \frac{c_j}{\sum_{j \sim i} c_j}, \quad \forall j \sim i, \quad \text{Influence}(j, i)_{\text{iCAR}(b)} = \frac{c_j}{c_i w_{ij}}, \quad \forall j \sim i. \]

The influence function \(\text{Influence}(j, i)_{\text{iCAR}(a)}\) in (20) quantifies relative influence of area \(j\) on area \(i\), relative to the neighbours of the area \(i\), in terms of its direct influence on the prediction of \(\psi_i\). When \(c_j\) is near-zero, area \(j\) exerts near-zero direct influence on its neighbours, but is not necessarily “disconnected” from its neighbours, which was suggested by Corpas-Burgos and Martinez-Beneito (2020). This is because area \(j\) could still be directly influenced by its neighbours, say, when \(c_i > 0\) and \(i \sim j\). Influence\((i, j)_{\text{iCAR}(a)} = c_i/\sum_{i \sim j} c_i > 0\).
On the other hand, Influence\((j, i)\)\(^{\text{CAR}(b)}\) quantifies direct influence of area \(j\) on area \(i\), which is partially controlled and partially determined by three quantities. It is an increasing function of \(c_j\) but decreasing function of \(c_i\): For given \(c_i\), the higher the \(c_j\), the greater the area \(j\) exerts direct influence on area \(i\). The greater the \(c_i\) and/or the higher the \(w_{ij}\), the less the area \(i\) is influenced by its neighbours. Area \(j\) exerts greater influence on area \(i\) if and only if \(c_j^2 w_{ij} > c_i^2 w_{ji}\).

The LCAR\((e)\) and LCAR\((f)\) share a common feature that, when \(\lambda = 0\), they both reduce to adaptive independent Gaussian distributions defined by a diagonal variance matrix diag\((\sigma^2c)\).

It is worth mentioning that the adaptive LCARs in the Category I are similar to the adaptive CARs of this category in the sense that their adaptive parameters regulate both spatial dependencies and heterogeneities.

In addition, in the context of Bayesian hierarchical modelling, the Rushworth et al. (2017) adaptive LCAR maybe viewed as an adaptive CAR of this category. For example, a LCAR (extension) of the adaptive iCAR\((\sigma, W)\) with \(\tilde{w}_{ij} = c_{ij}\) is a Rushworth et al. (2017) adaptive LCAR equivalent. It is readily seen this LCAR (and the Rushwoth et al. model) may be excessively parameterized. To gain identifiability, the LCAR can be simplified by letting \(\tilde{w}_{ij} = c_i c_j\) is a Corpus-Burgos and Martinez-Beneito (2020) adaptive LCAR\((d)\) in Table 3 (with \(c_i^{1/2}\) being replaced by \(c_i\), \(c_i \in (0, 1)\), \(\forall i\)).

### 8.3. Adaptive CARs of the category III

Adaptive CARs of this category can be derived via linear transformation of adaptive CARs of the Category I. The adaptive pCAR and LCAR proposed in MacNab (2018) are examples; they can be readily derived via linear transformation (19), where \(\tilde{\psi} \sim \text{pCAR}(c)\) or \(\tilde{\psi} \sim \text{LCAR}(c)\), pCAR\((c)\) and LCAR\((c)\) are adaptive CARs without scale parameter.

As the pCAR\((e)\) in Table 3 illustrates, the coefficients of influence (in CARs of this category) are functions of adaptive spatial and scale (or precision) parameters; the adaptive scale parameters regulate risk heterogeneities in the conditional variance functions.

The coregionalized adaptive pCAR offers the flexibility for modelling locally varying spatial (Markovian) dependencies in the latent components \(\tilde{\psi}\), regulated by an adaptive spatial dependence parameterization in its conditional mean functions. Linear transformation (19) introduces locally varying risk heterogeneities to \(\psi\), regulated by the adaptively parameterized coefficients of coregionalization \(\tilde{\sigma}\).

In the context of disease mapping, CARs of this category may be excessively parameterized. However, they may be useful when additional information, such as covariates, are available to infer these parameters (MacNab, 2018).

### 9. An illustrative example: Spatial and spatiotemporal modelling of COVID-19 infection risks

The purpose of this illustrative example is to show (potential) utilities of some of the disease mapping models and methods for analysing and monitoring communicable disease such as the COVID-19 infection risks during an ongoing epidemic or pandemic. I focus on potentials of the adaptive CAR models presented herein, in the context of Bayesian mapping of spatial and spatiotemporal COVID-19 infection risks, without or with covariates.

The overall analytic approach aimed for small area pandemic minoring by describing and understanding locally varying spatial risk dependencies, asymmetric local risk influence functions, within and between neighbourhood risks heterogeneities, and the resulting spatial risk discontinuities.

Details on models and related Bayesian methods for parameter estimation, learning, and inference, and details of the data analysed, and additional results and discussions, will be presented elsewhere. Here, for brevity, is a cursory report that outlines two sets of analysis, one is spatial modelling of county-level aggregates of daily infection cases for eighty-seven counties of Minnesota, USA, over the period of 2020/01/22–2021/02/14, named the cumulative period (data) hereafter. The second is spatiotemporal modelling of weekly aggregates of infection cases for the period of 2020/09/30–2021/01/20, a period the State-level 7-day averaged number of new cases exceeded 1000, named the peak period (data) hereafter (the State’s first major infection wave). I illustrate
Fig. 1. Posterior estimates of the county-specific (87 counties) adaptive parameters and relative risks for indicated models. Red line or dot: without covariate, blue line: with five covariates. The Minnesota county-level COVID-19 cumulative period data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Posterior estimates, median and standard deviation (sd), of the model parameters for the adaptive pCAR(b). The Minnesota bcounty-level COVID-19 cumulative period data.

| Parameter                              | Without covariate | With covariates |
|----------------------------------------|-------------------|-----------------|
|                                        | Median            | Median          |
|                                        | sd                | sd              |
| Intercept                              | 0.02              | 0.02            |
| Private transportation to work         | 0.88              | 0.68            |
| Age 55–64                              | −4.32             | 0.96            |
| Education less than high school        | 3.94              | 0.70            |
| College education                      | 0.95              | 0.60            |
| Unemployment                           | −4.56             | 1.35            |
| λ                                      | 0.91              | 0.67            |
| σ                                      | 0.40              | 0.32            |

spatiotemporal modelling of county-level weekly aggregates of 17 weeks. The covariates presented here are tentative and only used for illustrative purpose; they are census-based aggregates and will be explained and further discussed in a separate paper.

9.1. Spatial analysis of the cumulative period data

The Poisson data model (1) with a regression part was fitted, in which the vector of log relative risks \( \psi = (\psi_1, \ldots, \psi_{87}) \) was modelled by each of the adaptive CARs in Table 3. Overall,
the adaptive CARs led to comparable goodness-of-fit among the spatial models. Fig. 1 presents posterior estimates of the adaptive parameters and relative risks for two illustrative models; the results from the LCAR(a) illustrate spatial discontinuity, and those from the pCAR(b) illustrate locally varying neighbourhood risk dependencies. Fig. 1 and Table 4 suggest that the included (time-invariant) covariates explained modest (but noteworthy) amount of neighbourhood risk dependencies, modelled by the spatial dependence parameter $\lambda$; they also suggest that the included covariates explained modest amount of risk variability, modelled by the scale parameter $\sigma$.

9.2. Spatiotemporal analysis of the peak period data

The Poisson data model (1) was extended to a spatiotemporal data model

$$O_{it} \sim \text{Poisson}(\phi_{it}), \quad \phi_{it} = E_{it} \psi_{it}, \forall i, t,$$  \hspace{1cm} (21)

where

$$\log(\phi_{it}) = \beta_{0t} + \beta_{1t} X_{i1} + \cdots + \beta_{pt} X_{ip} + \psi_{it}, \quad p = 5,$$  \hspace{1cm} (22)

the regression part in Eq. (22) has time-varying coefficients for the included (time-invariant) covariates ($p = 5$).

The data were analysed via three sets of spatiotemporal models. I began with modelling the weekly data as time independent, and with weekly (i.e. time-varying) adaptive CARs, say,

$$\psi_{t} \sim \text{adaptive CAR}(c_{t}, \sigma_{t}).$$  \hspace{1cm} (23)

$$\psi_{t} = (\psi_{1t}, \psi_{2t}, \ldots, \psi_{nt})^T, \quad c_{t} = (c_{1t}, c_{2t}, \ldots, c_{nt})^T, \quad \forall t,$$

where CAR (23) is an example of adaptive CARs of the Category I. The main purpose of the analysis was to describe and understand weekly changes in (i) locally varying spatial risk dependencies, (ii) asymmetric neighbourhood risk influences, and (iii) spatial risk heterogeneities and discontinuities.

As anticipated, and illustrated in Fig. 2 for indicated adaptive CARs, the spatial patterns of the adaptive parameters and relative risks varied over the 17 weeks. The included covariates explained modest amount of variations in the adaptive parameters and the relative risks during the week 7, which was among the peak weeks of this period. The observed patterns were also consistent with the posterior estimates of time-varying regression coefficients presented in Fig. 3, which illustrates a likely scenario of time-varying effects of covariates, such as the small area and neighbourhood characteristics included in the model. Fig. 3 also indicates that the included covariates led to modest reductions in the time-varying scale parameters over the peak weeks (the weeks 7 to 10).

Fig. 4 suggests modest variations in the estimated coefficients of influence for a selection of counties, using the results from the adaptive pCAR(d) (a model that had one of the lowest DIC, see Table 5). It also suggests some modest changes in the local risk influences from week 7 to week 17, and minor or modest changes from unadjusted to adjusted risk influences.

For the second set of analysis, the spatiotemporal model defined by Eqs. (21) and (22) was fitted for eight options of time-invariant adaptive CAR parameterization. An example of the adaptive CARs of the category I is

$$\psi_{t} \sim \text{adaptive CAR}(c, \sigma), \quad \forall t,$$  \hspace{1cm} (24)

which is a simplification of Eq. (23). Table 5 presents the DIC results for the indicated adaptive CARs (1)–(8). Time- or variable-invariant adaptive CARs were previously used in Bayesian spatiotemporal models (Reich and Hodges, 2008; Rushworth et al., 2017; Sahu and Böhning, 2021) or multivariate spatial models (Corpas-Burgos and Martinez-Beneito, 2020). In this study, an intention of the analysis was to observe and compare the estimated time-invariant adaptive parameters with their counterparts of time-varying parameters. Fig. 5 illustrates the results for the pCAR(a). The patterns displayed from the estimated time-invariant spatial dependence parameters (the plots on the left) are notably different from those of the time-varying parameters for week 7 (the plots on the right). They convey different and important information: The estimated time-invariant $c$ may serve to explain and inform locally varying neighbourhood risk influences and spatial discontinuity.
Fig. 2. Illustrations of posterior estimates of county-specific (87 counties) adaptive parameters and relative risks for the indicated models. For the adaptive parameter estimates: Red lines—without covariate, blue lines—with covariates. For the relative risk estimates: Red dots: without covariate, blue lines: with covariates. The Minnesota county-level COVID-19 weekly data during peak period. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Fig. 3. Posterior estimates of the week-specific (17 weeks) regression coefficients and scale parameters for the indicated models. For the regression coefficients: dashed line and dots: posterior median, solid lines: lower and upper limits of 95% credible intervals. For the scale parameters: Red line: without covariate, blue line: with covariates. The Minnesota county-level COVID-19 weekly data during peak period. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Fig. 4. Posterior estimates of the coefficients of influence for indicated counties, calculated using the posterior medians of the adaptive parameters in pCAR(d) (see Table 3). Red dots: Unadjusted risk influences—estimated coefficients of risk influence based on indicated ST model without covariate; blue dots: Adjusted risk influences—estimated coefficients of risk influence based on indicated ST model with covariates. The Minnesota county-level COVID-19 weekly data during peak period.
Fig. 5. Posterior estimates (median and standard deviation) of county-specific (87 counties) adaptive parameters for the indicated models. Red dots: Unadjusted risk influences—estimated coefficients of risk influence based on indicated ST model without covariate; blue dots: Adjusted risk influences—estimated coefficients of risk influence based on indicated ST model with covariates. The Minnesota COVID-19 weekly data during peak period.

Table 5

| Model   | Without covariate | With covariates |
|---------|-------------------|-----------------|
|         | Dbar   | pD    | DIC    | Dbar   | pD    | DIC    |
| (1) pCAR(a) | 10693  | 1237  | 11930  | 10687  | 1242  | 11929  |
| (2) pCAR(b) | 10687  | 1243  | 11930  | 10682  | 1247  | 11929  |
| (3) pCAR(c) | 10688  | 1223  | 11911  | 10690  | 1225  | 11915  |
| (4) pCAR(d) | 10692  | 1212  | 11904  | 10689  | 1214  | 11903  |
| (5) LCAR(a) | 10704  | 1227  | 11931  | 10692  | 1230  | 11922  |
| (6) LCAR(b) | 10694  | 1227  | 11921  | 10682  | 1230  | 11912  |
| (7) LCAR(e) | 10677  | 1228  | 11905  | 10676  | 1227  | 11903  |
| (8) LCAR(f) | 10697  | 1208  | 11905  | 10693  | 1212  | 11905  |
| (9) SVC LCAR TpCAR1 | 10706  | 1222  | 11928  | 10706  | 1215  | 11921  |
| (10) SVC LCAR TpCAR2 | 10671  | 1121  | 11792  | 10675  | 1120  | 11795  |

over the first wave of the pandemic, while the weekly estimates of the adaptive parameters may serve to explain retrospectively, or monitor on going, temporal dynamics of locally varying risk dependencies, varying neighbourhood risk influences, and spatial discontinuities.

The eight adaptive CAR models led to comparable and consistent posterior estimates of relative risks, although modest differences were observed. Overall, the CARs of the Category II performed slightly better than the CARs of the Category I. The LCARs also had slightly smaller DIC scores,
compared with their pCAR counterparts. Of note is that the LCAR(e) (Corpas-Burgos and Martinez-Beneito, 2020) and the LCAR(d) and pCAR(f) of linear models of coregionalization (MacNab, 2018) led to comparable DIC results, although the LCAR(d) and pCAR(f) were shown to be more efficient, as indicated by the smaller pD scores and posterior relative risk standard deviations (not presented here). The overall DIC comparison, and the estimated adaptive parameters and relative risks, consistently suggested notable degrees of locally varying spatial dependencies, asymmetric local risk influences, and spatial discontinuities.

The third set of analysis concerned with CARs of the Category III, they are illustrative examples of exploring and explaining adaptive spatiotemporal risk interactions via a SVC model with spatially adaptive area-specific temporal CAR processes:

$$\text{vec}(\psi) = (I_T \otimes \text{diag}(\tilde{\sigma})) \text{vec}(\eta),$$  

$$(25)$$

where $\tilde{\sigma} = (\sigma s_1, \ldots , \sigma s_n)$, $\log(s) \sim \text{LCAR}(\lambda, W)$ and the vector of latent variables $\text{vec}(\eta)$ is modelled by a spatially adaptive temporal pCAR:

$$\text{vec}(\eta) \sim \text{pCAR}(c, W_T),$$  

$$(26)$$

where $W_T$ is a 17 by 17 first-order temporal connectivity matrix $t \sim t+1$ for “TpCAR1” and $t \sim t+1$ and/or $t \sim t-1$ for “TpCAR2”. The last model with the temporal pCAR2 led to the best DIC with notably improved statistical efficiency, as indicated by a notably reduced pD (in Table 5) and smaller posterior relative risk standard deviations (not presented here). The two sets of adaptive parameters $s$ and $c$ in the SVC models convey important information on asymmetric spatial and temporal risks dependencies and influences, which will be presented and discussed in a separate paper.

10. Discussion and looking into the future

The importance and popularity of Bayesian disease mapping in spatial epidemiology, population and public health, and medicine are seen in the large and still growing literature on spatial and spatiotemporal small area mapping of risks of COVID-19 infection and related health outcomes, and on applications of CAR models in COVID-19 related research. Identifications of high risk areas and patterns of disease incidence and mortality can provide valuable information for public health planning such as priority setting for allocating funds and localized disease prevention or intervention. In addition to mapping disease and illness, novel disease mapping models have been developed for spatial modelling of child growth (Gelfand and Vounatsou, 2003), geographical mapping of gene frequencies (Gelfand and Vounatsou, 2003), mapping health care utilization rates (Nathoo and Ghosh, 2012), and vaccination coverage mapping (Utazi et al., 2018), among others.

The (still on-going) COVID-19 pandemic, being an unprecedented global crisis, also offers unparalleled opportunities for the statistics community to put the established and newly developed disease mapping methods and tools, particularly the commonly used CAR, MCAR, and adaptive (M)CAR models, to rigorous assessment and test, to improve the existing ideas, methods, and tools, and to broaden the subject further. The colossal amount of COVID-19 related data available around the world are intrinsically spatial, spatiotemporal, multivariate (e.g. infection, hospitalization, mortality, uptake of vaccination, etc.), and multi-array (e.g. concerns spatial, temporal, and multivariate outcomes, in addition to other dimensions such as age, gender, risk/protective factors, and countries, for the varying public health intervention policies, updates of preventative measures, and availabilities of vaccines, among others). Making these data accessible to statistical analysis is critically important in our efforts to better understand the current (and future) pandemic and its impact on population and public health and to inform on pandemic intervention policies and strategies.

We see in issues of the Spatial Statistics, and (bio)statistical journals elsewhere, that innovative (multivariate) spatial and spatiotemporal disease mapping models have been developed to estimate, explain, and map the risks or spread of COVID-19 infection or related mortality (to name some most recent ones, Li and Dey, 2021; Slater et al., 2021; Huang et al., 2021; Feng, 2021; Lee et al., 2021; Sahu and Böhning, 2021) and for short-term forecasts of infection or infection-related mortality rates (Sahu and Böhning, 2021).
Disease mapping models are widely known for modelling rare and non-communicable diseases such as heart disease and cancer. CARs are commonly motivated for borrowing information and modelling smoothly varying disease risks as effects of omitted covariates. However, typical Bayesian disease mapping models are hierarchically formulated Bayesian spatial GLMMs that allows a variety of data models to be considered and applied. For example, in addition to the Poisson model (1) for rare events, we also have the options of binomial models for non-rare events Martuzzi and Elliott, 1996; MacNab, 2003a, frailty models for time-to-event data (Carlin and Banerjee, 2003; Lawson, 2018), and Gaussian models for continuous data (MacNab, 2021b).

When non-rare events are studied, (M)CARs and adaptive (M)CARs may be considered to allow for nuanced (intelligent) smoothing and for characterization and learning of (cross) spatial risks dependencies and/or (cross) spatial risk correlations, as well as spatial risk discontinuity. For instance, areal data of adequate numbers of events may contain useful information to inform on more flexible spatial dependence parameterization, perhaps aided by additional covariates. The works of Congdon (2008) and Rushworth et al. (2017), and the present case study of spatiotemporal modelling of peak period COVID-19 infection risks, are examples.

Risks of communicable diseases such as the COVID-19 infection are quite likely to be spatially varying and to exhibit spatial discontinuity. This is due to the nature of communicable disease outbreak and transmission in proximities, differences in implementation and/or uptake of public health intervention policies and preventive measures, etc., in addition to the underlying health inequality that can partially influence the risks of infection or infection-related hospitalization or mortality. For these reasons, adaptive CARs might be more plausible options for modelling small area COVID-19 related data, without or with associated (protective) risk factors. The Sahu and Böhning (2021) work presents an example.

In this article I gave a more detailed review on formulating adaptive CARs via options of adaptive spatial dependence or weight parameterization for a known connectivity matrix $W$. One attraction of this approach is that the resulting models have clear interpretations of locally varying spatial risk dependencies and neighbourhood risk influences. Estimation and inference of the model parameters lead to estimation and inference of spatial dependence (local influence) functions, with associated estimation uncertainties. The adaptive spatial dependence parameterization discussed herein may mitigate the impact of a potentially unrealistic assumption of local risk dependencies defined by a given connectivity matrix, by allowing location-specific spatial dependence parameters to be zero, and thereby enables characterization of spatial risk discontinuity, as illustrated in Congdon (2008). In addition, some of the recent proposals of adaptive parameterizations of multivariate CARs, are natural extensions of the univariate adaptive CARs; the adaptive MGMRFs discussed in MacNab (2018, 2021a) are examples.

The approach to allowing the connectivity matrix $W$ to be modelled as (binary) random variates within a CAR formulation leads to an alternative class of adaptive models. These models were initially motivated for, and were shown to be effective in, adjacency modelling and boundary detection (Lu and Carlin, 2005; Lu et al., 2007; Lee and Mitchell, 2012). While adaptive CARs of this class were also used for modelling spatial risk discontinuity (Lee et al., 2014; Sahu and Böhning, 2021), they lack the flexibility of modelling locally varying spatial dependencies that offered by adaptive spatial dependence parameterization. Nevertheless, the two classes of adaptive CARs are useful tools for their intended purposes and may be considered as complementary approaches in real-life applications. Adaptive CARs that accommodate both adaptive spatial dependence (or weight) parameterization and adjacency modelling are possible options to be further explored and developed.

The disease mapping community is poised to embrace new challenges and opportunities brought by the rise of data science and a (big) rich data era. Earlier disease mapping literature presented
alternative ideas of modelling structured spatial discontinuities. Examples are hidden Markov models (Green and Richardson, 2002) and mixture models (Denison and Holmes, 2001; Fernandez and Green, 2002). These earlier ideas and methods have shown limited advantage in mapping rare disease. However, they may regain their recognition and appeal in the context of modelling (big) data of rich content, information, and structure. Together with (adaptive) CARs and their multivariate extensions, these models and related methods for estimation and inference may be further explored and better understood for their roles in disease mapping applications.

Recent advances in multivariate CAR/GRMFs, a topic of my focus here and a (new) comprehensive book has written for (Martinez-Beneito and Botella-Rocamora, 2019), provide ideas, models, methods, and tools for jointly modelling multidimensional spatial data in general, and the COVID-19 infection and related data in particular. Recent proposals of the INLA-SPDE approach to Bayesian analysis of spatial data over high resolution grids mark a new frontier of Bayesian disease mapping and its applications (Lindgren et al., 2011; Moraga et al., 2017; Utazi et al., 2018).

In addition, recent proposals of directed acyclic graph auto-regressive models in the context of (multivariate) disease mapping (Datta et al., 2019; Gao et al., 2021) offer alternative ways to model spatial dependencies. In the context of multivariate disease mapping, Liang (2019) proposed MCARs in which the non-spatial dependencies were characterized by a directed acyclic graph representation. The fusion of the Datta et al. (2019) and Liang (2019) ideas might broaden the scope and complexity of Bayesian disease mapping and widen its applicability in applied scientific (e.g. health science) research. The data science and big data movements are also inspiring new ideas of Bayesian disease mapping methods for handling data of high dimensions, e.g. areal data of large $n$ and/or $p$ and/or $T$. A recent example is the idea of “divide and conquer” (Orozco-Acosta et al., 2021) for scalable approximate Bayesian inference of disease mapping of large $n$ using INLA.

Perhaps one of the most urgent need for advancing Bayesian disease mapping and its broad applications is devoted efforts for improved or new methods of Bayesian estimation, learning, and inference of (adaptive and/or multivariate) CAR. INLA offers a solution to computation for models of large $n$ but small or modest number of (prior) model parameters (Rue et al., 2017). Recent proposals of Gibbs sampling of GMRFs on large lattice (Marcotte and Allard, 2018; Brown et al., 2021) offer ideas and options to be explored, adapted, and improved. Alternative computational options that might be fruitfully explored and furthered are the hybrid or Hamiltonian Monte Carlo sampling methods (Girolami and Calderhead, 2011; Zhang et al., 2018; Hoffman et al., 2021) and the variational inference (Blei et al., 2017; Tan, 2020).

Acknowledgement

This research was funded in part by a discovery grant (RGPIN 238660-13) from the Natural Sciences and Engineering Research Council of Canada.

References

Ainsworth, L.M., Dean, C.B., 2006. Approximate inference for disease mapping. Comput. Statist. Data Anal. 50 (10), 2552–2570.

Arnold-Forster, A., 2020. Mapmaking and mapthinking: cancer as a problem of place in nineteenth-century England. Social History of Med 33 (2), 463–488.

Arnold-Forster, A., 2021. The Cancer Problem: Malignancy in Nineteenth-Century Britain. Oxford Scholarship online book.

Azevedo, D.R., Prates, M.O., Bandyopadhyay, D., 2021. Mspock: Alleviating spatial confounding in multivariate disease mapping models. J. Agric. Biol. Environ. Stat. 1–28.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2004. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC.

Baptista, H., Mendes, J.M., MacNab, Y.C., Xavier, M., Almeida, J.C., 2016. A Gaussian random field mode for similarity-based smoothing in Bayesian disease mapping. Stat Methods Med Res 25 (4), 1166–1184.

Bernardinelli, L., Montomoli, C., 1992. Empirical bays versus fully bayesian analysis of geographical variation in disease risk. Stat. Med. 11 (8), 983–1007.

Bernardinelli, L., Pascutto, C., Best, N.G., Gilks, W.R., 1998. Disease mapping with errors in covariates. Stat. Med. 16 (7), 741–752.

Besag, J. 1974. Spatial interaction and the statistical analysis of lattice systems (with discussions). J. R. Stat. Soc. Ser. B Stat. Methodol. 36, 192–236.

Besag, J., York, J., Mollie, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math 43, 1–20.
Best, N., Richardson, S., Thomson, A., 2005. A comparison of Bayesian spatial models for disease mapping. Stat. Methods Med. Res. 14 (1), 35–59.
Blei, D.M., Kucukelbir, A., McAuliffe, J.D., 2017. Variational inference: a review for statisticians. J. Amer. Statist. Assoc. 112 (518), 859–877.
Breslow, N.E., 1984. Extra-Poisson variation in log-linear models. Appl. Stat. 33 (1), 38–44.
Breslow, N.E., Clayton, D.G., 1993. Approximate inference in generalized linear mixed models. J. Amer. Statist. Assoc. 88 (421), 9–25.
Breslow, N.E., Day, N.E., 1980. Statistical Methods in Cancer Research. Volume 1 - the Analysis of Case-Control Studies. International Agency for Research on Cancer, Lyon.
Brewer, M.J., Nolan, A.J., 2007. Variable smoothing in Bayesian intrinsic autoregressions. Environmetrics 18 (8), 841–857.
Brezger, A., Fahrmeir, L., Hennerfeind, A., 2007. Adaptive Gaussian Markov random fields with applications in human brain mapping. Appl Stat 56 (3), 327–345.
Brown, A.D., McMahan, C.S., Watson, S.C., 2021. Sampling strategies for fast updating of Gaussian Markov random fields. Amer. Statist. 75 (1), 52–65.
Cameletti, M., Gómez-Rubio, V., Blangiardo, M., 2019. Bayesian modelling for spatially misaligned health and air pollution data through the INLA-SPDE approach. Spatial Stat 31, 2211–6753.
Carlin, B.P., Banerjee, S., 2003. Hierarchical multidivariate CAR models for spatio-temporally correlated survival data (with discussion). In: Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P., Heckerman, D., Smith, A.E.M., West, M. (Eds.), Bayesian Statistics. vol. 7, Oxford University Press, Oxford, pp. 45–63.
Carlin, B.P., Gelfand, A.E., 1990. Approaches for empirical Bayes confidence intervals. J. Amer. Statist. Assoc. 85 (409), 105–114.
Carlin, B.P., Gelfand, A.E., 1991. A sample reuse method for accurate parametric empirical Bayes confidence intervals. J. R. Statist. Soc. Ser. B Stat. Methodol. 53 (1), 189–200.
Carlin, B.P., Louis, T.A., 2000. Bayes and Empirical Bayes Methods for Data Analysis. Chapman and Hall/CRC.
Carlin, B.P., Louis, T.A., 2000b. Empirical Bayes: Past. Present and future. J. Amer. Statist. Assoc. 95 (452), 1286–1289.
Clayton, D., Bernardinelli, L., 1992. Bayesian methods for mapping disease risk. In: Elliott, P., Cuzick, J., English, D., Stern, R. (Eds.), Geographical and Environment Epidemiology: Methods for Small Area Studies. Oxford University Press, Oxford, 205220.
Clayton, D.G., Bernardinelli, L., Montomoli, C., 1993. Spatial correlation in ecological analysis. Int. J. Epidemiol. 22 (6), 1193–1202.
Clayton, D.G., Kaldor, J., 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics 43, 671–681.
Congdon, P., 2001. Bayesian Statistical Modelling. Wiley.
Congdon, P., 2008. A spatially adaptive conditional autoregressive prior for area health data. Stat. Methodol. 5 (6), 3172–3127.
Corpas-Burgos, F., Martinez-Beneito, M.A., 2020. On the use of adaptive spatial weight matrices from disease mapping multivariate analyses. Stoch Environ Res Risk Assess 34, 531–544.
Cressie, N., 1993. Statistics for Spatial Data (Revised Edition). Wiley, New York.
Cressie, N., Wikle, C.K., 2015. Statistics for Spatio-Temporal Data. Wiley, New York.
Datta, A., Banerjee, S., Hodges, J.S., Gao, L., 2019. Spatial disease mapping using directed acyclic graph auto-regressive (DAGAR) models. Bayesian Anal. 14 (4), 1221–1244.
Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B Stat. Methodol. 39 (1), 1–22.
Denison, D.G.T., Holmes, C.C., 2001. Bayesian partitioning for estimation of disease risks. Biometrics 57, 143–149.
Etteberrija, J., Ugrate, M.D., Goicoa, T., Militino, A.F., 2015. On predicting cancer mortality using ANOVA-type P-spline models. REVSTAT-Statistical J 13 (1), 21–40.
Feng, C., 2021. Spatial–temporal generalized additive model for modeling COVID-19 mortality risk in Toronto, Canada. Spatial Stat.
Fernandez, C., Green, P.J., 2002. Modelling spatially correlated data via mixture: A Bayesian approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 64 (4), 805–826.
Gao, H., Bradley, J.R., 2019. Bayesian analysis of areal data with unknown adjacencies using the stochastic edge mixed effects model. Spatial Stat 31, 2211–6753.
Gao, L., Datta, A., Banerjee, S., 2021. Multivariate directed acyclic graph auto-regressive (MDAGAR) models for spatial diseases mapping. In: UCLA: Biostatistics. Retrieved from https://escholarship.org/uc/item/88c7t942.
Gelfand, A.E., Schmidt, A.M., Banerjee, S., Sirmans, C.F., 2004. Nonstationary multivariate process modelling through spatially varying coregionalization (with discussion). Test 13, 263–312.
Gelfand, A.E., Smith, A.F.M., 1990. Sampling-based approaches to calculating marginal densities. J. Amer. Statist. Assoc. 85 (410), 398–409.
Gelfand, A.E., Vounatsou, P., 2003. Proper multivariate conditional autoregressive models for spatial data analysis. Biostatistics 4, 11–25.
Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., 1995. Bayesian Data Analysis. Chapman and Hall/CRC.
Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1998. Markov Chain Monte Carlo in Practice. Chapman and Hall/CRC.
Girolami, M., Calderhead, B., 2011. Riemann manifold Langevin and Hamiltonian Monte Carlo methods (with discussions). J. R. Stat. Soc. Ser. B Stat. Methodol. 73, 123–214.
Goicoa, T., Ugrate, M.D., Etteberrija, J., Militino, A.F., 2016. Age–space–time CAR models in Bayesian disease mapping. Stat. Med. 30, 2391–2405.
Greco, F.P., Trivisano, C., 2009. A multivariate CAR model for improving the estimation of relative risks. Stat. Med. 28, 1707–1224.

Green, P.J., Richardson, S., 2002. Hidden Markov models and disease mapping. J. Amer. Statist. Assoc. 97 (460), 1–16.

Haviland, A., 1871. The geographical distribution of disease in England and Wales. The British Med J 7, 5–6.

Haviland, A., 1888. The geographical distribution of cancerous disease in the British isles. Lancet March 3, 412–414.

Held, L., 2000. Bayesian modelling of inseparable space–time variation in disease risk. Stat. Med. 19 (17–18), 2555–2567.

Held, L., Besag, J., 1998. Modelling risk from a disease in time and space. Stat. Med. 17 (18), 2045–2060.

Held, L., Best, N.G., 2001. A shared component model for detecting joint and selective clustering of two diseases. J. R. Stat. Soc. Ser. C. Appl. Stat. 164 (1), 73–85.

Hoffman, M.D., Radul, A., Sountsov, P., 2021. An Adaptive MCMC Scheme for Setting Trajectory Lengths in Hamiltonian Monte Carlo. In: Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS) 2021, San Diego, California, USA. PMLR: Volume 130.

Huang, G., Patrick, E., Brown, P.E., 2021. Population-weighted exposure to air pollution and COVID-19 incidence in Germany. Spatial Stat 41, 6753.

Hughes, J., 2015. Copcar: A exible regression model for areal data. J. Comput. Graph. Statist. 24, 733–755.

Hughes, J., Hanan, M., 2013. Dimension reduction and alleviation of confounding for spatial generalized linear mixed effects models. J. the R Stat Soc B (Methodology) 75 (1), 139–159.

Jack, E., Lee, D., Dean, N., 2019. Estimating the changing nature of Scotland’s health inequalities by using a multivariate spatiotemporal model. J. the R Stat Soc. Ser A, (Statistics in Society) 182 (3), 1061–1080.

Jin, X., Carlin, B.P., Banerjee, S., 2007. Order–free co-regionalized areal data models with application to multiple-disease mapping. J. R. Stat. Soc. Ser. B Stat. Methodol. 69 (5), 817–838.

Kim, H., Sun, D., Tsutakawa, R.K., 2001. A bivariate Bayes method for improving the estimates of mortality rates with a twofold conditional autoregressive model. J. Amer. Statist. Assoc. 96 (456), 1506–1521.

Lawson, A.B., 2018. Bayesian Disease Mapping Hierarchical Modeling in Spatial Epidemiology (Third Edition). Chapman and Hall/CRC.

Lawson, A., Biggeri, A., Bohning, D., Lesaffre, E., Viel, J.-F., Bertolli, R., 1999. Disease Mapping and Risk Assessment for Public Health. Wiley.

Lee, D., 2011. A comparison of conditional autoregressive models used in Bayesian disease mapping. Spatial and Spatio-Temporal Epidemiol 2 (2), 79–89.

Lee, D., 2013. CARBayes: an R package for Bayesian spatial modeling with conditional autoregressive priors. J. Stat. Softw. 55 (13), 1–24.

Lee, D.J., Durbán, M., 2011. P-spline ANOVA-type interaction models for spatio-temporal smoothing. Stat Modell 11 (1), 49–69.

Lee, D., Lawson, A., 2016. Quantifying the spatial inequality and temporal trends in maternal smoking rates in glasgow. Ann Appl Stat 10 (3), 1427–1446.

Lee, D., Mitchell, R., 2012. Boundary detection in disease mapping studies. Biostatistics 13 (3), 415–426.

Lee, D., Robertson, C., Marques, C., 2021. Quantifying the small-area spatio-temporal dynamics of the Covid-19 pandemic in Scotland during a period with limited testing capacity. Spatial Stat.

Lee, D., Rushworth, A., Sahu, S.K., 2014. A Bayesian localized conditional autoregressive model for estimating the health effects of air pollution. Biometrics 70 (2), 419–429.

Lee, D., Sarran, C., 2015. Controlling for unmeasured confounding and spatial misalignment in long-term air pollution and health studies. Environmetrics 26 (2), 477–487.

Leroux, B.G., Lei, X., Breslow, N., 1999. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: Halloran, M.E., Berry, D. (Eds.), Statistical Models in Epidemiology, the Environment and Clinical Trials. Springer, New York, 135178.

Li, X., Dey, D.K., 2021. Estimation of COVID-19 mortality in the United States using spatio-temporal Conway maxwell Poisson model. Spatial Stat.

Liang, Y., 2019. Graph-based multivariate conditional autoregressive models. Stat Theory and Related Fields 3 (2), 158–169.

Lindgren, F., Rue, H., Lindstrom, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 73 (4), 423–498.

Lu, H., Carlin, B.P., 2005. Bayesian area wombling for geographical boundary analysis. Geograph Anal 37 (3), 265–285.

Lu, H., Reilly, C.S., Banerjee, S., Carlin, B.P., 2007. Bayesian areal Wombling via adjacency modelling. Environ Ecol Stat 14, 433–452.

MacNab, Y.C., 2003a. A Bayesian hierarchical model for accident and injury surveillance. Accid. Anal. Prev. 35 (1), 91–102.

MacNab, Y.C., 2003b. Hierarchical Bayesian spatial modelling of small-area rates of non-rare disease. Stat. Med. 22 (10), 1761–1773.

MacNab, Y.C., 2003c. Hierarchical Bayesian modelling of spatially correlated health service outcome and utilization rates. Biometrics 59 (2), 305–316.

MacNab, Y.C., 2004. Bayesian spatial and ecological models for small-area accident and injury analysis. Accid. Anal. Prev. 36 (6), 1019–1028.

MacNab, Y.C., 2007. Spline smoothing in Bayesian disease mapping. Environmetrics 18 (7), 727–744.

MacNab, Y.C., 2009. Bayesian multivariate disease mapping and ecological regression with errors in covariates. Stat. Med. 28 (9), 1369–1385.

MacNab, Y.C., 2010. On Bayesian shared component disease mapping and ecological regression with errors in covariates. Stat. Med. 29 (11), 1239–1249.
MacNab, Y.C., 2011. On Gaussian Markov random fields and Bayesian disease mapping. Stat Methods Med Res 20 (1), 49–68.
MacNab, Y.C., 2014. On identification in Bayesian disease mapping and ecological-spatial regression. Stat Methods Med Res 23 (2), 134–155.
MacNab, Y.C., 2016a. Linear models of coregionalization for multivariate lattice data: a general framework for coregionalized multivariate CAR models. Stat. Med. 35, 3827–3850.
MacNab, Y.C., 2016b. Linear models of coregionalization for multivariate lattice data: Order-dependent and order-free MCARs. Statist Methods Med Res 25 (4), 1118–1144.
MacNab, Y.C., 2018. Some recent work on multivariate Gaussian Markov random fields (with discussions). TEST 27 (3), 497–541.
MacNab, Y.C., 2020. Bayesian estimation of multivariate Gaussian Markov random fields with constraint. Stat. Med. 39 (30), 4767–4788.
MacNab, Y.C., 2021. On coregionalized multivariate Gaussian Markov random fields - Part I: Construction and parameterization. Manuscript under peer-review..
MacNab, Y.C., 2021. On coregionalized multivariate Gaussian Markov random fields – Part II: Bayesian estimation and inference, with simulation and case studies. Manuscript under peer-review.
MacNab, Y.C., Dean, C.B., 2001. Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. Biometrics 57 (3), 949–956.
MacNab, Y.C., Farrell, P., Gustafson, P., Wen, S., 2004. Estimation in Bayesian disease mapping. Biometrics 60, 865–873.
MacNab, Y.C., Gustafson, P., 2007. Regression B-spline smoothing in Bayesian disease mapping: with an application to patient safety surveillance. Stat. Med. 26 (24), 4455–4474.
MacNab, Y.C., Kmetic, A., Gustafson, P., Shaps, S., 2006. An innovative application of Bayesian disease mapping methods to patient safety research: the Canadian iatrogenic injury study. Stat. Med. 25 (23), 3960–3980.
MacNab, Y.C., Lin, Y., 2009. On empirical Bayes penalized quasi-likelihood inference in GLMMs and in disease mapping and ecological modeling. Comput. Statist. Data Anal. 53 (8), 2950–2967.
Manton, K.G., Woodbury, M.A., Stallard, E., Riggan, W.B., Creason, J.P., Pellom, A.C., 1989. Empirical Bayes procedures for stabilizing maps of u.s. cancer mortality rates. J. Amer. Statist. Assoc. 84 (407), 637–650.
Marcotte, D., Allard, D., 2018. Gibbs sampling on large lattice with GMRF. Comput. Geosci. 111, 190–199.
Mardia, K.V., 1988. Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. J. Multivariate Anal. 24, 265–284.
Martinez-Beneito, M.A., 2013. A general modeling framework for multivariate disease mapping. Biometrika 100 (3), 539–553.
Martinez-Beneito, M.A., 2020. Some link between conditional and coregionalized multivariate Gaussian Markov random fields. Spatial Stat 40, 100383.
Martinez-Beneito, M.A., Botella-Rocamora, P., 2019. Disease Mapping: From Foundations To Multidimensional Modeling. CRC Press.
Martinez-Beneito, M.A., Botella-Rocamora, P., Barerjee, S., 2017. Towards a multidimensional approach to Bayesian disease mapping. Bayesian Anal. 12 (1), 239–259.
Martuzzi, M., Elliott, P., 1996. Empirical Bayes estimation of small area prevalence of non-rare conditions. Stat. Med. 15 (17), 1867–1873.
McCullagh, P., Nelder, J.A., 1983. Generalized Linear Models (Second Edition). Chapman and Hall.
Moraga, P., Cramb, S.M., Mengersen, K.L., et al., 2017. A geostatistical model for combined analysis of point-level and area-level data using INLA and SPDE. Spatial Stat 21, 27–41.
Nathoo, S., Ghosh, P., 2012. Skew-elliptical spatial random effect modeling for areal data with application to mapping health utilization rates. Stat. Med. 32 (2), 290–306.
Orozco-Acosta, E., Adin, A., Ugarte, M.D., 2021. Scalable Bayesian modelling for smoothing disease risks in large spatial data sets using INLA. Spatial Stat 41, 2211–6753.
Paciorek, C.J., 2010. The importance of scale for spatial-confounding bias and precision of spatial regression estimators. Statist. Sci. 25 (1), 107–125.
Prates, M.O., Assuncao, R.M., Rodrigues, E.C., 2019. Alleviating spatial confounding for areal data problems by displacing the geographical centroids. Bayesian Anal. 14, 623–647.
Prates, M., Azevedo, D., MacNab, Y.C., M., Willig, 2021. Non-separable spatio-temporal models via transformed Gaussian Markov random fields. Under peer review.
Prates, M.O., Dey, D.K., Willing, M.R., Yan, J., 2015. Transformed Gaussian Markov random fields and spatial modeling of species abundance. Spatial Stat 14, 382–399.
Reich, B.J., Hodges, J.S., 2008. Modeling longitudinal periodontal data: A spatially-adaptive model with tools for specifying priors and checking fit. Biometrics 64, 790–799.
Reich, B.J., Hodges, J.S., Zadnik, V., 2006. Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models. Biometrics 62 (4), 1197–1206.
Richardson, S., Thomson, T., Best, N., Elliott, P., 2004. Interpreting posterior relative risk estimates in disease-mapping studies. Environ Health Perspect 112 (9), 1016–1025.
Rue, H., Held, L., 2005. Gaussian Markov Random Fields - Theory and Applications. Chapman & Hall, New York.
Rue, H., Riebler, A., Sbye, S.H., Illian, J.B., Simpson, D.P., Lindgren, F.K., 2017. Bayesian computing with INLA: A review. Annu. Rev. Stat. Appl. 4 (1), 395–421.
Rushworth, A., Lee, D., Sarran, C., 2017. An adaptive spatiotemporal smoothing model for estimating trends and step changes in disease risk. Appl. Stat. 66 (1), 141–157.
Sahu, S.K., Böhning, D., 2021. Bayesian spatio-temporal joint disease mapping of Covid-19 cases and deaths in local authorities of England. Spatial Stat.
Sain, S.R., Cressie, N.A.C., 2007. A spatial analysis of multivariate lattice data. J. Econometrics 140, 226–259.
Sain, S.R., Furrer, R., Cressie, N.A.C., 2011. A spatial analysis of multivariate output from regional climate models. Ann. Appl. Stat. 5 (1), 150–175.
Slater, J.J., Brown, P.E., Rosenthal, J.S., Mateu, J., 2021. Capturing spatial dependence of COVID-19 case counts with cellphone mobility data. Spatial Stat In press.
Smith, A.F.M., Gelfand, A.E., 1992. Bayesian statistics without tears: a sampling–resampling perspective. Amer. Statist. 46 (2), 84–88.
Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). J. R. Stat. Soc. Ser. B Stat. Methodol. 64 (4), 583–639.
Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2003. WinBUGS User manual.
Sun, D., Tsutakawa, R.K., Speckman, P.L., 1999. Posterior distribution of hierarchical models using CAR(1) distributions. Biometrika 86, 341–350.
Tan, L.S.L., 2020. Use of model reparametrization to improve variational Bayes. J. R. Stat. Soc. Ser. B Stat. Methodol. 83 (1), 30–57.
Thomas, A., Best, N., Lunn, D., Arnold, R., Spiegelhalter, D., 2004. GeoBUGS User manual.
Tsutakawa, R.K., 1988. Mixed model for analyzing geographic variability in disease rates. J. Amer. Statist. Assoc. 83, 37–42.
Tsutakawa, R.K., Shoop, G.L., Marienfeld, C.J., 1985. Empirical bayes estimation of cancer mortality rates. Stat. Med. 4 (2), 201–212.
Tzala, E., Best, N., 2008. Bayesian latent variable modelling of multivariate spatio-temporal variation in cancer mortality. Stat. Methods Med. Res. 17 (1), 97–118.
Ugarte, M.D., Adin, A., Goicoa, T., 2017. One-dimensional, two-dimensional, and three dimensional B-splines to specify space–time interactions in Bayesian disease mapping: Model fitting and model identifiability. Spatial Stat 22 (2), 451–468.
Ugarte, M.D., Goicoa, T., Ibanez, B., Militino, A.F., 2009. Evaluating the performance of spatio-temporal Bayesian models in disease mapping. Environmetrics 20 (6), 647–665.
Ugarte, M.D., Goicoa, T., Militino, A.F., 2010. Spatio-temporal modeling of mortality risks using penalized splines. Environmetrics 21 (3–4), 270–289.
Utazi, C.E., Thorley, J., Alegana, V.A., Ferrari, M.J., Nilsen, K., Takahashi, S., Metcalf, C.J.E., Lessler, J., Tatem, A.J., 2018. A spatial regression model for the disaggregation of areal unit based data to high-resolution grids with application to vaccination coverage mapping. Stat. Methods Med. Res. 28 (10–11), 3226–3241.
Vicente, G., Goicoa, T., Ugarte, M.D., 2020. Bayesian inference in multivariate spatio-temporal areal models using INLA: analysis of gender-based violence in small areas. Stochastic Environ Res Risk Assess 34, 1421–1440.
Vicente, G., Goicoa, T., Ugarte, M.D., 2021. Multivariate Bayesian spatio-temporal P-spline models to analyze crimes against women. Biostatistics kxab042.
Wakefield, J., Morrise, S., 1999. Spatial dependence and errors-in-variables in environmental epidemiology. Bayesian Statistics 6, 657–684.
Wakefield, J., 2007. Disease mapping and spatial regression with count data. Biostatistics 8 (2), 158–183.
Wakefield, J., Salway, R., 2001. A statistical framework for ecological and aggregate studies. J. the R Statistical Soc: Ser A (Statistics in Society) 164, 119–137.
Waller, L.A., Carlin, B.P., Xia, H., Gelfand, A.E., 1997. Hierarchical spatio-temporal mapping of disease rates. J. Amer. Statist. Assoc. 92 (438), 607–617.
Waller, L.A., Gotway, G.A., 2004. Applied Spatial Statistics for Public Health Data. Wiley.
Zhang, C., Shahbaba, B., Zhao, H., 2018. Variational Hamiltonian Monte Carlo via score matching. Bayesian Anal. 13 (2), 485–506.
Zimmerman, D.L., Hoef, J.M.V., 2021. On deconfounding spatial confounding in linear models. Amer. Statist. http://dx.doi.org/10.1080/00031305.2021.1946149.