

Non-separable spatio-temporal models via transformed multivariate Gaussian Markov random fields

Marcos O. Prates¹  | Douglas R. M. Azevedo² |
Ying C. MacNab³ | Michael R. Willig⁴

¹Department of Statistics, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

²Apsilon, Warsaw, Poland

³School of Population and Public Health, The University of British Columbia, Vancouver, British Columbia, Canada

⁴Department of Ecology & Evolutionary Biology, Center for Environmental Sciences & Engineering and Institute of the Environment, University of Connecticut, Storrs, Connecticut, USA

Correspondence

Marcos O. Prates, Department of Statistics, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil.

Email: marcosop@est.ufmg.br

Abstract

Models that capture spatial and temporal dynamics are applicable in many scientific fields. Non-separable spatio-temporal models were introduced in the literature to capture these dynamics. However, these models are generally complicated in construction and interpretation. We introduce a class of non-separable transformed multivariate Gaussian Markov random fields (TMGMRF) in which the dependence structure is flexible and facilitates simple interpretations concerning spatial, temporal and spatio-temporal parameters. Moreover, TMGMRF models have the advantage of allowing specialists to define any desired marginal distribution in model construction without suffering from spatio-temporal confounding. Consequently, the use of spatio-temporal models under the TMGMRF framework leads to a new class of general models, such as spatio-temporal Gamma random fields, that can be directly used to model Poisson intensity for space-time data. The proposed model was applied to identify important environmental characteristics that affect variation in the abundance of *Nenia tridens*, a dominant species of gastropod in a well-studied tropical ecosystem, and to characterize its spatial and temporal trends, which are particularly critical during the Anthropocene, an epoch

of time characterized by human-induced environmental change associated with climate and land use.

KEYWORDS

Bayesian method, generalized linear mixed model, MCMC, spatial confounding, TGMRF, TMGMRF

1 | INTRODUCTION

In many fields of science, spatio-temporal models are useful to better understand and more realistically represent the dynamics of systems. This is particularly true for ecological systems during the Anthropocene (Steffen et al., 2007; Zalasiewicz et al., 2010), a time of rapid, human-induced environmental change linked to climate and land use. Ecological systems (suites of species that co-occur in time and space, and that interact with each other, as well as with matter and energy, to form systems) are complex, involving dynamics associated with abiotic (e.g. temperature, precipitation) and biotic (e.g. land use composition and configuration) characteristics. Because the Anthropocene is characterized by unprecedented rates of change, it is important to understand and predict spatio-temporal dynamics of populations that can inform management and policy with the ultimate goal of reducing the likelihood of species extinction and consequent loss of ecosystem services that are essential for human well-being. The urgency of the situation is reflected in recent suggestions that the planet is now entering its sixth major extinction period as well as in the controversy surrounding the announcement of biological Armageddon (Lister & Garcia, 2018; Schowalter et al., 2019; Willig et al., 2019).

Generalized linear mixed models (GLMMs; Breslow & Clayton, 1993) represent a flexible class of models that are capable of accommodating random effects in a simple manner. In this class of models, it is customary to choose an appropriate link function to model the conditional mean with covariates and random effects. Transformed Gaussian Markov random fields (TGMRF; Prates et al., 2015) appear as an effective tool for modelling spatial data. In TGMRFs, it is possible to directly choose the distribution of the conditional mean, including covariates, and to define the desired spatial structure. Unlike the traditional structure of spatial GLMMs that typically defines an appropriate link function and then model sources of outcome variation via the link function, which in some models can make interpretation difficult, TGMRFs allow direct inclusion of different sources of variation directly into the mean, rather than a function of it.

Spatial confounding (Reich et al., 2006) has gained attention because it can bias or inflate the variance of fixed effects estimates, making a significant factor appear to be non-significant or to reverse the conclusion about the covariate effect (Azevedo, Prates, et al., 2020). Recently, solutions to spatial confounding have appeared from multiple perspectives. The most common solution is to alleviate spatial confounding by model reparametrization (Azevedo, Bandyopadhyay, et al., 2020a; Azevedo, et al., 2021; Hanks et al., 2015; Hughes & Haran, 2013; Prates et al., 2019; Thaden & Kneib, 2018). Another venue to remove spatial confounding, which is relatively unexplored, involves separation of marginal distributions and dependencies (Hughes, 2015; Prates et al., 2015). More recently, Azevedo, Prates, et al. (2020) studied and proposed a solution for spatial confounding in misaligned models, where the spatial structure of the random effects are not the same as the spatial structure of the covariates.

A simple way to include spatial dependence in statistical models is to use spatially structured random effects. For areal data, the most common spatial structure is the conditional

autoregressive model (CAR; Besag, 1974). Although CAR models are useful for fitting spatial data, their structure is not directly applied to multivariate problems. Multivariate conditional autoregressive models (MCARs; Carlin & Banerjee, 2003; Gelfand & Vounatsou, 2003; Jin et al., 2005, 2007) were proposed to extend CAR models when multiple variables are observed in the same space. The idea is to control for the correlation structure between variables. Sain et al. (2011), Rodrigues (2012) and MacNab (2018) have presented alternatives to define the cross-correlation between regions and variables.

In this paper, we propose a non-separable, flexible and interpretable spatio-temporal dependence structure and an extension of TGMRFs to multivariate problems. This new formulation facilitates a clear and direct interpretation of the contributions of spatial, temporal and spatio-temporal components. In addition, the proposed model prevents spatio-temporal confounding via a copula structure that guarantees by construction, the separation of fixed and random effects. This is an advantage because, to the best of our knowledge, little is known about the extent to which spatio-temporal random effects might confound fixed effects estimates (Adin et al., 2021).

We leverage a long-term (17 years) ecological study (Bloch & Willig, 2006; Willig et al., 1998, 2007, 2014) to illustrate the utility of our multivariate TGMRF approach. More specifically, we construct and interpret spatio-temporal models for counts of *Nenia tridens*, an abundant species, that dominates the gastropod fauna in forests of Puerto Rico. This is particularly relevant because these ecosystems are disturbance-mediated: the mapping of environmental characteristics onto geographic space changes over time in response to climatic events (e.g. cyclonic storms and droughts) and subsequent secondary succession, with consequences to the abundance and distribution of resident species (Willig et al., 2021). Fortunately, spatially explicit data are available for counts of species as well as for habitat characteristics that are known to influence abundance over time.

Section 2 highlights the ecological relevance and importance of the data. Section 3 summarizes several multivariate dependence structures in the literature. The TMGMRF formulation for spatio-temporal settings and how inference is performed are presented in Section 4. A detailed simulation study about the proposed method appears in Section 5. Section 6 revisits the ecological application showing the empirical and modelled results. A final conclusion and discussion are presented in Section 7.

2 | ECOLOGICAL CHARACTERISTICS

Gastropods (snails and slugs) are the second most species-rich group of animals in the world (Prié, 2019). They are ubiquitous heterotrophs (decomposers) and provide essential ecosystem functions associated with energy flow and nutrient cycling (Prather et al., 2013). Previous research has documented their habitat associations and responses to disturbances such as tree-fall gaps (Alvarez & Willig, 1993), hurricanes (Prates et al., 2011; Willig & Camilo, 1991), and previous land-use history (Willig et al., 1998) in the Luquillo Experimental Forest of Puerto Rico. Thus, gastropods in the Luquillo Experimental Forest represent an ideal system to explore as an illustrative case for modelling spatio-temporal demographics within a dynamic environmental context. Moreover, *Nenia tridens* is one of the most numerically dominant gastropods in tabonuco forest on the Luquillo Forest Dynamics Plot (LFDP), and has a heterogeneous spatial distribution, making it of particular ecological importance (Bloch & Willig, 2006; Willig et al., 1998).

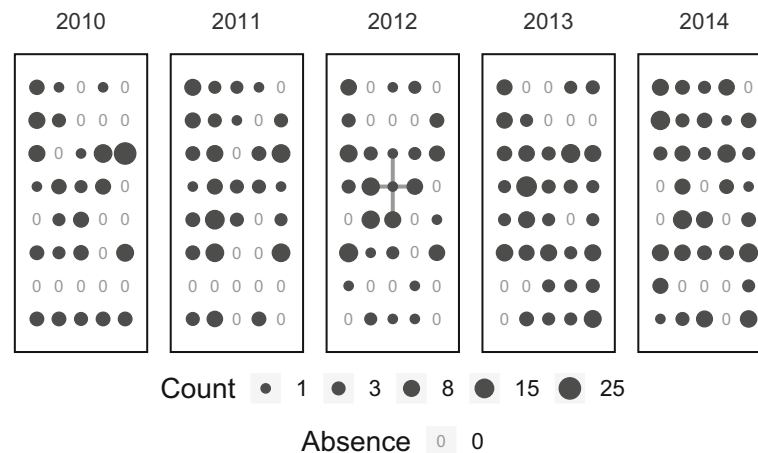


FIGURE 1 Graphic representation of the spatial distribution of counts of *Nenia tridens* and on the sampling lattice for each of 5 illustrative years. Circles represent counts and grey lines represent the neighbourhood structure adopted.

Between 2000 and 2017, data on counts (minimum known alive) of *Nenia tridens* were quantified on the LFDP (Figure 1). The LFDP is a 16-ha rectilinear grid that comprises an 8×5 lattice of 40 points (circles of 3 m radius), with 60 m spacing between adjacent points (Willig et al., 1998). A suite of covariates characterized each of the 40 points and represent habitat characteristics. Some varied in space but not time: Elevation (metres above sea level) and slope (inclination of land in degrees). As a consequence of disturbance and succession, others varied in space and time: density of vegetation or FolDenAll (foliar intercepts by plant, regardless of species identity, in the understorey), density of Sierra Palm or FolDenPa (foliar intercepts by *Prestoea acuminata* in the understorey), litter cover (ordinal representation of amount of litter on the forest floor, from 0 to 2) and canopy openness (estimate of penetration of light to forest understorey). To avoid computational problems, all covariates were centred and scaled so that interpretations involve deviations from the mean. In Section 6, our methodology is applied to this data set and the space–time parameters are interpreted from an environmental perspective.

3 | DEPENDENCE STRUCTURE

For areal data, when observations represent a well-defined region, the most traditional model used to capture spatial dependence is the CAR model. In a multivariate or spatio-temporal context, there are a variety of ways to formulating models; however, the approaches of Mardia (1988) and Sain et al. (2011) are the most natural extensions of the univariate CARs (see MacNab, 2018, for a recent review).

3.1 | Multivariate conditional autoregressive

A direct extension of the CAR model occurs when more than one dependent variable is observed over the same region. This family of multivariate models is known as MCAR (Gelfand & Vounatsou, 2003).

Let n be the number of regions of interest and p the number of variables observed. Define $\mathbf{Y}_1$ and $\boldsymbol{\Theta}_1$ a vector of observations and spatial random effects, respectively, ordered by region, then

$$\begin{aligned}\mathbf{Y}_1 &= (Y_{11}, Y_{12}, \dots, Y_{1p}, Y_{21}, \dots, Y_{2p}, \dots, Y_{n1}, \dots, Y_{np}), \\ \boldsymbol{\Theta}_1 &= (\theta_{11}, \theta_{12}, \dots, \theta_{1p}, \theta_{21}, \dots, \theta_{2p}, \dots, \theta_{np}).\end{aligned}$$

Now define $\mathbf{Y}_2$ an observation vector and $\boldsymbol{\Theta}_2$ a spatial random effect sorted by variable, thus

$$\begin{aligned}\mathbf{Y}_2 &= (Y_{11}, Y_{21}, \dots, Y_{n1}, Y_{12}, \dots, Y_{n2}, \dots, Y_{1p}, \dots, Y_{np}), \\ \boldsymbol{\Theta}_2 &= (\theta_{11}, \theta_{21}, \dots, \theta_{n1}, \theta_{12}, \dots, \theta_{n2}, \dots, \theta_{np}).\end{aligned}$$

A p -variate CAR can be defined by conditional distributions for θ_{ij} as

$$(\theta_{ij} | \boldsymbol{\theta}_{-ij}) \sim N_{np} \left(\sum_{ij \sim kl} b_{ij,kl} \theta_{kl}, \tau_{ij} \right),$$

where $\boldsymbol{\theta}_{-ij}$ represents the vector $\boldsymbol{\theta}$ without the ij entry and $ij \sim kl$ are defined as the neighbours of a variable j in region i with a variable l in region k . Applying Brook's Lemma, it is possible to calculate the joint distribution of $\boldsymbol{\Theta}$, whereas $\boldsymbol{\Theta}$ is either $\boldsymbol{\Theta}_1$ or $\boldsymbol{\Theta}_2$, as

$$\Pi(\boldsymbol{\Theta}) \propto \exp \left\{ -\frac{1}{2} \boldsymbol{\Theta}' \mathbf{Q} \boldsymbol{\Theta} \right\}, \quad (1)$$

where $\mathbf{Q}$ is a precision matrix and $q_{ij,kl} = \frac{-b_{ij,kl}}{\tau_{ij}}$. Therefore, as in the univariate case, it is necessary to guarantee that $\mathbf{Q}$ is symmetric and positive definite. Different choices of the coefficients $b_{ij,kl}$ and τ_{ij} determine the methodologies that are available from the literature.

Given the general representation of Equation (1), an alternative way of interpreting and understanding the multivariate distribution is considering its conditional mean

$$E(\theta_{ij} | \boldsymbol{\theta}_{-ij}) = \underbrace{\sum_{k \neq i} b_{ij,kj} \theta_{kj}}_{\mathbb{A}} + \underbrace{\sum_{l \neq j} b_{ij,il} \theta_{il}}_{\mathbb{B}} + \underbrace{\sum_{(kl) \neq (ij)} b_{ij,kl} \theta_{kl}}_{\mathbb{C}}, \quad (2)$$

and conditional variance

$$\text{Var}(\theta_{ij} | \boldsymbol{\theta}_{-ij}) = \tau_{ij}^2, \quad (3)$$

This representation allows for a direct interpretation of the sums $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$ in Equation (2): $\mathbb{A}$ measures spatial dependence within the same variable, $\mathbb{B}$ measures the dependence between variables in the same region, and $\mathbb{C}$ measures the spatial dependence between different variables.

Next, we revisit the key MCAR proposals in the literature, each represents a specific parametrization of Equation (1) and its dependence interpretation represented by the conditional mean structure in Equation (2).

3.1.1 | Gelfand and Vounatsou (2003) and Carlin and Banerjee (2003)

Ordering the data by region, $\mathbf{Y}_1$, the authors parametrize the n -site and p -variable MCAR precision matrix $\mathbf{Q}_1$, as:

$$\mathbf{Q}_1 = (\mathbf{D}_w - \rho \mathbf{W}) \otimes \mathbf{\Lambda},$$

where $\mathbf{\Lambda}$ is a $p \times p$ inverse covariance matrix with elements Λ_{jl} , $\mathbf{W}$ is the CAR connectivity matrix with elements $w_{ii} = 0$, $w_{ik} = 1$ if $i \sim k$ or $w_{ik} = 0$ otherwise, $\mathbf{D}_w$ is a diagonal matrix with values w_{i+} with $w_{i+} = \sum_k w_{ik}$ from matrix $\mathbf{W}$, and ρ is a spatial parameter.

The MCAR precision matrix $\mathbf{Q}_1$ is a Kronecker product of a CAR (spatial) precision matrix and a (non-spatial) precision matrix of a p -variate Gaussian distribution. This structure is named separable MCAR, because its covariance matrix can be written as a product of a spatial and a non-spatial covariance matrix. A MCAR is said to be non-separable MCAR if its precision matrix cannot be expressed as a product of spatial and non-spatial precision matrix.

With this parametrization, the $b_{ij,kl}$ and τ_{ij} in Equations (2) and (3) are given by:

$$b_{ij,kl} = \begin{cases} \rho \frac{w_{ik}}{w_{i+}}, & \text{if } j = l \text{ and } i \neq k, \\ -\frac{\Lambda_{jl}}{\Lambda_{jj}}, & \text{if } j \neq l \text{ and } i = k, \text{ and} \\ \rho \frac{w_{ik}}{w_{i+}} \frac{\Lambda_{jl}}{\Lambda_{jj}}, & \text{if } j \neq l \text{ and } i \neq k, \end{cases}$$

$$\tau_{ij} = \frac{1}{w_{i+} \Lambda_{jj}}.$$

When the data are ordered by variable, $\mathbf{Y}_2$, we have the $\mathbf{Q}_2$ equivalent of $\mathbf{Q}_1$ as $\mathbf{Q}_2 = \mathbf{\Lambda} \otimes (\mathbf{D}_w - \rho \mathbf{W})$.

The above-mentioned parameterization for $b_{ij,kl}$ leads to the following conditional mean

$$E(\theta_{ij} | \theta_{-ij}) = \sum_{k \neq i} \rho \frac{w_{ik}}{w_{i+}} \theta_{kj} - \sum_{l \neq j} \frac{\Lambda_{jl}}{\Lambda_{jj}} \theta_{il} + \sum_{k, l \neq i, j} \rho \frac{w_{ik}}{w_{i+}} \frac{\Lambda_{jl}}{\Lambda_{jj}} \theta_{kl}. \quad (4)$$

The first summation represents within-variable spatial dependence, where ρ is a common spatial dependence parameter for each of the p variables. The second summation represents non-spatial dependence between variables at the same locations, while the third summation represents cross-spatial dependence between variables at different neighbouring locations. Notice that the previously seen spatial parameter ρ also appears in the third summation, where it serves as a common cross-spatial dependence parameter for any two of the p -variables. As noted and illustrated in MacNab (2018), this parametrization may not lead to intuitively appealing cross-spatial dependence interpretation. The negative sign in the second summation implies that the cross-spatial dependence between variables at different neighbouring locations, represented by the third summation, may be inconsistent with the non-spatial dependence between variables at the same locations. For example, let variables j and l be correlated positively. This leads to a negative Λ_{jl} in the precision matrix $\mathbf{\Lambda}$, and $\rho \Lambda_{jl} < 0$ in Equation (4) when $\rho > 0$, which suggests a positive spatial dependency of each variable, positive non-spatial dependency between variables j and l at

the same locations, but negative cross-spatial dependency between the two variables at different neighbouring locations.

3.1.2 | Jin et al. (2007)

The representation of Jin et al. (2007) has a restriction in order and can be sorted only by variable, thus the authors define the matrix $\mathbf{Q}_2$ as

$$\mathbf{Q}_2 = \begin{pmatrix} (\mathbf{D}_w - \gamma_{11}\mathbf{W})\Lambda_{11} & \dots & (\mathbf{D}_w - \gamma_{1p}\mathbf{W})\Lambda_{1p} \\ \vdots & \ddots & \vdots \\ (\mathbf{D}_w - \gamma_{1p}\mathbf{W})\Lambda_{1p} & \dots & (\mathbf{D}_w - \gamma_{pp}\mathbf{W})\Lambda_{pp} \end{pmatrix}, \quad (5)$$

where $\mathbf{\Gamma} = (\gamma_{jl}) = \mathbf{\Lambda}^{1/2} \boldsymbol{\rho} \mathbf{\Lambda}^{1/2}$, $\mathbf{\Lambda}^{1/2} (\mathbf{\Lambda}^{1/2})^\top = \mathbf{\Lambda} = (\Lambda_{jl})$, $\boldsymbol{\rho} = (\rho_{jl})$ is a p by p symmetric (or a diagonal) matrix of spatial dependence parameters (MacNab, 2018).

With this structure, we can find $b_{ij,kl}$ and τ_{ij} as

$$b_{ij,kl} = \begin{cases} \gamma_{jj} \frac{w_{ik}}{w_{i+}}, & \text{if } j = l \text{ and } i \neq k, \\ -\frac{\Lambda_{jl}}{\Lambda_{jj}}, & \text{if } j \neq l \text{ and } i = k, \\ \gamma_{jl} \frac{w_{ik}}{w_{i+}} \frac{\Lambda_{jl}}{\Lambda_{jj}}, & \text{if } j \neq l \text{ and } i \neq k, \end{cases}$$

$$\tau_{ij} = \frac{1}{w_{i+} \Lambda_{jj}}.$$

Like the previous formulation, this representation may also lead to inconsistent cross-spatial and non-spatial dependence (see MacNab, 2018, for illustrative examples). From $\mathbf{Q}_2$ in Equation (5) we have

$$E(\theta_{ij} | \boldsymbol{\theta}_{-ij}) = \sum_{k \neq i} \gamma_{jj} \frac{w_{ik}}{w_{i+}} \theta_{kj} - \sum_{l \neq j} \frac{\Lambda_{jl}}{\Lambda_{jj}} \theta_{il} + \sum_{k, l \neq i, j} \gamma_{jl} \frac{w_{ik}}{w_{i+}} \frac{\Lambda_{jl}}{\Lambda_{jj}} \theta_{kl},$$

which, when $\boldsymbol{\rho} = \rho \mathbf{I}_n$, leads to the Gelfand and Vounatsou (2003) and Carlin and Banerjee (2003) MCAR, also known as the separable MCAR. The Jin et al. (2007) parametrization is more flexible since it allows a symmetric (or diagonal) matrix of spatial dependence parameters instead of a unique spatial parameter ρ . Except when considering a single spatial parameter, the Jin et al. (2007) MCAR is a non-separable MCAR, because its precision matrix cannot be expressed as a product of a spatial and non-spatial precision matrix (MacNab, 2018).

3.1.3 | Sain et al. (2011)

Sain et al. (2011) proposed an alternative framework for MCAR formulation where three types of neighbourhoods were considered in characterizing the dependence structure of a model: (1) spatial neighbours of region i (Figure 2a), (2) neighbours of the same region i between variables (Figure 2b), (3) spatial neighbours of region i across different variables (Figure 2c).

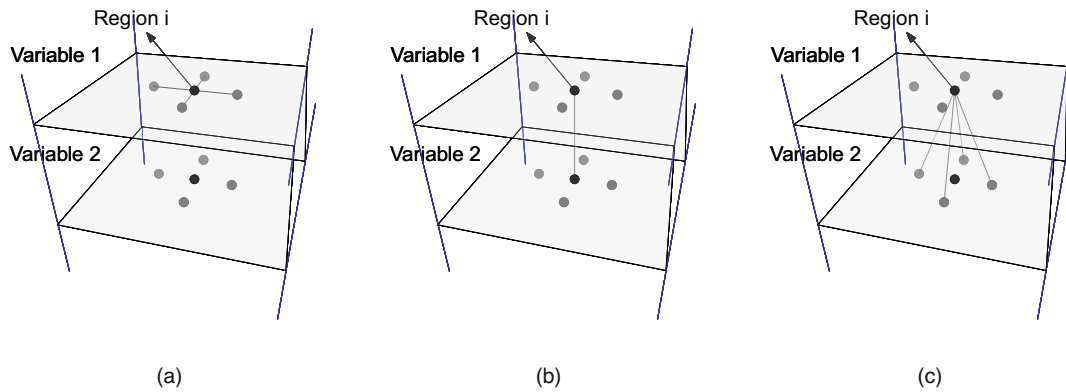


FIGURE 2 Types of neighbourhood. (a) Spatial neighbourhood within the variable. (b) Neighbourhood between the same region and different variables. (c) Neighbourhood between a particular region and its spatial neighbour regions across different variables. [Colour figure can be viewed at wileyonlinelibrary.com]

Let $\mathbf{C} = \mathbf{C}^\top$ be a symmetric matrix of spatial dependence parameters, the $\mathbf{Q}_1$ matrix of the Sain et al. MCAR is defined as

$$\mathbf{Q}_1 = (\mathbf{I}_n \otimes \Lambda^{-\frac{1}{2}})(\mathbf{I}_n \otimes \mathbf{A} - \mathbf{W} \otimes \mathbf{C})(\mathbf{I}_n \otimes \Lambda^{-\frac{1}{2}}), \quad (6)$$

where

$$\Lambda = \begin{pmatrix} \Lambda_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \Lambda_p^2 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 1 & \dots & -\rho_{1p}^c \\ \vdots & \ddots & \vdots \\ -\rho_{p1}^c & \dots & 1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} c_{11} & \dots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{p1} & \dots & c_{pp} \end{pmatrix},$$

where $\mathbf{A} = \mathbf{A}^\top$ is a partial correlation matrix, $\rho_{jl}^c, \forall j, l$ are non-spatial partial correlation parameters. Notice that the diagonal element c_{jj} in $\mathbf{C}$ controls the within-variable spatial dependence for the variable j and the off-diagonal elements c_{jl} regulates cross-spatial dependence between variables j and l .

For this parametrization, the $b_{ij,kl}$ and τ_{ij} are defined by

$$b_{ij,kl} = \begin{cases} c_{jj}w_{ik}, & \text{if } j = l \text{ and } i \neq k, \\ \rho_{jl}^c \frac{\Lambda_j}{\Lambda_l}, & \text{if } j \neq l \text{ and } i = k, \\ w_{ik}c_{jl} \frac{\Lambda_j}{\Lambda_l}, & \text{if } j \neq l \text{ and } i \neq k, \end{cases}$$

and $\tau_{ij} = \Lambda_j^2$ are variable-specific scale parameters. The equation for the conditional mean is now

$$E(\theta_{ij} | \boldsymbol{\theta}_{-ij}) = \sum_{k \neq i} c_{jj}w_{ik}\theta_{kj} + \sum_{l \neq j} \rho_{jl}^c \frac{\Lambda_j}{\Lambda_l} \theta_{il} + \sum_{k, l \neq i, j} c_{jl}w_{ik} \frac{\Lambda_j}{\Lambda_l} \theta_{kl}. \quad (7)$$

This representation has four salient differences compared to the previous two parametrizations: (a) the contribution of the second summation is positive; (b) each variable is characterized by different spatial parameters (c_{jj}) in the first summation; (c) different parameters (ρ_{jl}^c)

accommodate the dependence between variables at the same locations in the second summation; and (d) the spatial parameters (c_{jl}) in the third summation control for cross-spatial dependence between variables.

Although the model seems flexible, its interpretation is not trivial because the conditional mean depends on the (scale) parameters $\Lambda_j, \forall j$. Moreover, the summations in Equation (7) are not weighted, and as a consequence, the resulting precision matrix has complex positivity requirement on both the spatial and non-spatial dependence parameters (see MacNab, 2018, 2020, for details). Again, with the exception of having a single spatial parameter, the Sain et al. (2011) MCARs are non-separable MCARs (MacNab, 2018).

3.1.4 | MacNab (2018)

MacNab (2018) extends the Sain et al. (2011) MCAR by allowing a set of scaling factors, denoted $\mathbf{D}_m = \text{diag}(m_1, m_2, \dots, m_n)$, to be introduced to $\mathbf{Q}_1$:

$$\mathbf{Q}_1 = (\mathbf{I}_n \otimes \Lambda^{-\frac{1}{2}})(\mathbf{I}_n \otimes \mathbf{A} - \tilde{\mathbf{W}} \otimes \mathbf{C})(\mathbf{I}_n \otimes \Lambda^{-\frac{1}{2}}), \quad (8)$$

which leads to the following MCAR conditional mean

$$E(\theta_{ij}|\theta_{-ij}) = \sum_{k \neq i} c_{kj} \frac{w_{ik}}{m_i} \theta_{kj} + \sum_{l \neq j} \rho_{jl}^c \frac{\Lambda_j}{\Lambda_l m_i} \theta_{il} + \sum_{k, l \neq i, j} c_{jl} \frac{w_{ik}}{m_i} \frac{\Lambda_j}{\Lambda_l} \theta_{kl}. \quad (9)$$

and $\tau_{ij} = m_i \Lambda_j^2$, where $\tilde{\mathbf{W}} = \mathbf{D}_m^{-1/2} \mathbf{W} \mathbf{D}_m^{-1/2}$. Notice that the $\mathbf{Q}_1$ in Equation (8) is the $\mathbf{Q}_1$ in Equation (6) with the spatial connectivity matrix $\mathbf{W}$ being replaced by a spatial weight matrix $\tilde{\mathbf{W}}$.

When m_i is the total number of terms in Equation (9), which is also the total numbers of neighbours for site i of the variable j , the MCAR Equation (8) is named a p -fold CAR, and it is the generalization of the twofold CAR proposed in Kim et al. (2001). Unlike the MCAR (7), which is the MCAR (8) with $\mathbf{D}_m = \mathbf{I}_n$, sufficient positivity constraints are more readily available for MCAR (8) when $m_i \geq 2, \forall i$ (see MacNab, 2018, for details). Furthermore, notice that when $\Lambda_j = \Lambda, \forall j$, the conditional mean of the MCAR (8) is simplified to

$$E(\theta_{ij}|\theta_{-ij}) = \sum_{k \neq i} c_{kj} \frac{w_{ik}}{m_i} \theta_{kj} + \sum_{l \neq j} \rho_{jl}^c \frac{1}{m_i} \theta_{il} + \sum_{k, l \neq i, j} c_{jl} \frac{w_{ik}}{m_i} \theta_{kl}, \quad (10)$$

where Expression (10) defines the multidimensional (spatial and non-spatial) dependence structure of the MCAR.

After defining the dependence structures for the multivariate case, it is clear that a spatio-temporal setup can be seen as an equivalent case where instead of having p variables in a map, we have one variable observed over the whole map in a discrete period of T times.

4 | SPATIO-TEMPORAL MODELLING USING TMGMRFs

In the spatial setup, the TGMRF (Prates et al., 2015) was proposed as a flexible alternative to GMRF (Rue & Held, 2005). In this class, the marginal distribution is chosen according to each application, providing flexibility in being capable of accommodating asymmetry, heavy tails or

other characteristics, thereby maintaining many desirable properties of the GMRF (Prates, 2011). For example, in a Poisson regression, using a gamma prior for modelling the marginal distributions of the relative risks is more flexible in comparison to use of the Gaussian prior (CARs or MCARs) for log relative risks. The shape parameter of the gamma prior facilitates modelling different degrees of skewness, whereas the Gaussian prior is limited in its capacity to model skewness. Furthermore, the TGMRF uses a copula approach to separate the marginal structure of the model from the dependent one. This is also of interest because, by construction, it modulates these two aspects of modelling. Consequently, TGMRFs do not suffer from confounding between fixed and random effects.

A TGMRF is obtained by transforming the marginal distribution of the GMRFs to a desired one. Let $\epsilon = (\epsilon_1, \dots, \epsilon_n)'$ be a multivariate normal vector with mean $\mathbf{0}$ and sparse correlation matrix Ξ , $\epsilon \sim N_n(\mathbf{0}, \Xi)$, consequently ϵ is a GMRF. Let $\mathbf{Z} = (Z_1, \dots, Z_n)'$ and $Z_i = F_i^{-1}\{\Phi(\epsilon_i)\}$, $i = 1, \dots, n$, where $F_i(x)$ is the cumulative distribution function (cdf) of an absolutely continuous function with respect to the support of x and Φ is the cdf of the $N(0, 1)$. So, each Z_i has marginal distribution f_i (probability density distribution [pdf] of F_i) and jointly a TGMRF with marginals $\mathbf{F}$ and correlation structure Ξ , denoted by $\mathbf{Z} \sim \text{TGMRF}_n(\mathbf{F}, \Xi)$. The $\mathbf{Q} = \Xi^{-1}$ brings a more intuitive interpretation of the conditionals distribution of $\mathbf{Z}$, and we parametrize the TGMRF by its precision matrix $\mathbf{Q}$ and denoted by $\mathbf{Z} \sim \text{TGMRF}_n(\mathbf{F}, \mathbf{Q})$.

TGMRFs can be used to directly model Poisson intensities or Bernoulli rates, taking into account a marginal distribution of interest and spatial dependence (Prates et al., 2015). For example, let $\mu_{n \times 1}$ be vector of the mean parameters in a Poisson regression, the TGMRF is defined as a joint distribution for μ as

$$\mu \sim \text{TGMRF}_n(\mathbf{F}, \mathbf{Q}),$$

where $\mathbf{F} = (F_1, \dots, F_n)$, F_i is a desired and adequate cdf for the marginal distribution of μ_i with pdf f_i and precision matrix $\mathbf{Q}$.

From a spatio-temporal perspective let $\mathbf{Y} = (\mathbf{Y}'_1, \mathbf{Y}'_2, \dots, \mathbf{Y}'_T)'$ be a random vector observed at T times and n regions with $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \dots, Y_{nt})'$ for $t = 1, \dots, T$. The $nT \times q$ covariate matrix is defined as $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_q)$ with $\mathbf{X}_j = (X_{11}, \dots, X_{n1}, X_{12}, \dots, X_{nT})'$ for $j = 1, \dots, q$ and random effects $\epsilon = (\epsilon'_1, \epsilon'_2, \dots, \epsilon'_T)'$ with $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \dots, \epsilon_{nt})'$ following a $N_{nt}(\mathbf{0}, \mathbf{Q})$.

If the distribution of the random variables Y_{it} belongs to the exponential family with mean $\mu_{it} = E(Y_{it} | \mathbf{X}, \epsilon_{it})$, then the joint distribution of μ can be modeled by a transformed multivariate GMRF (TMGMRF) as

$$\mu \sim \text{TMGMRF}_{nt}(\mathbf{F}, \mathbf{Q}),$$

where $\mathbf{F} = (F_{11}, \dots, F_{1T}, F_{21}, \dots, F_{nT})$, F_{it} is the cdf related to the marginal distribution of μ_{it} and $\mathbf{Q}$ is the precision matrix of μ .

Let $\xi = (\beta, \rho, \nu)$, where $\rho = (\rho_s, \rho_t, \rho_{st})$ are the spatial, temporal and spatio-temporal dependence parameters, respectively, and ν are hyperparameters of the distribution $\mathbf{F}$. A spatio-temporal hierarchical TMGMRF model can be defined as:

$$\begin{aligned} Y_{it} | \mu_{it} &\sim \pi(y | \mu_{it}), i = 1, \dots, n; t = 1, \dots, T, \\ \mu &\sim \text{TMGMRF}_{nt}(\mathbf{F}_{\xi, \mathbf{X}}, \mathbf{Q}_{\rho}), \\ \beta &\sim \pi(\beta), \\ \nu &\sim \pi(\nu), \\ \rho &\sim \pi(\rho), \end{aligned} \tag{11}$$

where $F_{\xi, \mathbf{X}}$ may depend on the covariates $\mathbf{X}$, regression coefficient vector $\boldsymbol{\beta}$, dispersion parameter(s) $\mathbf{v}$ and spatial, temporal and spatio-temporal parameters ρ_s , ρ_t and ρ_{st} , respectively. The precision matrix $\mathbf{Q}_\rho$ will depend on only the dependence parameters $\boldsymbol{\rho}$.

As previously emphasized, this formulation will not suffer from spatio-temporal confounding because it separates the marginal effects of the dependence structure. Moreover, it facilitates a flexible representation of marginals distributions. To avoid over parametrization and to construct a precision matrix capable of carrying the flexibility of model (11) combined with an intuitive parameter interpretation, we propose a simplification of the p -fold CAR, by incorporating MCAR (10) $c_{ij} \equiv \rho_s$, $\rho_{jl}^c \equiv \rho_t$ and $c_{jl} \equiv \rho_{st}$. With this parameterization, we have the conditional mean and variance defined as

$$\begin{aligned} E(\theta_{ij} | \boldsymbol{\theta}_{-ij}) &= \sum_{k \neq i} \rho_s \frac{w_{ik}}{m_i} \theta_{kj} + \sum_{l \neq j} \rho_t \frac{v_{jl}}{m_i} \theta_{il} + \sum_{k, l \neq i, j} \rho_{st} \frac{w_{ik} v_{jl}}{m_i} \theta_{kl}, \\ \text{Var}(\theta_{ij} | \boldsymbol{\theta}_{-ij}) &= \frac{\tau}{m_i}, \end{aligned} \quad (12)$$

where ρ_s accommodate the spatial dependence between regions, ρ_t represent the temporal dependence between time t and its previous ($t - 1$) and its next ($t + 1$), mimicking an autoregressive model in time with order 1 and ρ_{st} model the dependence between area i in time t and its spatial neighbours in time $t - 1$ and $t + 1$.

Notice that the conditional mean in Expression (12) defines the precision matrix $\mathbf{Q}_\rho$ in Expression (11), where

$$\mathbf{Q}_\rho = \mathbf{I}_n \otimes \mathbf{A}(\rho_t) - \tilde{\mathbf{W}} \otimes \mathbf{C}(\rho_s, \rho_{st}). \quad (13)$$

This MCAR proposal is non-separable because its precision matrix has two spatial dependence parameters (ρ_s and ρ_{st}) and cannot be expressed as product of a spatial and a temporal precision matrix.

4.1 | Marginal models and inference

When a traditional GLMM is used to fit a Poisson model, it is common to use the log-link function. It is easy to prove that under this link function the marginal distribution for the conditional mean is log-normal. Under TMGMRFs models, we can set the family, mean and variance of these distributions, to obtain the parametrization presented in Table 1.

An equivalent approach to the usual GLMM under log-link function is the log-normal (LN) model. Other distributions provide flexibility to the model. As can be seen, the gamma independent (GI), gamma scale (GSC) and gamma shape (GSH) models have different marginal variance functions. Depending on the type of the application, one function might be more plausible than the others. Additionally, with gamma priors on the areal-specific means, the regression part can be modelled on the areal-specific scale parameters (the GSC model), the areal-specific shape parameters (the GSH model) or as an areal mixture of the shape and scale parameters (the GI model). As a consequence, it is clear that the TMGMRF easily offers a variety of alternative models.

Although the traditional log-normal model (i.e. the model with Poisson likelihood, and MCAR prior on log means) has elements of the dependence structure in its marginal (last row of Table 1), it has a quadratic dependence in the variance with respect to the mean regressor, as in the

TABLE 1 The marginal models with their respective parametrizations, means and variances. Here, the Gamma distribution has expected value as a/b

Model	Model parametrization	$E(\mu_{ij})$	$V(\mu_{ij})$
Gamma independent (GI)	$\Gamma(\nu \exp\{2\mathbf{X}_{ij}\boldsymbol{\beta}\}, \nu \exp\{\mathbf{X}_{ij}\boldsymbol{\beta}\})$	$\exp\{\mathbf{X}_{ij}\boldsymbol{\beta}\}$	$\frac{1}{\nu}$
Gamma scale (GSC)	$\Gamma(\nu, \nu \exp\{-\mathbf{X}_{ij}\boldsymbol{\beta}\})$	$\exp\{\mathbf{X}_{ij}\boldsymbol{\beta}\}$	$\frac{1}{\nu} \exp\{\mathbf{X}_{ij}\boldsymbol{\beta}\}^2$
Gamma shape (GSH)	$\Gamma(\nu \exp\{\mathbf{X}_{ij}\boldsymbol{\beta}\}, \nu)$	$\exp\{\mathbf{X}_{ij}\boldsymbol{\beta}\}$	$\frac{1}{\nu} \exp\{\mathbf{X}_{ij}\boldsymbol{\beta}\}$
Log-normal (LN)	$LN\left(\mathbf{X}_{ij}\boldsymbol{\beta}, \frac{Q_{ij}^{-1}}{\nu}\right)$	$\exp\left\{\mathbf{X}_{ij}\boldsymbol{\beta} + 0.5\frac{1}{\nu}Q_{ij}^{-1}\right\}$	$\exp\{2\mathbf{X}_{ij}\boldsymbol{\beta} + \nu Q_{ij}^{-1}\} \left(\exp\left\{\nu Q_{ij}^{-1}\right\} - 1\right)$

case of the GSC model, making them comparable. However, the flexibility of the GSH model to incorporate the skewness of the mean distribution represents a unique novelty arising from the association of the mean regressors with the shape parameter. Importantly, the ν parameter does not have an equivalent role in these models and we do not expect the same estimated value for this parameter under model misspecification.

Assume we have $Y_{ij}|\mu_{ij} \sim \text{Poisson}(\mu_{ij})$, $\forall i = 1, \dots, n$ and $j = 1, \dots, t$. Let $\mathbf{Q}_\rho$ be the structure matrix of the spatio-temporal MCAR defined by Equation (12) and let $\boldsymbol{\beta}$ be a coefficient vector of dimension q . We used a Gibbs Sampling algorithm with Metropolis–Hastings step for each parameter in the modelling. Priors distributions were set to be flat on their domain even for the dependence parameters.

To compare methods, we used the widely applicable information criterion (WAIC; Watanabe, 2010), the logarithm of the pseudo marginal likelihood (LPML; Dey et al., 1997; Geisser & Eddy, 1979) and the deviance information criterion (DIC; Spiegelhalter et al., 2002). A broader discussion of these criteria can be found in Gelman et al. (2014).

To allow reproducibility and provide access for a wider range of practitioners, an R package has been created that can be installed following the instructions in the TGMRF: Transformed Gaussian Markov Random Fields repository <https://github.com/douglasmesquita/TGMRF>.

5 | SIMULATION STUDY

To evaluate our method, we performed a simulation study. The global sample size is always fixed at 300 but the spatio-temporal design varies across scenarios. The MCMC setup was calibrated after empirical tests that showed that a chain with 1000 samples thinned by 10 to reduce autocorrelation after 5000 iterations of burn-in (15,000 iterations in total) is sufficient to achieve convergence and estimate parameters.

As our method is applied in a spatio-temporal setting, we divided our study into three parts. First, we investigated the ability of our method to recover parameters under a situation in which we have temporal but not spatial independence. Second, we explored a scenario where there is spatial but not temporal independence. Finally, we considered a more realistic scenario in which spatial, temporal and spatio-temporal dependence is present.

For all scenarios, data were generated from one of the models (Table 1). A total of 100 data sets was generated for each proposed model. Nonetheless, we fitted the data set using the model introduced in Equation (11) with all marginals from Table 1 and precision matrix, $\mathbf{Q}_\rho$, defined by Equation (13). For all simulated scenarios, the data were generated in a 6×5 grid with 10 sampling times per point, also, the fixed effects $\boldsymbol{\beta}$ were set as $(1, -0.1)$. The dependence parameters $\boldsymbol{\rho} = (\rho_s, \rho_t, \rho_{st})$ vary according to a tripartite generating scheme: (1) spatial dependence, $\boldsymbol{\rho} = (2.18, 0, 0)$; (2) temporal dependence, $\boldsymbol{\rho} = (0, 3.88, 0)$; and (3) spatio-temporal dependence, $\boldsymbol{\rho} = (0.97, 1.71, 0.77)$.

To demonstrate the accuracy of the method, we present the results of Scenario 3 in Table 2. Results are summarized as modes, standard deviations and mean squared errors (MSEs). The different choices for ν were such that the mean marginal variance $V(\mu_{ij})$ of each model was set at ≈ 10 (Table 2). The point estimates of the parameters are well recovered for the true generating mode with a low MSE. Even under model misspecification β_1 and $\boldsymbol{\rho}$ are nicely recovered in all models. Because of the copula separation of the TMGMRF the dependence parameters in $\boldsymbol{\rho}$ are not predicated on the choice of the marginal link. The traditional LN model does not have the same marginal mean as that of the other proposals, for this reason, β_0 for the LN model is not

TABLE 2 Simulation study for the spatio-temporal scenario. The ν parameter are not comparable across models. Results are shown as mode (standard deviation) and mean squared error (MSE) for 100 simulated data sets

Specified model													
True model	Parameters	True value	GI			GSC			GSH			LN	
			Mode (SD)	MSE		Mode (SD)	MSE		Mode (SD)	MSE		Mode (SD)	MSE
GI	β_0	1.00	0.94(0.06)	0.0034		0.94(0.06)	0.0033		0.94(0.06)	0.0038		0.40(0.07)	0.3556
	β_1	-0.10	-0.09(0.05)	0.0001		-0.12(0.07)	0.0003		-0.12(0.06)	0.0005		-0.16(0.08)	0.0038
	ρ_s	0.97	1.07(0.52)	0.0110		1.04(0.51)	0.0058		1.07(0.52)	0.0101		0.99(0.53)	0.0006
	ρ_t	1.71	1.61(0.80)	0.0117		1.57(0.80)	0.0204		1.63(0.80)	0.0072		1.47(0.82)	0.0586
	ρ_{st}	0.77	0.51(0.39)	0.0655		0.53(0.38)	0.0588		0.52(0.39)	0.0627		0.61(0.40)	0.0260
	ν	0.10	0.13(0.03)	0.0009		0.81(0.13)	0.5103		0.33(0.06)	0.0510		0.11(0.02)	0.0001
GSC	β_0	1.00	0.98(0.04)	0.0005		0.98(0.05)	0.0004		0.98(0.04)	0.0005		0.77(0.05)	0.0527
	β_1	-0.10	-0.07(0.04)	0.0012		-0.09(0.05)	0.0000		-0.08(0.05)	0.0002		-0.10(0.05)	0.0000
	ρ_s	0.97	1.06(0.62)	0.0087		1.07(0.61)	0.0100		1.03(0.61)	0.0039		0.99(0.61)	0.0006
	ρ_t	1.71	1.21(1.07)	0.2552		1.21(1.06)	0.2583		1.23(1.05)	0.2384		1.09(1.05)	0.3827
	ρ_{st}	0.77	0.57(0.46)	0.0400		0.54(0.46)	0.0508		0.57(0.46)	0.0412		0.65(0.47)	0.0133
	ν	2.00	0.31(0.72)	2.8471		2.08(1.29)	0.0069		0.81(1.33)	1.4156		0.29(1.28)	2.9173
GSH	β_0	1.00	0.94(0.06)	0.0032		0.94(0.06)	0.0033		0.94(0.06)	0.0035		0.41(0.07)	0.3432
	β_1	-0.10	-0.07(0.05)	0.0007		-0.12(0.07)	0.0003		-0.11(0.06)	0.0000		-0.15(0.08)	0.0027
	ρ_s	0.97	1.06(0.52)	0.0092		1.06(0.51)	0.0081		1.06(0.52)	0.0097		1.02(0.53)	0.0028
	ρ_t	1.71	1.59(0.80)	0.0157		1.56(0.80)	0.0222		1.57(0.80)	0.0219		1.47(0.82)	0.0588
	ρ_{st}	0.77	0.52(0.38)	0.0640		0.53(0.38)	0.0579		0.52(0.38)	0.0641		0.61(0.39)	0.0247
	ν	0.27	0.13(0.03)	0.0195		0.83(0.13)	0.3098		0.33(0.06)	0.0035		0.11(0.02)	0.0259
LN	β_0	1.00	1.26(0.04)	0.0688		1.26(0.04)	0.0687		1.26(0.04)	0.0672		1.00(0.04)	0.0000
	β_1	-0.10	-0.05(0.04)	0.0021		-0.10(0.05)	0.0000		-0.08(0.05)	0.0004		-0.10(0.05)	0.0000
	ρ_s	0.97	1.01(0.52)	0.0015		1.02(0.52)	0.0028		0.99(0.52)	0.0007		1.11(0.54)	0.0211
	ρ_t	1.71	1.45(0.83)	0.0673		1.40(0.83)	0.0961		1.37(0.84)	0.1156		1.54(0.90)	0.0304
	ρ_{st}	0.77	0.59(0.42)	0.0336		0.60(0.42)	0.0287		0.61(0.42)	0.0254		0.57(0.43)	0.0391
	ν	0.27	0.15(0.05)	0.0134		1.79(0.35)	2.3014		0.52(0.13)	0.0647		0.25(0.14)	0.0002

comparable with the Gamma proposals. The same consideration can be made for ν , since this parameter is not comparable along with the different marginals, thus its estimated values are unrelated.

Similar observations are made for scenarios 1 and 2 (see Tables S.1 and S.2 in the Supplementary Material S.1). Overall, the regression coefficient estimates (β_1) and dependence parameters in ρ are stable across the proposed marginals. Furthermore, the TMGMRF correctly detects the type of the data dependence according to with the generating scenario. Specifically, in scenario 1, only ρ_s is significantly different from 0; in scenario 2, only ρ_t is significantly different from 0; and in scenario 3, ρ_s , ρ_t and ρ_{st} are each significantly different from 0. Therefore, we conclude that our method can accurately recover spatial, temporal and spatio-temporal characteristics as well as fixed effects coefficients, and scale or variability parameters.

6 | ABUNDANCE OF *NENIA TRIDENS*

Our research integrates several fundamental principles of ecology (Scheiner & Willig, 2008) by exploring the bases of the heterogeneous distribution of organisms in space and time, and by linking such dynamics to the heterogeneous distribution of abiotic and biotic factors that represent local habitat characteristics, many of which are affected by disturbance and subsequent secondary succession. Indeed, this integration is a paramount challenge in ecology and biodiversity science and has critical ramifications for wildlife management and conservation action. Nonetheless, most ecological research considers spatio-temporal dynamics over relatively short periods of 3–6 years, thereby missing opportunities to evaluate long-term dynamics associated with long-term environmental variability. In contrast, we have taken advantage of long-term population data (Bloch & Willig, 2006; Willig et al., 1998, 2007) in a well-studied tropical ecosystem (Brokaw et al., 2012) that is subject to climate-induced disturbances (i.e. cyclonic storms and droughts) to illustrate the utility of our new statistical model and to evaluate the insights it provides for ecological understanding.

We investigated spatial, temporal and spatio-temporal trends in the abundance of *N. tridens* as well as in the environmental characteristics that may affect such variability. For this, we fitted model (11) with $\mathbf{Q}_\rho$ and marginals discussed at Section 4, and consider two possible fits. In one, we include the effects of covariates that are constant over time, whereas in another, we additionally allow regression coefficients to vary in time by setting $\boldsymbol{\beta} = (\beta_{11}, \dots, \beta_{1t}, \beta_{21}, \dots, \beta_{3t})^\top$ with independent normal priors. The second approach was proposed to ascertain if any patterns arise when fitting temporal fixed effects for the covariates.

As can be seen in Figure 3 the evolution of the coefficients over time does not suggest any pattern. Consequently, we believe that the constant fixed effect model is more parsimonious and should provide equivalent insight to those of the other models.

The spatial and the temporal parameters for the model with constant fixed effects (Table 3) were significantly greater than 0, whereas the spatio-temporal dependence was significantly smaller than 0. To study the strength of the results obtained for ρ_s , ρ_t and ρ_{st} , we compared the posterior estimates with the marginal limits, based on the diagonal dominance criterion, calculated for ρ_s when $\rho_t = \rho_{st} = 0$ and analogously for ρ_t and ρ_{st} . These limits are $\rho_s^{\max} = 2.25$, $\rho_t^{\max} = 4.00$ and $\rho_{st}^{\max} = 1.75$. This implies that $\hat{\rho}_s / \rho_s^{\max} \approx 0.76$ and

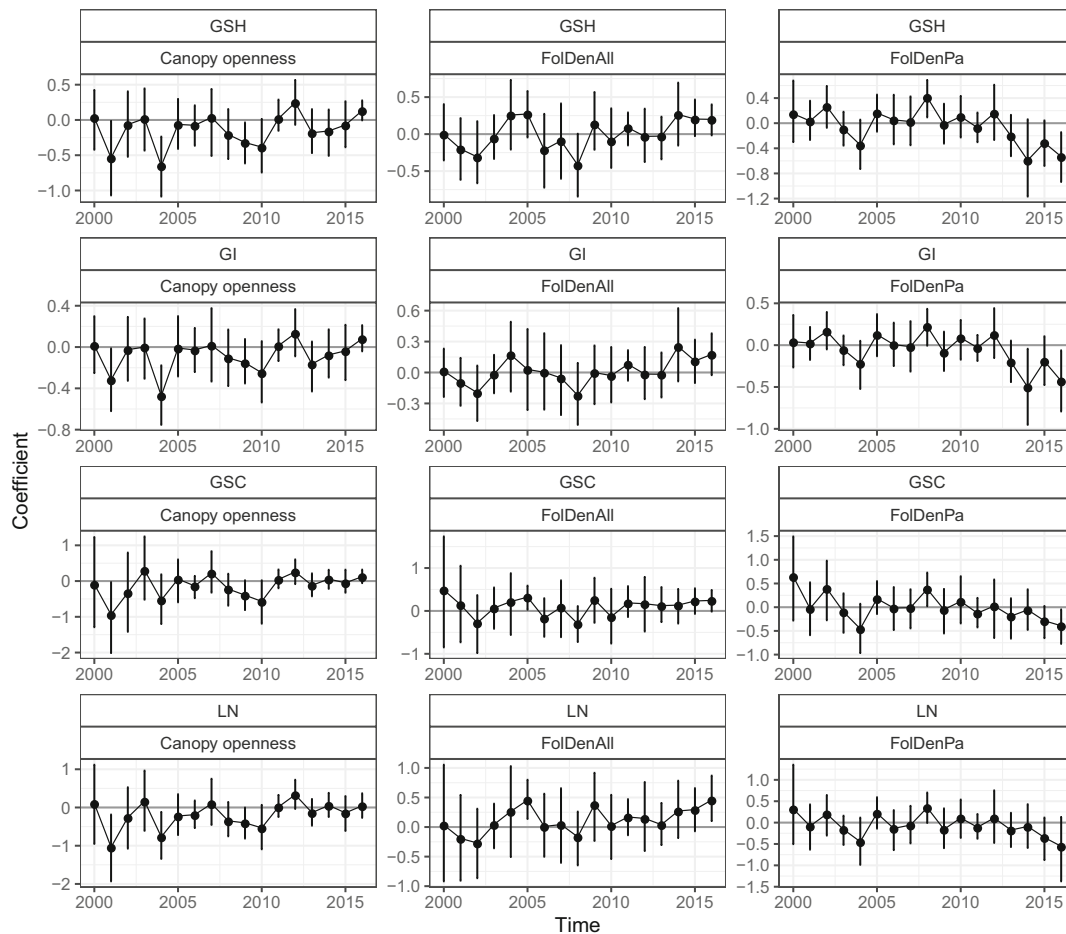


FIGURE 3 Time-varying regression coefficients for Canopy openness, FolDenAll and FolDenPa with their respective 95% credible intervals.

$\hat{\rho}_t/\rho_t^{\max} \approx 0.85$, evidencing a strong association, while $\hat{\rho}_{st}/\rho_{st}^{\max} \approx -0.15$, indicating a mild negative cross-dependence.

Spatial dependence of abundance of a particular site with regard to abundances at neighbouring sites is likely due to the effect of immigration and emigration among those sites. These lead to the greater similarity among sites in abundance than expected by chance. Temporal dependence of abundance between consecutive time periods arises from the demographic process such as site-specific birth rates and death rates. In ecological terms, the correspondence between site-specific abundance of *N. tridens* at one time on abundances of *N. tridens* at neighbouring sites at a subsequent time are not independent. Rather, low abundance at a site at one time is associated with high abundance at neighbouring sites at a future time. Conversely, high abundance at a site at one time is associated with low abundances at neighbouring sites at a future time. As long as the fundamental niche of a species does not change over the time of the study (i.e. we are examining ecological rather than evolutionary dynamics), individuals should be responding to the same suite of environmental characteristics. Indeed, previous research on *N. tridens* (Secrest et al., 1996) has shown that abundance is related to the same characteristics of the environment in two areas of tabonuco forest that differ from each other in the intensity of disturbance from

TABLE 3 Parameters estimation of the parameters for the application. Results are shown as mode (standard deviation) and high posterior density (HPD) intervals

Parameter	GI		GSC		GSH		LN	
	Mode (SD)	HPD 90%	Mode (SD)	HPD 90%	Mode (SD)	HPD 90%	Mode (SD)	HPD 90%
Intercept	0.78 (0.09)	(0.64, 0.91)	0.72 (0.11)	(0.55, 0.89)	0.69 (0.12)	(0.49, 0.85)	-0.18 (0.11)	(-0.39, -0.03)
Elevation	0.01 (0.06)	(-0.09, 0.11)	-0.03 (0.12)	(-0.21, 0.17)	0.02 (0.10)	(-0.14, 0.19)	-0.04 (0.16)	(-0.29, 0.22)
Slope	0.04 (0.04)	(-0.03, 0.10)	0.00 (0.06)	(-0.10, 0.10)	0.04 (0.06)	(-0.04, 0.15)	-0.02 (0.07)	(-0.13, 0.09)
FolDenAll	0.04 (0.04)	(-0.03, 0.11)	0.12 (0.06)	(0.02, 0.23)	0.08 (0.06)	(0.00, 0.18)	0.18 (0.07)	(0.08, 0.30)
Litter cover								
Low	Ref.		Ref.		Ref.		Ref.	
Medium	0.23 (0.08)	(0.10, 0.35)	0.33 (0.13)	(0.14, 0.55)	0.34 (0.13)	(0.12, 0.55)	0.42 (0.14)	(0.18, 0.66)
High	0.40 (0.11)	(0.21, 0.56)	0.50 (0.16)	(0.24, 0.72)	0.52 (0.16)	(0.28, 0.76)	0.57 (0.17)	(0.32, 0.86)
FolDenPa	-0.02 (0.04)	(-0.09, 0.05)	-0.05 (0.06)	(-0.14, 0.07)	-0.03 (0.06)	(-0.13, 0.06)	-0.07 (0.07)	(-0.19, 0.03)
Canopy openness	-0.02 (0.03)	(-0.08, 0.03)	-0.01 (0.05)	(-0.09, 0.07)	-0.03 (0.05)	(-0.10, 0.05)	-0.08 (0.06)	(-0.17, 0.01)
ρ_s	1.74 (0.16)	(1.45, 1.92)	1.69 (0.20)	(1.38, 1.94)	1.72 (0.15)	(1.44, 1.90)	1.57 (0.18)	(1.21, 1.79)
ρ_t	3.39 (0.15)	(3.17, 3.63)	3.40 (0.16)	(3.13, 3.64)	3.40 (0.15)	(3.11, 3.60)	3.12 (0.18)	(2.82, 3.38)
ρ_{st}	-0.26 (0.14)	(-0.50, -0.06)	-0.25 (0.17)	(-0.51, 0.02)	-0.26 (0.14)	(-0.48, -0.04)	-0.03 (0.14)	(-0.25, 0.22)
ν	0.07 (0.01)	(0.06, 0.09)	0.51 (0.04)	(0.44, 0.58)	0.20 (0.02)	(0.16, 0.23)	0.09 (0.01)	(0.07, 0.10)
DIC	2189.73		2195.93		2188.92		2291.65	
-2*LPML	4496.43		3588.66		3990.64		3934.76	
WAIC	2024.48		2026.46		2021.31		2115.74	

The bold quantities represent the selected model by the criterion.

Hurricane Hugo. More specifically, the mean and variance of population abundances differed greatly between the two areas, as did the mean values for environmental characteristics, nonetheless the habitat characteristics that predict abundance did not differ significantly.

Critically, the mapping of environmental characteristics in space depends on time as a consequence of habitat changes induced by disturbance and subsequent secondary succession. The idiosyncratic appearance of small scale disturbances (e.g. tree fall gaps) between survey periods could give rise to a negative cross spatio-temporal patterns. For example, gap formation at a previously productive site can reduce abundance of *N. tridens* to zero, but not affect abundances at surrounding sites that continue to be high. The relatively small magnitude of this general effect likely arises because the number of sites affected by such small scales disturbances is relatively small at any particular time interval. Alternatively, if sites within the forest are at different stages of ecological succession as a consequence of disturbances (i.e. patch dynamics; Pickett & Rogers, 1997; Pickett & White, 2013; Willig et al., 2007) at any point in time, and if the rate of change in abundance is non-linear during succession, then abundance of any site in predicting future abundance at surrounding sites will be time-specific as well as site-specific, contributing to negative cross spatio-temporal association.

Based on the model selection criteria presented in Table 3 and introduced in Section 4.1, the GSH model was preferable because it has the best performance in two (DIC and WAIC) of the three model selection criteria (GSC is preferable according to the LPML criterion). Thus, the conventional log-normal approach does not provide the best fit. The FolDenAll represents the foliar density of all plants in the understorey of the forest, whereas litter cover estimates the volume of leaf litter on the forest floor. Gastropods in general, and *N. tridens* in particular, use such live vegetation for the substrate on which to persist, or for food (the leaves themselves or algae, diatoms, or fungi that grow on them). Leaf litter enhances humidity and decreases temperature on the forest floor. Gastropods are very sensitive to desiccation, especially during periods of activity. High humidity in the litter can mitigate microclimatic characteristics of the understorey (e.g. during droughts or in tree fall gaps induced by cyclonic storms) that allow gastropods to persist and be active. Moreover, leaf litter is a substrate on which micro-organisms grow that represent food sources for gastropods such as *N. tridens*. Thus, the importance of these two characteristics is explicable in terms of the natural history of *N. tridens*, and corroborates the results of previous research (Secrest et al., 1996).

7 | FINAL REMARKS

Herein an overview of many multivariate areal spatial models was considered and re-interpreted. Using the conditional mean and variance we show that parameter interpretation between most of the different proposals is not intuitive. With that in mind, we adapted the p-fold MCAR such that parameters have an intuitive ecological interpretation.

Such reparametrization of the multivariate structure is formulated in a spatio-temporal context and combined with the TGMRF approach. The TMGMRF provides flexibility in the marginal distribution of the mean response and separates the mean structure from the dependence structure, thereby avoiding spatio-temporal confounding. As a by-product of this research, we provide the analysed data and an R package (Section 4.1) for this family of models, called TGMRF, for use by empiricists.

Spatio-temporal variation in counts of *N. tridens* is quite complex because of the environmental dynamics associated with disturbance and subsequent secondary succession in this

tropical forest. Nonetheless, a strong positive spatial and temporal association is present in contrast with a weak, but significant, negative spatio-temporal dependence. The occurrence of small scale disturbances between survey periods is a possible explanation for this negative cross spatio-temporal patterns (for more details about this discussion see Section 6). After controlling for spatio-temporal dynamics, two environmental characteristics, the density of vegetation in the understorey and litter cover, accounted for significant variation in mean abundance at each site.

Finally, as the model can be applied with regard to any hierarchical model in future studies, likelihoods other than the Poisson, included in the R package, as are other distribution families for the marginal, thereby providing flexibility and enhancing the utility of the software and its underlying statistical approach. Models that can effectively ascertain the effects of space, time, and their interactions, all in the context of dynamically changing environmental characteristics, are critical tools for ecologists in the Anthropocene. Because the proposed approach and statistical tools are provided in R, these approaches should become widely adopted in a variety of ecological contexts and for any species of organism. Finally, multivariate application for jointly modelling different species living in tabonuco forest can provide different insights about the complex dynamics of the ecological system.

ACKNOWLEDGEMENTS

Marcos O. Prates acknowledges CNPq (436948/2018-4 and PQ-307457/2018-4), FAPEMIG and CAPES for partial financial support. For the synthetic and modelling portions of this research, Michael R. Willig acknowledges support an OPUS grant from NSF (DEB-1950643). Moreover, the field component of this research was facilitated by a number of grants (DEB-0218039, DEB-0620910, DEB-1239764, DEB-1546686 and DEB-1831952) from the US National Science Foundation to the Institute of Tropical Ecosystem Studies, University of Puerto Rico and the International Institute of Tropical Forestry as part of the Long-Term Ecological Research Program in the Luquillo Experimental Forest. Additional support was provided by the USDA Forest Service, the University of Puerto Rico, the Department of Biological Sciences at Texas Tech University and the Center for Environmental Sciences and Engineering at the University of Connecticut. The staff of El Verde Field Station provided valuable logistical support in Puerto Rico. Finally, we thank the mid-sized army of students and colleagues who have assisted with the collection of field data over the years. Ying C. MacNab acknowledges the funding support from an NSERC Discovery grant (RGPIN 238660-13).

ORCID

Marcos O. Prates  <https://orcid.org/0000-0001-8077-4898>

REFERENCES

- Adin, A., Goicoa, T., Hodges, J.S., Schnell, P.M. & Ugarte, M.D. (2021) Alleviating confounding in spatio-temporal areal models with an application on crimes against women in India. *Statistical Modelling*. Available from: <https://doi.org/10.1177/1471082X211015452>
- Alvarez, J. & Willig, M.R. (1993) Effects of treefall gaps on the density of land snails in the Luquillo experimental forest of Puerto Rico. *Biotropica*, 25, 100–110.
- Azevedo, D.R., Bandyopadhyay, D., Prates, M.O., Abdel-Salam, A.-S.G. & Garcia, D. (2020) Assessing spatial confounding in cancer disease mapping using R. *Cancer Reports*, 3, e1263.
- Azevedo, D.R.M., Prates, M.O. & Bandyopadhyay, D. (2020) Alleviating spatial confounding in spatial frailty models. *arXiv preprint arXiv:2008.06911*.
- Azevedo, D.R., Prates, M.O. & Bandyopadhyay, D. (2021) MSPOCK: alleviating spatial confounding in multivariate disease mapping models. *Journal of Agricultural, Biological and Environmental Statistics*, 26, 1–28.

- Besag, J. (1974) Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36, 192–225.
- Bloch, C.P. & Willig, M.R. (2006) Context-dependence of long-term responses of terrestrial gastropod populations to large-scale disturbance. *Journal of Tropical Ecology*, 22, 111–122.
- Breslow, N.E. & Clayton, D.G. (1993) Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88, 9–25.
- Brokaw, N., Crowl, T. & Lugo, A. (2012) *A Caribbean forest tapestry: the multidimensional nature of disturbance and response*. Oxford: Oxford University Press.
- Carlin, B.P. & Banerjee, S. (2003) Hierarchical multivariate CAR models for spatio-temporally correlated survival data. *Bayesian Statistics*, 7, 45–63.
- Dey, D.K., Chen, M.H. & Chang, H. (1997) Bayesian approach for nonlinear random effects models. *Biometrics*, 53, 1239–1252.
- Geisser, S. & Eddy, W.F. (1979) A predictive approach to model selection (Corr: V75 p765). *Journal of the American Statistical Association*, 74, 153–160.
- Gelfand, A.E. & Vounatsou, P. (2003) Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics*, 4, 11–15.
- Gelman, A., Hwang, J. & Vehtari, A. (2014) Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24, 997–1016.
- Hanks, E.M., Schliep, E.M., Hooten, M.B. & Hoeting, J.A. (2015) Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification. *Environmetrics*, 26, 243–254.
- Hughes, J. (2015) copCAR: a flexible regression model for areal data. *Journal of Computational and Graphical Statistics*, 24, 733–755.
- Hughes, J. & Haran, M. (2013) Dimension reduction and alleviation of confounding for spatial generalized linear mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75, 139–159.
- Jin, X., Carlin, B.P. & Banerjee, S. (2005) Generalized hierarchical multivariate CAR models for areal data. *Biometrics*, 61, 950–961.
- Jin, X., Banerjee, S. & Carlin, B.P. (2007) Order-free co-regionalized areal data models with application to multiple-disease mapping. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69, 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. *Journal of the American Statistical Association*, 96, 1506–1521.
- Lister, B.C. & Garcia, A. (2018) Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proceedings of the National Academy of Sciences*, 115, E10397–E10406.
- MacNab, Y.C. (2018) Some recent work on multivariate Gaussian Markov random fields. *Test*, 27, 497–541.
- MacNab, Y.C. (2020) Bayesian estimation of multivariate Gaussian Markov random fields with constraint. *Statistics in Medicine*, 39, 4767–4788.
- Mardia, K. (1988) Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *Journal of Multivariate Analysis*, 24, 265–284.
- Pickett, S.T.A. & Rogers, K.H. (1997) Patch dynamics: the transformation of landscape structure and function. In: Bissonette, J.A. (Ed.) *Wildlife and landscape ecology*. New York, NY: Springer. Available from: https://doi.org/10.1007/978-1-4612-1918-7_4
- Pickett, S.T. & White, P.S. (2013) *The ecology of natural disturbance and patch dynamics*. Amsterdam: Elsevier.
- Prates, M.O. (2011) *Link specification and spatial dependence for generalized linear mixed models*. Storrs, CT: University of Connecticut.
- Prates, M.O., Dey, D.K., Willig, M.R. & Yan, J. (2011) Intervention analysis of hurricane effects on snail abundance in a tropical forest using long-term spatiotemporal data. *Journal of Agricultural, Biological, and Environmental Statistics*, 16, 142–156.
- Prates, M.O., Dey, D.K., Willig, M.R. & Yan, J. (2015) Transformed Gaussian Markov random fields and spatial modeling of species abundance. *Spatial Statistics*, 14, 382–399.
- Prates, M.O., Assunção, R.M. & Rodrigues, E.C. (2019) Alleviating spatial confounding for areal data problems by displacing the geographical centroids. *Bayesian Analysis*, 14, 623–647.
- Prather, C.M., Pelini, S.L., Laws, A., Rivest, E., Woltz, M., Bloch, C.P. et al. (2013) Invertebrates, ecosystem services and climate change. *Biological Reviews*, 88, 327–348.

- Prié, V. (2019) Molluscs. In: White, W., Culver, D. & Pipan, T. (Eds.) *Encyclopedia of caves*, Amsterdam: Elsevier, pp. 725–731.
- 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, 1197–1206.
- Rodrigues, E.C. (2012) *Estruturas de Covariância em Modelos Espaciais Bayesianas*. Ph.D. thesis, Universidade Federal de Minas Gerais.
- Rue, H. & Held, L. (2005) *Gaussian Markov random fields: theory and applications*. Boca Raton, FL: CRC Press.
- Sain, S.R., Furrer, R. & Cressie, N. (2011) A spatial analysis of multivariate output from regional climate models. *The Annals of Applied Statistics*, 5, 150–175.
- Scheiner, S.M. & Willig, M.R. (2008) A general theory of ecology. *Theoretical Ecology*, 1, 21–28.
- Schowalter, T.D., Willig, M.R., Presley, S.J. & Pandey, M. (2019) Warnings of an “insect apocalypse” are premature. *Frontiers in Ecology and the Environment*, 17, 547–547.
- Secrest, M.F., Willig, M.R. & Peppers, L.L. (1996) The legacy of disturbance on habitat associations of terrestrial snails in the Luquillo experimental forest, Puerto Rico. *Biotropica*, 28, 502–514.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & Van Der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 583–639.
- Steffen, W., Crutzen, P.J. & McNeill, J.R. (2007) The anthropocene: are humans now overwhelming the great forces of nature. *AMBIO: A Journal of the Human Environment*, 36, 614–622.
- Thaden, H. & Kneib, T. (2018) Structural equation models for dealing with spatial confounding. *The American Statistician*, 72, 239–252.
- Watanabe, S. (2010) Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571–3594.
- Willig, M.R. & Camilo, G.R. (1991) The effect of hurricane Hugo on six invertebrate species in the Luquillo experimental forest of Puerto Rico. *Biotropica*, 23, 455–461.
- Willig, M.R., Secrest, M.F., Cox, S.B., Camilo, G.R., Cary, J.F., Alvarez, J. et al. (1998) Long-term monitoring of snails in the Luquillo experimental forest of Puerto Rico: heterogeneity, scale, disturbance, and recovery. *Man and the Biosphere Series*, 21, 293–322.
- Willig, M.R., Bloch, C.P., Brokaw, N., Higgins, C., Thompson, J. & Zimmermann, C.R. (2007) Cross-scale responses of biodiversity to hurricane and anthropogenic disturbance in a tropical forest. *Ecosystems*, 10, 824–838.
- Willig, M.R., Bloch, C.P. & Presley, S.J. (2014) Experimental decoupling of canopy opening and debris addition on tropical gastropod populations and communities. *Forest Ecology and Management*, 332, 103–117.
- Willig, M., Woolbright, L., Presley, S., Schowalter, T., Waide, R., Scalley, T.H. et al. (2019) Populations are not declining and food webs are not collapsing at the Luquillo experimental forest. *Proceedings of the National Academy of Sciences*, 116, 12143–12144.
- Willig, M.R., Presley, S.J. & Cullerton, E.I. (2021) A canonical metacommunity structure over 3 decades: ecologically consistent but spatially dynamic patterns in a hurricane-prone montane forest. *Oecologia*, 196, 919–933.
- Zalasiewicz, J., Williams, M., Steffen, W. & Crutzen, P. (2010) The new world of the anthropocene. *Environmental Science & Technology*, 44, 2228–2231.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Prates, M.O., Azevedo, D.R.M., MacNab, Y.C. & Willig, M.R. (2022) Non-separable spatio-temporal models via transformed multivariate Gaussian Markov random fields. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 71(5), 1116–1136. Available from: <https://doi.org/10.1111/rssc.12567>