Received: 13 October 2022 | Revised: 12 September 2023 | Accepted: 12 October 2023

DOI: 10.1111/stan.12330

ORIGINAL ARTICLE

WILEY

Scaling priors for intrinsic Gaussian Markov random fields applied to blood pressure data

Maria-Zafeiria Spyropoulou[1] | **James Bentham**[2]

[1]School of Sport and Exercise Science, University of Kent, Canterbury, UK

[2]School of Statistics, Mathematics, and Actuarial Science, University of Kent, Canterbury, UK

Correspondence
Maria-Zafeiria Spyropoulou, School of Sport and Exercise Science, University of Kent, Canterbury CT2 7FS, UK.
Email: mzs@kent.ac.uk

Funding information
Vice-Chancellor's Research Scholarship, University of Kent

An Intrinsic Gaussian Markov Random Field (IGMRF) can be used to induce conditional dependence in Bayesian hierarchical models. IGMRFs have both a precision matrix, which defines the neighborhood structure of the model, and a precision, or scaling, parameter. Previous studies have shown the importance of selecting the prior for this scaling parameter appropriately for different types of IGMRF, as it can have a substantial impact on posterior estimates. Here, we focus on cases in one and two dimensions, where tuning of the prior is achieved by mapping it to the marginal *SD* of an IGMRF of corresponding dimensionality. We compare the effects of scaling various IGMRFs, including an application to real two-dimensional blood pressure data using MCMC methods.

KEYWORDS
hyperpriors, intrinsic Gaussian Markov random fields, MCMC, precision, scaling, two-dimensional data

1 | INTRODUCTION

Intrinsic Gaussian Markov Random Fields (IGMRFs) are used widely as prior distributions in Bayesian hierarchical models, particularly for modeling spatial or temporal data, as they capture conditional dependence through their precision matrices (Rue & Held, 2005). We examine one- and two-dimensional IGMRFs, where the latter can capture dependence between a pair of

variables at multiple time points. IGMRFs are of various types, and can be specified to induce particular neighborhood structures for the precisions, either by varying weights, introducing certain behavior at boundaries or within the precision matrix, or by considering different sets of neighbors (Terzopoulos, 1988).

We present results for the two-dimensional case that generalize previous work in one dimension on scaling different types of field. In that case, an IGMRF was used as the prior for capturing non-linear trends, with a hyperprior for the precision parameter (Sørbye & Rue, 2014); we select priors for precision parameters so that the same degree of scaling is applied to bivariate data as in the one-dimensional case. We have shown previously that models can be fitted to two-dimensional blood pressure data using MCMC algorithms (Spyropoulou, 2023a). In the work of Simpson, Rue, Riebler, Martins, and Sørbye (2017), penalized complexity (PC) priors are presented; these are priors that can be vague, weakly, or strongly informative depending on the tuning of the scaling parameter. PC priors are defined in terms of four principles, which provide beneficial properties such as invariance to reparameterization, a connection with Jeffreys' prior, and robustness to the flexibility parameter chosen by the user. However, we suggest a prior that is not restricted to any principle and is also invariant to reparameterization. In this paper, we show that scaling requires particular care in two dimensions, where differences in behavior between IGMRFs may be larger than in a single dimension.

The paper is structured as follows. Section 2 describes the behavior of IGMRFs in one and two dimensions, while Section 3 describes the mapping between the precision parameter and the marginal *SD* for two-dimensional IGMRFs. Analyses using real two-dimensional blood pressure data are presented in Section 4, with a discussion of our findings and suggestions for future work in Section 5.

2 | USE OF IGMRFS AS PRIORS

2.1 | Motivation

Blood pressure is bivariate, with measurements comprising systolic and diastolic values (SBP and DBP, respectively). Although a realistic one-dimensional model of trends in mean SBP or DBP at national level is available (Danaei et al., 2011; Finucane, Paciorek, Danaei, & Ezzati, 2014), it cannot make estimates for both variables simultaneously, and no information is captured on interactions between them. We have developed a two-dimensional extension, including analogous terms to the original model (Spyropoulou, 2023b). Specifically, a vector $\mathbf{y}_{h,i}$ of the means of SBP, DBP, and their interaction, indexed by age group h and study i in country j, is assumed to be distributed.

$$\mathbf{y}_{h,i} \sim \mathcal{N}(\mathbf{a}_{j[i]} + \mathbf{b}_{j[i]}t_i + \mathbf{u}_{j[i],t_i} + \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\gamma}_i(z_h) + \mathbf{e}_i, \mathbf{SD}^2_{h,i}/n_{h,i} + \boldsymbol{\tau}^2_i). \tag{1}$$

The model includes country-level linear intercepts and slopes, $\mathbf{a}_{j[i]}$ and $\mathbf{b}_{j[i]}$, time-varying nonlinear terms, $\mathbf{u}_{j[i],t_i}$, covariate effects $\boldsymbol{\beta}$, terms in age $\boldsymbol{\gamma}_i$, study-specific random effects $\mathbf{e}_i$, age-varying study-specific random effects $\mathbf{w}_{h,i}$ with mean equal to zero corresponding to $\boldsymbol{\tau}^2_i$, and noise $\boldsymbol{\epsilon}_{h,i}$, assumed iid Gaussian. Lastly, $\mathbf{SD}^2_{h,i}/n_{h,i}$ expresses the variance for age group h in study i with a sample size of $n_{h,i}$. In the earlier work, a one-dimensional second-order IGMRF was used as a prior for the $\mathbf{u}_{j[i]}$ terms, which we have extended to the two-dimensional case including changes in the degree of scaling applied.

2.2 | IGMRFs of one and two dimensions

We begin by comparing the behavior of one-dimensional first-order and second-order IGMRFs with the two-dimensional second-order case (Rue & Held, 2005). An IGMRF can be defined as

$$\pi(\mathbf{u}) = (2\pi)^{-(n-k)/2}(|\mathbf{Q}|^*)^{1/2}\exp\Big(-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu})^T\mathbf{Q}(\mathbf{u}-\boldsymbol{\mu})\Big). \tag{2}$$

Let $\mathbf{Q}$ be an $n\times n$ symmetric positive semi-definite matrix with rank $n-k>0$ (Rue & Held, 2005), where k and n denote the order of the IGMRF and the total number of nodes, respectively. $|\mathbf{Q}|^*$ expresses the generalized determinant of the $\mathbf{Q}$ precision matrix taking the product of the $n-k$ nonzero eigenvalues, Sørbye and Rue (2014). Then $\mathbf{u}=(u_1,u_2,\ \dots\ ,u_n)^T$ is an improper GMRF of rank $n-k$ with parameters $(\mu,\mathbf{Q})$. The vector $\mathbf{u}$ is represented by a graph $G=(V,E)$ where V represents vertices or nodes and E represents edges or the connections between nodes such that

$$Q_{ij}\neq 0 \Leftrightarrow \{i,j\}\in E, \quad \forall i\neq j. \tag{3}$$

The n nodes define the dimensionality of the covariance matrix, where for n nodes we have $n\times n$ dimensions in the precision matrix. Considering $\mu=0$ and focusing on the structure of $\mathbf{Q}$, Equation (2) is written as:

$$\pi(\mathbf{u}) \propto (|\mathbf{Q}|^*)^{1/2}\exp\Big(-\frac{1}{2}\mathbf{u}^T\mathbf{Q}\mathbf{u}\Big). \tag{4}$$

As described previously by Rue and Held (2005), for a vector of observations $\mathbf{u}$ of length n, the one-dimensional first-order model assumes independent increments

$$\Delta u_s = u_{s+1}-u_s \sim \mathcal{N}(0,\lambda^{-1}), \quad s=1,\dots,n-1, \tag{5}$$

with joint density

$$\pi(\mathbf{u}|\lambda) \propto \lambda^{(n-1)/2}\exp\left(-\frac{\lambda}{2}\sum_{s=1}^{n-1}(u_{s+1}-u_s)^2\right). \tag{6}$$

Writing Equation (6) in a more compact form we have:

$$\begin{aligned}\pi(\mathbf{u}|\lambda) &\propto \lambda^{(n-1)/2}\exp\Big(-\frac{\lambda}{2}(\mathbf{D}\mathbf{u})^T\mathbf{D}\mathbf{u}\Big)\\ &\propto \lambda^{(n-1)/2}\exp\Big(-\frac{1}{2}\mathbf{u}^T\lambda\mathbf{P}\mathbf{u}\Big)\\ &\propto \lambda^{(n-1)/2}\exp\Big(-\frac{1}{2}\mathbf{u}^T\mathbf{Q}\,\mathbf{u}\Big),\end{aligned} \tag{7}$$

where $\mathbf{D}$ expresses differences within $\mathbf{u}$, with $\mathbf{P}=\mathbf{D}^T\mathbf{D}$ and $\mathbf{Q}=\lambda\mathbf{P}$; $\mathbf{D}$ and $\mathbf{Q}$ are defined as

$$\mathbf{D}=\begin{pmatrix}1 & -1 & & & & \\ & 1 & -1 & & & \\ & & 1 & -1 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & & 1 & -1\end{pmatrix}, \quad \mathbf{Q}=\lambda\mathbf{P}=\lambda\begin{pmatrix}1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & & -1 & 1\end{pmatrix}.$$

The second-order model assumes independent increments and joint density

$$\Delta^2 u_s = u_{s+2} - 2u_{s+1} + u_s \sim \mathcal{N}(0, \lambda^{-1}), \quad s = 1, \ldots, n-2. \tag{8}$$

$$\pi(\mathbf{u}|\lambda) \propto \lambda^{(n-2)/2} \exp\left(-\frac{\lambda}{2} \sum_{s=1}^{n-2} (u_{s+2} - 2u_{s+1} + u_s)^2 \right). \tag{9}$$

The compact form for the second-order equivalent is:

$$\pi(\mathbf{u}|\lambda) \propto \lambda^{(n-2)/2} \exp\left(-\frac{1}{2} \mathbf{u}^T \lambda \mathbf{P}\, \mathbf{u} \right). \tag{10}$$

with

$$\mathbf{D} = \begin{pmatrix} 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & 1 & -2 & 1 \end{pmatrix}, \quad \mathbf{Q} = \lambda \mathbf{P} = \lambda \begin{pmatrix} 1 & -2 & 1 & & \\ -2 & 5 & -4 & 1 & \\ 1 & -4 & 6 & -4 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & 1 & -4 & 5 & -2 \\ & & 1 & -2 & 1 \end{pmatrix}.$$

In two dimensions, a second-order model constructed on a torus assumes independent two-dimensional second-order increments (Rue & Held, 2005), and for variables indexed d and s we have

$$\begin{aligned} \Delta_0^2 u_{d,s} &= (\Delta_{(1,0)}^2 + \Delta_{(0,1)}^2) u_{d,s} \\ &= u_{d+2,s} - 2u_{d+1,s} + 2u_{d,s} - 2u_{d,s+1} + u_{d,s+2} \\ &= u_{d+1,s} - 4u_{d,s} + u_{d-1,s} + u_{d,s+1} + u_{d,s-1} \sim \mathcal{N}(0, \lambda^{-1}), \end{aligned} \tag{11}$$

with joint density

$$\pi(\mathbf{u}|\lambda) \propto \lambda^{(n_1 \times n_2 - 3)/2} \exp\left(-\frac{\lambda}{2} \sum_{d=2}^{n_1-1} \sum_{s=2}^{n_2-1} (\Delta_{(1,0)}^2 u_{d,s} + \Delta_{(0,1)}^2 u_{d,s})^2 \right), \tag{12}$$

where n_1 and n_2 represent the total number of nodes for each variable.

Our models are time-varying, so the assumption of an IGMRF on a torus is not appropriate, and the two-dimensional second-order density in this case (Yue & Speckman, 2010) has the more complex form

$$\begin{aligned} \pi(\mathbf{u}|\lambda) \propto \lambda^{(n_1 \times n_2 - 3)/2} \exp\Bigg(& -\frac{\lambda}{2} \sum_{d=2}^{n_1-1} \sum_{s=2}^{n_2-1} \{\Delta_0^2 u_{d,s}\}^2 + \{\Delta_1 u_{1,1}\}^2 + \{\Delta_2 u_{n_1,1}\}^2 \\ & + \{\Delta_3 u_{1,n_2}\}^2 + \{\Delta_4 u_{n_1,n_2}\}^2 + \sum_{d=2}^{n_1} (\{\Delta_5 u_{d,1}\}^2 + \{\Delta_6 u_{d,n_2}\}^2) \\ & + \sum_{s=2}^{n_2} (\{\Delta_7 u_{1,s}\}^2 + \{\Delta_8 u_{n_1,s}\}^2) \Bigg). \end{aligned} \tag{13}$$

A special case of (13) arises when the variables have the same number of nodes, that is, $n = n_1 = n_2$

$$\pi(\mathbf{u}|\lambda) \propto \lambda^{(n^2-3)/2} \exp\Bigg(-\frac{\lambda}{2} \sum_{d=2}^{n-1} \sum_{s=2}^{n-1} \{\Delta_0^2 u_{d,s}\}^2 + \{\Delta_1 u_{1,1}\}^2 + \{\Delta_2 u_{n,1}\}^2 + \{\Delta_3 u_{1,n}\}^2 + \{\Delta_4 u_{n,n}\}^2 + \sum_{d=2}^{n} (\{\Delta_5 u_{d,1}\}^2 + \{\Delta_6 u_{d,n}\}^2) + \sum_{s=2}^{n} (\{\Delta_7 u_{1,s}\}^2 + \{\Delta_8 u_{n,s}\}^2) \Bigg). \tag{14}$$

The compact form for the second-order equivalent will be:

$$\pi(\mathbf{u}|\lambda) \propto \lambda^{(n-2)/2} \exp\Big(-\frac{1}{2} \mathbf{u}^T \lambda \mathbf{P} \mathbf{u} \Big), \tag{15}$$

with $\mathbf{Q}$ defined as (Spyropoulou, 2023b; Yue & Speckman, 2010)

$$\mathbf{A}_1 = \begin{pmatrix} 6 & -5 & 1 & & \\ -5 & 12 & -6 & 1 & \\ 1 & -6 & 12 & -6 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & 1 & -6 & 12 & -5 \\ & & 1 & -5 & 6 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} -5 & 2 & & & \\ 2 & -7 & 2 & & \\ & 2 & -7 & 2 & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & 2 & -7 & 2 \\ & & & 2 & -5 \end{pmatrix}.$$

$$\mathbf{A_3} = \text{diag}(1, 1, 1, 1, 1).$$

$$\mathbf{A}_4 = \begin{pmatrix} 12 & -7 & 1 & & \\ -7 & 20 & -8 & 1 & \\ 1 & -8 & 20 & -8 & 1 \\ & 1 & -8 & 20 & -7 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & 1 & -7 & 12 \end{pmatrix}, \quad \mathbf{A}_5 = \begin{pmatrix} -6 & 2 & & & \\ 2 & -8 & 2 & & \\ & 2 & -8 & 2 & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & 2 & -8 & 2 \\ & & & 2 & -6 \end{pmatrix}.$$

$$\lambda \mathbf{P} = \lambda \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{A}_3 & & \\ \mathbf{A}_2 & \mathbf{A}_4 & \mathbf{A}_5 & \mathbf{A}_3 & \\ \mathbf{A}_3 & \mathbf{A}_5 & \mathbf{A}_4 & \mathbf{A}_5 & \mathbf{A}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & \mathbf{A}_3 & \mathbf{A}_5 & \mathbf{A}_4 & \mathbf{A}_2 \\ & & \mathbf{A}_3 & \mathbf{A}_2 & \mathbf{A}_1 \end{pmatrix}. \tag{16}$$

For each $\mathbf{Q}$, as the dimensionality of the matrices increases, the central row is repeated.

In summary, in each case we have time-varying nonlinear effects $\mathbf{u} \sim \mathcal{N}(0, (\lambda \mathbf{P})^{-1})$. They follow an IGMRF with structure matrix, $\mathbf{P}$, and precision parameter, λ, which is always a scalar (Yue & Speckman, 2010).

3 | SPECIFYING HYPERPRIORS FOR TWO-DIMENSIONAL IGMRFS

IGMRFs have structure matrices and marginal variances that change depending on their type. Hyperpriors need to be chosen and assigned appropriate ranges for particular models based on their structure, particularly their dimensionality and numbers of nodes (Sørbye & Rue, 2014). For our model, we need to scale λ appropriately for a two-dimensional second-order IGMRF with boundaries, and given that blood pressure data from national-level surveys are available from around 1980, we consider models with up to 40 nodes. Here, we derive reference *SD*s and use them to select appropriate values for specific hyperpriors.

3.1 | Reference *SD*

We can describe an IGMRF using an alternative definition where

$$\sigma^2_\lambda(u_i) = \lambda^{-1}\Sigma^*_{ii}. \tag{17}$$

Here, Σ^*_{ii} is the diagonal element of the covariance matrix in position i. For the standard normal distribution, we have

$$\mathbf{u}\sqrt{\lambda} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{P}^{-1}\right), \quad \lambda\sigma^2_\lambda(u_i) = \Sigma^*_{ii}. \tag{18}$$

Therefore for $\lambda = 1$, we have

$$\sigma^2_{\{\lambda=1\}}(u_i) = \Sigma^*_{ii}. \tag{19}$$

Combining the results in (17) and (19), we have

$$\sigma^2_\lambda(u_i) = \lambda^{-1}\Sigma^*_{ii} = \lambda^{-1}\sigma^2_{\{\lambda=1\}}(u_i). \tag{20}$$

This implies that for any fixed precision λ, the marginal *SD* of the components of a Gaussian vector $\mathbf{u}$ can be expressed as a function of λ (Sørbye & Rue, 2014) by

$$\sigma_\lambda(u_i) = \lambda^{-1/2}\sigma_{\{\lambda=1\}}(u_i), \quad i = 1, \ldots, n. \tag{21}$$

For a given IGMRF $\mathbf{u}$ with random precision λ, we can calculate a reference *SD* for fixed $\lambda = 1$, and then approximate the marginal *SD* for each component of $\mathbf{u}$ by Sørbye and Rue (2014)

$$\sigma_\lambda(u_i) \approx \lambda^{-1/2}\sigma_{ref}(\mathbf{u}), \quad i = 1, \ldots, n. \tag{22}$$

The reference *SD* is calculated using the geometric mean, an appropriate measure for a set of positive numbers (Sørbye & Rue, 2014). The reference *SD* for $\mathbf{u}$ in the one-dimensional case is then

$$\sigma_{ref}(\mathbf{u}) = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log\sigma_{\{\lambda=1\}}(u_i)\right) = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\log\Sigma^*_{ii}\right), \tag{23}$$

where Σ^*_{ii} values express the diagonal elements of the generalized inverse matrix $\mathbf{\Sigma}^* = \mathbf{Q}^{-1}$ calculated when $\lambda = 1$. Specifically, this is calculated as $\mathbf{Q}^{-1} = \mathbf{\Gamma}^T \Lambda^{-1} \mathbf{\Gamma}$, where $\mathbf{\Gamma}$ are the eigenvectors and $\mathbf{\Lambda}$ the eigenvalues of $\mathbf{Q}$ when $\lambda = 1$. Since $\mathbf{Q}$ and $\mathbf{\Sigma}^*$ are both $n \times n$ dimensional for one-dimensional IGMRFs, the n diagonal values are used to calculate the geometric mean.

We extend the calculation of $\sigma_{ref}(\mathbf{u})$ to two-dimensional second-order IGMRFs. The precision matrix is now $(n_1 \times n_2) \times (n_1 \times n_2)$ dimensional, where n_1 and n_2 are the total number of nodes for the first and second variables, respectively. The scaling is no longer for n_1 or n_2 values, but their product $n_1 \times n_2$, with a special case when $n = n_1 = n_2$

$$\sigma_{ref}(\mathbf{u}) = \exp\left(\frac{1}{n^2} \sum_{i=1}^{n^2} \log \sigma_{\{\lambda=1\}}(u_i) \right) = \exp\left(\frac{1}{n^2} \sum_{i=1}^{n^2} \frac{1}{2} \log \mathbf{\Sigma}^*_{ii} \right). \tag{24}$$

Again, $\mathbf{\Sigma}^*_{ii}$ denotes the diagonal elements of the generalized inverse matrix $\mathbf{\Sigma}^* = \mathbf{Q}^{-1} = \mathbf{\Gamma}^T \Lambda^{-1} \mathbf{\Gamma}$ for $\lambda = 1$. The precision matrix $\mathbf{Q}$, and therefore $\mathbf{\Sigma}^*$, is $n^2 \times n^2$ dimensional, hence there are n^2 elements in the diagonal.

It has been shown previously in Sørbye and Rue (2014) that using 100 nodes, the reference *SD*s in one dimension are $\sigma_{\text{ref}}(\mathbf{u}) = 3.89$ and $\sigma_{\text{ref}}(\mathbf{u}) = 41.39$ for the first-order and second-order cases, respectively; applying the result in (24), we find that $\sigma_{\text{ref}}(\mathbf{u}) = 7.24$ for the two-dimensional second-order case for 100 nodes in each dimension. As shown in Figure 1, the behavior by node of the marginal *SD*s also varies substantially between the three IGMRFs. For a particular hyperprior, larger variances would be allowed for the one-dimensional second-order IGMRF than its two-dimensional equivalent, and both would have larger variances than the one-dimensional first-order case. Equivalently, to allow the same variance we need to impose an upper limit on the marginal *SD*

$$Pr(\sigma(u_i) > U) \approx Pr\left(\frac{\lambda}{\sigma^2_{\text{ref}}(\mathbf{u})} < \frac{1}{U^2} \right) = \alpha, \tag{25}$$

where α is a fixed small probability (Sørbye & Rue, 2014). By assigning a hyperprior to $\lambda(\sigma^2_{\text{ref}}(\mathbf{u}))^{-1}$, the interpretation of the hyperprior remains the same for the different models.

These results complement others (Lindgren & Rue, 2008; Lindgren, Rue, & Lindström, 2011), where k equally sized subintervals are created between original nodes $u_1, u_2, \ldots$, to give equidistant nodes $u'_1, u'_2, ..., u'_{k+1}$. In the first-order one-dimensional case, the precision using the new nodes is $(k\lambda)^{-1}$, for the second-order equivalent, the precision using the new nodes is $(k^3\lambda)^{-1}$, and finally, for the second-order two-dimensional IGMRF, the precision using the new nodes is $(k^2\lambda)^{-1}$.

3.2 | Specifications using Gaussian hyperpriors

Applying a Gaussian hyperprior, the upper limit expressed in probabilistic form in (25) is

$$U = \left(\frac{b\sigma^2_{\text{ref}}(\mathbf{u})}{F^{-1}(\alpha,\mu,1)} \right)^{1/2}, \tag{26}$$

where $F^{-1}(\cdot)$ denotes the quantiles of the Gaussian distribution (Sørbye & Rue, 2014). Analytically, the calculations are presented in (A1) shown also in (Spyropoulou, 2023b). For a given value of α, we can then interpret the mean and *SD* parameters, μ and b, in terms of this upper limit.

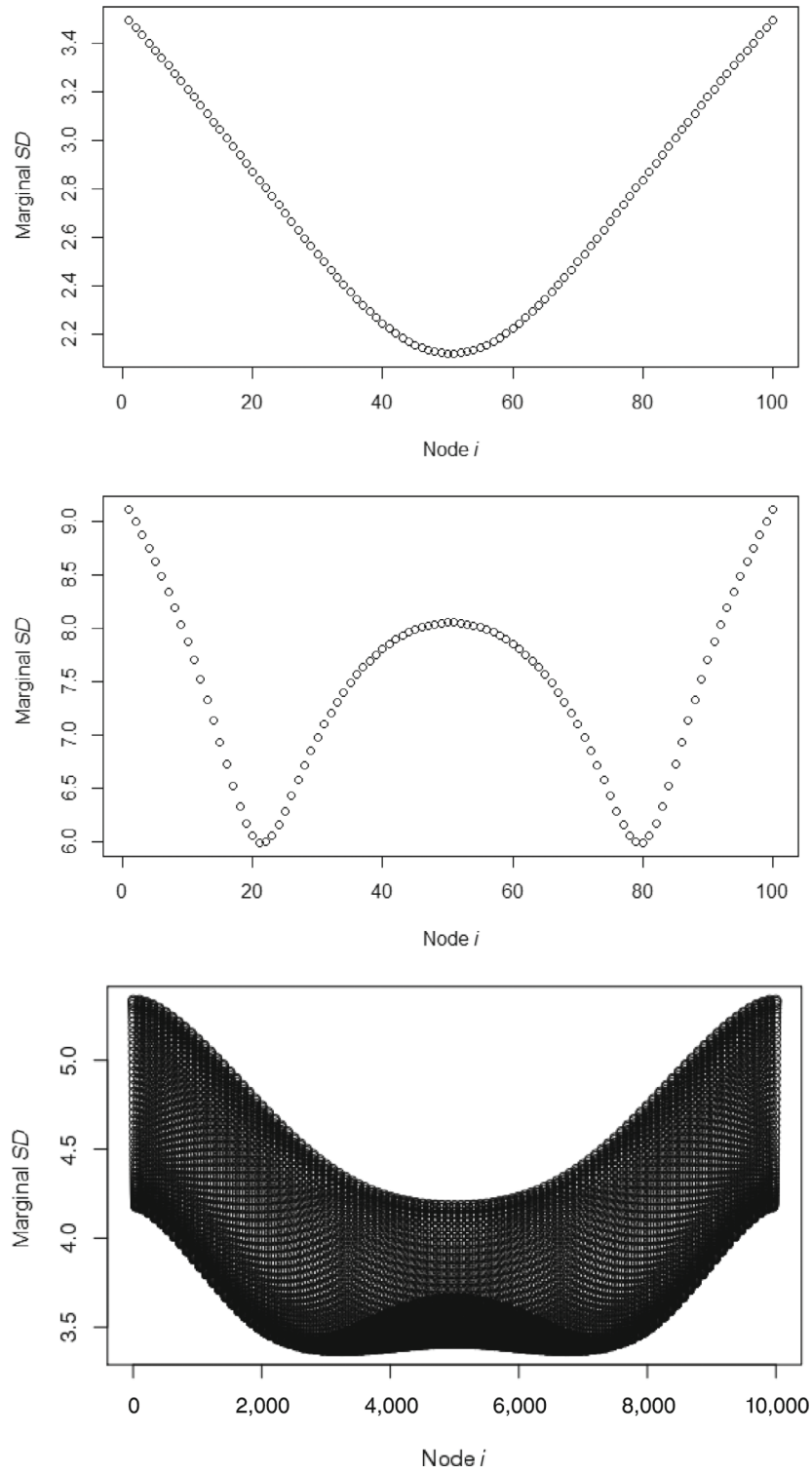

FIGURE 1 Marginal *SD*s of one-dimensional first-order and second-order Sørbye and Rue (2014), and two-dimensional second-order intrinsic Gaussian Markov random fields, calculated using fixed precision $\lambda = 1$.

To recalculate hyperpriors for different IGMRFs, we can use the same mean parameter μ for each model and calculate a new *SD* parameter. By using the upper limit provided in (26), the new *SD* parameter is expressed as

$$b_{\text{new}} = \frac{U^2 F^{-1}(\alpha, \mu, 1)}{\sigma^2_{\text{ref}}(\mathbf{u})} \tag{27}$$

We see that the new value of the *SD* in (27) depends on σ^2_{ref}, which captures the precision matrix for a specific type of IGMRF. It is then only necessary to recalculate the *SD* parameter, b, to account for the different shapes and sizes of the graph for a specific IGMRF (Sørbye & Rue, 2014). This can be done for the three types of IGMRF considered using

$$b_{rw2} = b_{rw1} \times \frac{\sigma^2_{\text{ref}}(\mathbf{u}_{rw1})}{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2})}, \quad b_{rw2D} = b_{rw2} \times \frac{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2})}{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2D})}, \quad b_{rw2D} = b_{rw1} \times \frac{\sigma^2_{\text{ref}}(\mathbf{u}_{rw1})}{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2D})} \tag{28}$$

Here, *rw1* and *rw2* refer to the one-dimensional first- and second-order IGMRFs (Sørbye & Rue, 2014), and *rw2D* to the two-dimensional second-order IGMRF.

3.3 | Types of two-dimensional second-order IGMRFs

While trends in single risk factors for noncommunicable diseases are increasingly well understood, there has been little analysis of bivariate or multivariate distributions even though interactions are of substantial health importance. To provide a guide for researchers applying these types of models to real data, we compare IGMRFs with fixed order but different numbers of nodes and boundary conditions, and IGMRFs of varying order and dimensionality. We do so firstly for two-dimensional second-order IGMRFs with four structure matrices: Torus 1 and Torus 2 (Rue & Held, 2005; Thon, Rue, Skrøvseth, & Godtliebsen, 2012), and Bound 1 (Yue & Speckman, 2010) and Bound 2 (Terzopoulos, 1988). Torus 1 has a structure matrix defined on a torus, while Torus 2 has a similar structure matrix but with boundaries at its four corners, $u_{1,1}, u_{n_1,1}, u_{1,n_2}, u_{n_1,n_2}$. Bound 1 and Bound 2 have boundary effects and induce the same neighbors in the structure matrix for each node, but give different weightings to these neighbors. For our application to time-varying mean blood pressure data and for similar applications, the bounded IGMRFs are appropriate and should be used.

In Table 1, we see that Torus 2 consistently has the smallest reference *SD*, with the changes in each IGMRF being similar proportionally when the number of nodes is increased. Bound 2 has the largest reference *SD*, followed by Bound 1, which we used in our two-dimensional model of blood pressure (Spyropoulou, 2023b). These findings show that it is clearly necessary to scale the hyperparameter each time the precision matrix or number of nodes is changed, especially when boundary conditions are introduced.

Table 2 shows the effect of changing the order or dimensionality of the IGMRFs. For small numbers of nodes, the one-dimensional first-order random walk has the largest reference *SD*. However, this increases slowly with the number of nodes, and for 10 nodes or more there is a consistent pattern, with the one-dimensional second-order random walk having the largest reference *SD*s, while the two-dimensional second-order random walk behaves more similarly to the one-dimensional first-order random walk. These results emphasize the importance of scaling the hyperparameter when dimensionality or order is changed.

TABLE 1 Reference *SD*s σ_{ref} for second-order two-dimensional intrinsic Gaussian Markov random fields.

Nodes	Torus 1	Torus 2	Bound 1	Bound 2
5	0.30	0.05	0.41	0.53
6	0.35	0.06	0.48	0.63
8	0.44	0.08	0.62	0.81
10	0.54	0.09	0.76	1.00
12	0.63	0.11	0.90	1.19
14	0.73	0.12	1.04	1.39
16	0.82	0.14	1.19	1.58
18	0.93	0.16	1.33	1.77
20	1.02	0.17	1.47	1.96
25	1.25	0.21	1.83	2.44
30	1.51	0.25	2.19	2.92
40	2.00	0.33	2.91	3.88

TABLE 2 Reference *SD*s, σ_{ref}, for one-dimensional first-order and second-order intrinsic Gaussian Markov random fields (IGMRFs), and two-dimensional second-order IGMRFs.

Nodes	$\sigma_{\text{ref}}(\mathbf{u}_{rw1})$	$\sigma_{\text{ref}}(\mathbf{u}_{rw2})$	$\sigma_{\text{ref}}(\mathbf{u}_{rw2D})$
5	0.85	0.53	0.41
6	0.94	0.65	0.48
8	1.09	0.96	0.62
10	1.22	1.35	0.76
12	1.34	1.75	0.90
14	1.45	2.20	1.04
16	1.55	2.68	1.19
18	1.65	3.19	1.33
20	1.74	3.73	1.47
25	1.94	5.20	1.83
30	2.13	6.82	2.19
40	2.46	10.49	2.91

4 | APPLICATION TO BLOOD PRESSURE DATA

We compare hyperprior scaling for one- and two-dimensional second-order IGMRFs using blood pressure data (Spyropoulou, 2023b). The scaling varies both by dimensionality and the number of nodes, corresponding to the number of years considered. The hyperpriors must be set for each of the four precision parameters, λ_c, λ_r, λ_s, and λ_g, that are used at different levels of a Bayesian hierarchical model, with countries nested in regions, super-regions and the globe. When using 40 years of data, we calculate the following values

$$\sigma_{\text{ref}}(\mathbf{u}_{rw2}) = 10.49 \qquad \sigma_{\text{ref}}(\mathbf{u}_{rw2D}) = 2.91. \tag{29}$$

The distribution under consideration for the country-level precision parameter is $\lambda_c \sim \mathcal{N}(\mu, b)$, with μ and b parameters that are assigned the values

$$\mu = 7, \qquad b = 2, \qquad \alpha = 0.001. \tag{30}$$

Here, b is the adjusted parameter to which we must apply the correct scaling. From (29) and (30), the upper bounds for the one-dimensional and two-dimensional second-order IGMRFs are

$$\begin{aligned} U_{rw2} &= \left(\frac{b\sigma^2_{\text{ref}}(\mathbf{u}_{rw2})}{F^{-1}(\alpha, \mu, 1)} \right)^{1/2} = 7.5. \\ U_{rw2D} &= \left(\frac{b\sigma^2_{\text{ref}}(\mathbf{u}_{rw2D})}{F^{-1}(\alpha, \mu, 1)} \right)^{1/2} = 2.08. \end{aligned} \tag{31}$$

Gamma quantiles have no negative values, so to reproduce earlier results (Sørbye & Rue, 2011) using a Gaussian rather than a Gamma distribution, we need to proceed as if we have truncation below at zero. It has been shown (Gelman, 2006; Lunn, Spiegelhalter, Thomas, & Best, 2009; Roos & Held, 2011) that overfitting is common using Gamma priors and so a half-Gaussian prior for the *SD* is suggested as a suitable alternative. By taking the median, we have

$$\text{median}(U_{rw2}, U_{rw2D}) = U = 4.79. \tag{32}$$

The new *SD* parameters for the hyperpriors are:

$$\begin{aligned} b_{rw2} &= \frac{U^2 F^{-1}(\alpha, \mu, 1)}{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2})} = 0.81 \\ b_{rw2D} &= \frac{U^2 F^{-1}(\alpha, \mu, 1)}{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2D})} = 10.59. \end{aligned} \tag{33}$$

Alternatively, knowing $b_{rw2} = 0.81$,

$$b_{rw2D} = b_{rw2} \frac{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2})}{\sigma^2_{\text{ref}}(\mathbf{u}_{rw2D})} = 10.59.$$

We also observe clear patterns as the adjusted parameter, b, is varied. In earlier work on blood pressure modeling (Danaei et al., 2011), the *SD* of a one-dimensional second-order IGMRF, b_{rw2}, was set to 3. In the case of five nodes, scaling makes this equivalent to 5.01 for the two-dimensional case, b_{rw2D}, while the adjusted parameter, b, is also equal to 3, as shown in Table 3. We can see variations in the tuning of b_{rw1}, b_{rw2}, and b_{rw2D} as the number of nodes changes, and in particular cases, they each coincide with the adjusted parameter b. For example, we see that the adjusted parameter, b is equal to b_{rw1} when the number of nodes is 10; for five nodes, the adjusted parameter is equal to b_{rw2}; for 30 nodes, the adjusted parameter is equal to b_{rw2D}. Table 4 shows the scaling applied to *SD*s for a model with 11 nodes. With an adjusted parameter $b = 0.90$, scaling λ_c gives 0.53 and 1.84 for the one- and two-dimensional second-order models, respectively, for example.

TABLE 3 Scaling *SD* parameters as the adjusted parameter b and number of nodes are varied.

	$b = 1$			$b = 2$			$b = 3$		
Nodes	b_{rw1}	b_{rw2}	b_{rw2D}	b_{rw1}	b_{rw2}	b_{rw2D}	b_{rw1}	b_{rw2}	b_{rw2D}
5	0.39	1.00	1.67	0.78	2.00	3.34	1.17	3.00	5.01
6	0.48	1.00	1.84	0.96	2.00	3.67	1.43	3.00	5.49
8	0.78	1.00	2.40	1.55	2.00	4.80	2.33	3.00	7.19
10	1.00	0.82	2.58	2.00	1.63	5.16	3.00	2.45	7.74
12	1.00	0.59	2.22	2.00	1.17	4.43	3.00	1.76	6.65
14	1.00	0.43	1.94	2.00	0.87	3.89	3.00	1.30	5.83
16	1.00	0.33	1.70	2.00	0.67	3.40	3.00	1.00	5.09
18	1.00	0.27	1.54	2.00	0.53	3.08	3.00	0.80	4.62
20	1.00	0.22	1.40	2.00	0.43	2.80	3.00	0.65	4.20
25	1.00	0.14	1.12	2.00	0.28	2.25	3.00	0.42	3.37
30	1.06	0.10	1.00	2.11	0.21	2.00	3.17	0.31	3.00
40	1.40	0.08	1.00	2.80	0.15	2.00	4.20	0.23	3.00
50	1.74	0.06	1.00	3.48	0.12	2.00	5.23	0.18	3.00
100	3.46	0.03	1.00	6.93	0.06	2.00	10.39	0.09	3.00

TABLE 4 Scaling *SD* parameters for one- and two-dimensional intrinsic Gaussian Markov random fields with 11 nodes.

λ	b	b_{rw1}	b_{rw2}	b_{rw2D}
λ_c	0.90	0.77	0.53	1.84
λ_r	1.20	1.03	0.71	2.45
λ_s	1.59	1.36	0.94	3.24
λ_g	3.55	3.04	2.10	7.24

Together, these results allowed us to apply the same degree of smoothness in the two-dimensional second-order case as in the earlier work, scaling b_{rw2D} correctly given the variation in the blood pressure data.

5 | SUMMARY AND FUTURE WORK

We have shown the importance of correct scaling of hyperpriors for precision parameters in IGMRFs. This scaling varies with the dimensionality, order, and size of the IGMRFs, and also depends strongly on the structure of the precision matrices. We have presented general results in two dimensions, and specific applications to one- and two-dimensional models of blood pressure using real data.The implementation of the calculations was carried out using R, with the code being publicly available online (Spyropoulou, 2023a) for both the one- and two-dimensional cases.

Future work could include applying PC priors, which are invariant to transformations, to the precision parameter for two-dimensional random effects (Simpson et al., 2017) which corresponds to a second-order IGMRF, with $\mathbf{u} \sim \mathcal{N}(0, \lambda^{-1}\mathbf{P}^{-1})$. However, specific steps would need to be made as the priors need to be specified using the four principles defined for PC priors (Simpson et al., 2017) and also to refer to a covariance matrix of a bivariate nature. In addition, a model can have two types of random effects, constructed, and unconstructed. Therefore, they can have dependent precision parameters, and so a joint bivariate or multivariate distribution should express this dependence (Simpson et al., 2017). For the two-dimensional second-order IGMRF, the precision parameter is univariate but we could investigate the use of multivariate PC priors, which have the property that no further scaling is required as the number of nodes is varied.

ACKNOWLEDGMENTS

Specialist and High Performance Computing systems were provided by Information Services at the University of Kent. We would also like to thank the editor and the reviewers of our paper.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Maria-Zafeiria Spyropoulou https://orcid.org/0000-0001-8992-7133
James Bentham https://orcid.org/0000-0001-5733-9417

REFERENCES

Danaei, G., Finucane, M. M., Lin, J. K., Singh, G. M., Paciorek, C. J., Cowan, M. J., … , Ezzari, M. (2011). National, regional, and global trends in systolic blood pressure since 1980: Systematic analysis of health examination surveys and epidemiological studies with 786 country-years and 5·4 million participants. *Lancet*, *377*(9765), 568–577.

Finucane, M. M., Paciorek, C. J., Danaei, G., & Ezzati, M. (2014). Bayesian estimation of population-level trends in measures of health status. *Statistical Science*, *29*(1), 18–25.

Gelman, A. (2006). Multilevel (hierarchical) modeling: What it can and cannot do. *Technometrics*, *48*(3), 432–435.

Lindgren, F., & Rue, H. (2008). On the second-order random walk model for irregular locations. *Scandinavian Journal of Statistics*, *35*(4), 691–700.

Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(4), 423–498.

Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique and future directions. *Statistics in Medicine*, *28*(25), 3049–3067.

Roos, M., & Held, L. (2011). Sensitivity analysis in Bayesian generalized linear mixed models for binary data. *Bayesian Analysis*, *6*(2), 259–278.

Rue, H., & Held, L. (2005). *Gaussian Markov random fields: Theory and applications*. New York, NY: CRC Press.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., & Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, *32*(1), 1–28.

Sørbye, S. H., & Rue, H. (2011). Simultaneous credible bands for latent Gaussian models. *Scandinavian Journal of Statistics*, *38*(4), 712–725.

Sørbye, S. H., & Rue, H. (2014). Scaling intrinsic Gaussian Markov random field priors in spatial modelling. *Spatial Statistics*, *8*, 39–51.

Spyropoulou, M.-Z. (2023b). *Bayesian hierarchical modelling for two-dimensional blood pressure data*. (PhD thesis). University of Kent.

Spyropoulou, M.-Z. (2023a). *Scaling priors of IGMRF in 2D*. Retrieved from https://github.com/mariza22

Terzopoulos, D. (1988). The computation of visible-surface representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *10*(4), 417–438.

Thon, K., Rue, H., Skrøvseth, S. O., & Godtliebsen, F. (2012). Bayesian multiscale analysis of images modeled as Gaussian Markov random fields. *Computational Statistics & Data Analysis*, *56*(1), 49–61.

Yue, Y., & Speckman, P. L. (2010). Nonstationary spatial Gaussian Markov random fields. *Journal of Computational and Graphical Statistics*, *19*(1), 96–116.

How to cite this article: Spyropoulou, M. -Z., & Bentham, J. (2023). Scaling priors for intrinsic Gaussian Markov random fields applied to blood pressure data. *Statistica Neerlandica*, 1–14. https://doi.org/10.1111/stan.12330

APPENDIX A

$$
\begin{aligned}
P\left(\frac{\sigma^2_{\text{ref}}(\mathbf{u})}{\lambda} > U^2\right) = \alpha \Rightarrow P\left(\lambda < \frac{\sigma^2_{\text{ref}}(\mathbf{u})}{U^2}\right) = \alpha \\
F_\lambda\left(\frac{\sigma^2_{\text{ref}}(\mathbf{u})}{U^2}\right) = \alpha \Rightarrow \frac{\sigma^2_{\text{ref}}(\mathbf{u})}{U^2} = F^{-1}_\lambda(\alpha) \\
U^2 = \frac{\sigma^2_{\text{ref}}(\mathbf{u})}{F^{-1}_\lambda(\alpha,\mu,b)} = \frac{\sigma^2_{\text{ref}}(\mathbf{u})}{(1/b)F^{-1}_\lambda(\alpha,\mu,1)} \\
U = \left(\frac{b\sigma^2_{\text{ref}}(\mathbf{u})}{F^{-1}_\lambda(\alpha,\mu,1)}\right)^{1/2}. \qquad \text{(A1)}
\end{aligned}
$$