Accepted Manuscript

Efficient likelihood computations for some multivariate Gaussian Markov random fields

L. Ippoliti, R.J. Martin, L. Romagnoli

PII: S0047-259X(17)30486-4
DOI: https://doi.org/10.1016/j.jmva.2018.07.007
Reference: YJMVA 4387

To appear in: *Journal of Multivariate Analysis*

Received date : 13 August 2017

Please cite this article as: L. Ippoliti, R.J. Martin, L. Romagnoli, Efficient likelihood computations for some multivariate Gaussian Markov random fields, *Journal of Multivariate Analysis* (2018), https://doi.org/10.1016/j.jmva.2018.07.007

Efficient likelihood computations for some multivariate Gaussian Markov random fields

L. Ippoliti[a,*], R.J. Martin[a], L. Romagnoli[b]

[a]*Dipartimento di Economia, Università degli Studi G. d'Annunzio, Viale Pindaro 42, 65127 Pescara, Italy*
[b]*Dipartimento di Economia, Università degli Studi del Molise, Via F. De Sanctis, 86100 Campobasso, Italy*

Abstract

Data collected from spatial locations are often multivariate. Gaussian conditional autoregressive (CAR) models, also known as Gaussian Markov random fields, are frequently used to analyze such continuous data, or as models for the parameters of discrete distributions. Two difficulties in Gaussian maximum likelihood estimation are ensuring that the parameter estimates are allowable values, and computing the likelihood efficiently. It is shown here that, for some commonly-used multivariate CAR models, checking for allowable parameter values can be facilitated, and the likelihood can be computed very quickly.

Keywords: Conditional autoregressive model, Gaussian Markov random fields, Lattice data, Maximum likelihood estimation, Multivariate observations, Regional data.

AMS 2000 subject classifications: 62H11, 62H12, 62H35

1. Introduction

Large amounts of essentially continuous spatial data are associated with the nodes or interiors of a regular rectangular lattice, or with irregularly spaced sites or irregularly shaped regions. For example, pixellated images are associated with the interiors of rectangular lattices, and some spatial sampling is at rectangular grid points; in contrast, epidemiological, ecological and environmental data are usually associated with irregular sites or regions. Frequently, there is more than one variable of interest. Spatial data observed at a few time points can also be treated as multivariate.

It is often reasonable to use, possibly after transformation, a Gaussian distribution to model continuous data. Gaussian models are also frequently used in hierarchical modeling of the parameters of discrete models, such as the Poisson and Logistic. Henceforth only Gaussian models are considered, which are defined by their mean and variance structure. These can be directly specified, but the likelihood can then be difficult to obtain. An alternative specification, which is considered here, and frequently used in applications, is that of conditional autoregressive (CAR) models, also known as Gaussian Markov random fields (GMRFs); see, e.g., Section 6.3.2 of [6] or [20]. These CAR models specify the mean and variance of the values at a site in terms of the values of a set, usually small, of nearby spatial sites.

For lattice data, the number of sites can be very large, and models for the mean-corrected observations are usually simplified, e.g., taken to be (approximately) stationary, or homogeneous. Several sets of dependence neighborhoods can be included, each with an associated parameter.

Many simple models with few parameters have been proposed for multivariate CARs on irregular regions. Examples are [8, 11–13, 18, 22]. In these papers, a simple form for the spatial dependence is used, based on a 0/1 neighborhood adjacency matrix. For irregular regional (or areal) data, the number of sites is often not large, with some models implying that both the strength of the dependence and the conditional variance vary with the number of neighbors.

*Corresponding author
Email addresses: ippoliti@unich.it (L. Ippoliti), luca.romagnoli@unimol.it (L. Romagnoli)

In both situations, it can be difficult to check if the model parameters are allowable, or equivalently if the resulting variance matrix is positive definite (pd). Moreover, Gaussian maximum likelihood estimation can be difficult and time consuming. In Section 2, general multivariate CAR models are discussed and compared, and a simplified additive model is proposed. This model helps to give a unification and clarification of many of the regional models that have been proposed. A technical result in Section 3 is used to consider some situations in which checking for pd of multivariate CARs, and likelihood evaluation, can be efficiently carried out. For irregular regions (Section 4.1), the results generalize the model of [11]. For the regular rectangular lattice (Section 4.2), the results can be applied to models that are multivariate extensions of commonly-used univariate CARs — those that are reversible and have reflection-symmetric neighbor dependence. In both cases, the theory provides an extremely fast likelihood evaluation with a large number of sites. Two real data sets are used in Section 5 to illustrate the results, and some possible extensions are discussed in Section 6. Some examples and additional information are given in the Appendix.

2. Preliminaries

In this section, some initial definitions are given, and some notation introduced. Two different forms of the multivariate CAR, and two different data orderings, are compared and linked. Under certain assumptions about the variance structure, a useful additive version of the general model is developed.

2.1. Definitions and notation

It is assumed without loss of generality that the random vector $\mathbf{y}$ has been mean-corrected, $\mathrm{E}(\mathbf{y}) = \mathbf{0}$, and that it is modeled solely by a Gaussian multivariate CAR with no measurement error. Provided they are consistent, only conditional means and variances need to be specified.

In general, if the index set of $\mathbf{y}$ is partitioned as a $\cup$ b, with $\mathbf{y}$ partitioned as $(\mathbf{y}_\mathrm{a}^\top \ \mathbf{y}_\mathrm{b}^\top)^\top$, and the precision matrix $\mathbf{P} = \{\mathrm{var}(\mathbf{y})\}^{-1}$ is correspondingly partitioned, then

$$\mathrm{var}(\mathbf{y}_\mathrm{a} \,|\, \mathbf{y}_\mathrm{b}) = \mathbf{P}_{\mathrm{aa}}^{-1} \quad \text{and} \quad \mathrm{E}(\mathbf{y}_\mathrm{a} \,|\, \mathbf{y}_\mathrm{b}) = \mathrm{E}(\mathbf{y}_\mathrm{a}) + \mathbf{P}_{\mathrm{aa}}^{-1}\mathbf{P}_{\mathrm{ab}}\{\mathbf{y}_\mathrm{b} - \mathrm{E}(\mathbf{y}_\mathrm{b})\}. \tag{1}$$

Conversely, $\mathrm{var}(\mathbf{y}_\mathrm{a} \,|\, \mathbf{y}_\mathrm{b})$ and $\mathrm{E}(\mathbf{y}_\mathrm{a} \,|\, \mathbf{y}_\mathrm{b})$ essentially specify $\mathbf{P}_{\mathrm{aa}}$ and $\mathbf{P}_{\mathrm{ab}}$. Usually, for CAR models, set 'a' is small and most elements of $\mathbf{P}_{\mathrm{ab}}$ are zero.

Suppose there are n sites (used generally to include regions), with p variables observed at each site, and that $y_{i,j}$ denotes the observation on variable $i \in \{1, \ldots, p\}$ at site $j \in \{1, \ldots, n\}$. Assume that there are no missing values. Let $\mathbf{Y} = (y_{i,j}) = (\mathbf{y}_1 \cdots \mathbf{y}_n)$ be the $p \times n$ matrix of observations. Column j of $\mathbf{Y}$, $\mathbf{y}_j = (y_{1,j} \cdots y_{p,j})^\top$, is the vector of the p values at site j. Row i of $\mathbf{Y}$ contains the n observations for variable i.

Let $\mathbf{y}_{(\mathrm{s})} = (\mathbf{y}_1^\top \cdots \mathbf{y}_n^\top)^\top = \mathrm{vec}(\mathbf{Y})$ denote the np-vector of p-variate observations ordered (consistently) by site. Let $\mathrm{var}(\mathbf{y}_{(\mathrm{s})}) = \mathbf{V}_{(\mathrm{s})}$, and let $\mathbf{P}_{(\mathrm{s})} = \mathbf{V}_{(\mathrm{s})}^{-1}$ denote the $np \times np$ precision matrix of $\mathbf{y}_{(\mathrm{s})}$. Let the $\mathbf{F}_{jj'}$ be $p \times p$ matrices with (i, i')-element $f_{jj',ii'}$, where $f_{jj,ii} = -1$, and let $\mathbf{F}$ be the $np \times np$ matrix with (j, j')-block $\mathbf{F}_{jj'}$. Let the $\mathbf{T}_1, \ldots, \mathbf{T}_n$ be symmetric $p \times p$ pd matrices with (i, i')-element $\tau_{j,ii'}$, and let $\mathbf{H}_j = \mathbf{T}_j^{-1} = (\eta_{j,ii'})$. Assume $\mathbf{P}_{(\mathrm{s})} = \mathrm{diag}(\mathbf{H}_1, \ldots, \mathbf{H}_n)$ times $-\mathbf{F}$, with $\mathbf{V}_{(\mathrm{s})} = -\mathbf{F}^{-1} \times \mathrm{diag}(\mathbf{T}_1, \ldots, \mathbf{T}_n)$. Then $|\mathbf{P}_{(\mathrm{s})}| = |-\mathbf{F}| \times \prod_j |\mathbf{H}_j|$, and $\mathbf{P}_{(\mathrm{s})}$ is pd if and only if $-\mathbf{F}$ is pd. The $\mathbf{T}_j$ (or $\mathbf{H}_j$) and the $\mathbf{F}_{jj'}$ usually depend on parameters, and assumptions are necessary to reduce their total possible number, and to ensure $\mathbf{P}_{(\mathrm{s})}$ is symmetric and pd.

The observations can alternatively be ordered (consistently) by variable rather than by site. Let $\mathbf{y}_{(\mathrm{v})}$ denote the vector of observations in variable order, so that $\mathbf{y}_{(\mathrm{v})} = \mathrm{vec}(\mathbf{Y}^\top)$. Let $\mathbf{V}_{(\mathrm{v})} = \mathrm{var}(\mathbf{y}_{(\mathrm{v})})$, with $\mathbf{P}_{(\mathrm{v})} = \mathbf{V}_{(\mathrm{v})}^{-1}$. Finally, let $\mathbf{J}_p$ denote a $p \times p$ matrix of ones.

2.2. Two specifications of a multivariate CAR

Univariate CARs ($p = 1$) are defined by the conditional distributions of the variable at each site given the values at all the other sites. There are two possible generalizations of this to multivariate CARs ($p > 1$). The more usual generalization [15], denoted the mvCAR form here, considers the n conditional distributions $\mathbf{y}_j | y_{-j}$ of the p-vector variable $\mathbf{y}_j$ at each site j given y_{-j}, where y_{-j} denotes the values $\mathbf{y}_{j'}$ at all other sites, i.e., $\{\mathbf{y}_{j'} : j' \neq j\}$. The other generalization, as in [13], the extended CAMCAR model in Section 5 of [21] and the model of [22], and Ippoliti et al. [10], denoted the uvCAR form here, considers the np individual conditional distributions $y_{i,j} | y_{-(i,j)}$ of each variable at

each site, where $y_{-(i,j)}$ denotes all other values than $y_{i,j}$, i.e., $\{y_{i',j'} : (i', j') \neq (i, j)\}$. Hence the uvCAR form can be regarded as a univariate CAR over np 'sites' (i.e., n sites for each variable, which can be pictured as stacked in an extra dimension). From Eq. (1) and its converse, it is clear that the two forms are just different ways of parametrizing the same multivariate CAR model. One form may sometimes be more convenient to specify than the other.

The mvCAR form has

$$\mathrm{E}(\mathbf{y}_j|y_{-j}) = \sum_{j' \neq j} \mathbf{F}_{jj'}\mathbf{y}_{j'} \quad \text{and} \quad \mathrm{var}(\mathbf{y}_j|y_{-j}) = \mathbf{T}_j\,,$$

where $\mathbf{F}_{jj} = -\mathbf{I}_p$ for all j, and the conditional variance $\mathbf{T}_j = \mathbf{H}_j^{-1}$ is in general not diagonal. Then the (j, j') block of $\mathbf{P}_{(\mathrm{s})}$ is $-\mathbf{H}_j\mathbf{F}_{jj'}$, and symmetry of $\mathbf{P}_{(\mathrm{s})}$ requires that $\mathbf{H}_j\mathbf{F}_{jj'} = \mathbf{F}_{j'j}^\top\mathbf{H}_{j'}$, or equivalently $\mathbf{T}_j\mathbf{F}_{j'j}^\top = \mathbf{F}_{jj'}\mathbf{T}_{j'}$ for all $j \neq j'$.

The uvCAR form has

$$\mathrm{E}(y_{i,j}|y_{-(i,j)}) = \sum_{(i',j')\neq(i,j)} f_{jj',ii'}y_{i',j'} \quad \text{and} \quad \mathrm{var}(y_{i,j}|y_{-(i,j)}) = \tau_{j,ii} = \eta_{j,ii}^{-1},$$

with $\mathbf{F}_{jj}$ non-diagonal, but $\mathbf{T}_j = \mathbf{H}_j^{-1}$ is diagonal. Then $\mathbf{P}_{(\mathrm{s})}$ has $\{(j-1)p + i, (j'-1)p + i'\}$-element $-\eta_{j,ii}f_{jj',ii'}$, and symmetry of $\mathbf{P}_{(\mathrm{s})}$ requires that $\eta_{j,ii}f_{jj',ii'} = \eta_{j',i'i'}f_{j'j,i'i}$ for all i, i', j, j'.

Note that when $p > 1$ the symmetry requirement on $\mathbf{P}_{(\mathrm{s})}$ usually implies that the conditional means $\mathrm{E}(y_{i,j}|y_{-(i,j)})$ or $\mathrm{E}(\mathbf{y}_j|y_{-j})$, and hence $\mathbf{F}$, cannot simply be specified; it is also necessary to take account of the conditional variances.

Given $\mathbf{P}_{(\mathrm{s})}$, the mvCAR form can be obtained by grouping the p rows and columns corresponding to each site, and the uvCAR form is obtained by considering each row of $\mathbf{P}_{(\mathrm{s})}$. Observe that the within-site cross-dependences $\eta_{j,ii}f_{jj,ii'}$ in the uvCAR form correspond to off-diagonal terms in $\mathbf{T}_j$ (or $\mathbf{H}_j$) in the mvCAR form, and vice versa. This is illustrated in Example A in Appendix A.1.

2.3. Some simplified models

The general model requires simplification. A special case of the mvCAR form has $\mathbf{F}_{jj'} = f_{jj'}\mathbf{I}_p$ with $f_{jj} = -1$ for all j, where $f_{jj'}$ for $j \neq j'$ may be a linear combination of parameters, i.e., $f_{jj',ii'} = f_{jj'}$ if $i = i'$, and 0 otherwise. Then $\mathbf{F} = \mathbf{F}_0 \otimes \mathbf{I}_p$, where $\mathbf{F}_0 = (f_{jj'})$. Symmetry of $\mathbf{P}_{(\mathrm{s})}$ in this case requires that $f_{jj'}\mathbf{H}_j = f_{j'j}\mathbf{H}_{j'}$, or $f_{j'j}\mathbf{T}_j = f_{jj'}\mathbf{T}_{j'}$, for all $j \neq j'$. The separable (or factorized) case has $\mathbf{F}_0$ symmetric, with $\mathbf{T}_j = \mathbf{T} = (\tau_{ii'})$ and $\mathbf{H}_j = \mathbf{H} = (\eta_{ii'})$ for all j, so that $\mathbf{P}_{(\mathrm{s})} = -\mathbf{F}_0 \otimes \mathbf{H}$. The uvCAR form of the separable case has $f_{jj',ii'} = f_{jj'}\eta_{ii'}/\eta_{ii}$ and $\mathrm{var}(y_{i,j}|y_{-(i,j)}) = \eta_{ii}^{-1}$.

More generally, it is often reasonable to assume in the mvCAR form that the conditional variance $\mathbf{T}_j$ is either constant, or only varies over sites by a constant, so that $\mathbf{T}_j = w_{0,j}^{-1}\mathbf{T}$ and $\mathbf{H}_j = w_{0,j}\mathbf{H}$ for all j. If $\mathbf{W}_0 = \mathrm{diag}(w_{0,1}, \ldots, w_{0,n})$, then the $p \times p$ diagonal blocks of $\mathbf{P}_{(\mathrm{s})}$ are $\mathbf{W}_0 \otimes \mathbf{H}$. To include the uvCAR form with $\tau_{j,ii} = \tau_{ii}$ for all j, replace $\mathbf{W}_0 \otimes \mathbf{H}$ by $-\mathbf{W}_0 \otimes \mathbf{\Phi}_0$, where $\mathbf{\Phi}_0 = \mathbf{H}\mathbf{A}_0$, with $\mathbf{A}_0$ and $\mathbf{\Phi}_0$ non-singular. In the mvCAR form $\mathbf{A}_0 = -\mathbf{I}_p$, while in the uvCAR form $\mathbf{H}$ is the diagonal $\mathbf{H}_\mathrm{D} = \mathrm{diag}(\eta_{11}, \ldots, \eta_{pp})$ and $\mathbf{A}_0$ is $-\mathbf{I}_p$ plus some non-zero off-diagonal terms. Many models that have been proposed with this assumption can be expressed in the following way.

With $\mathbf{W}_0$ and $\mathbf{\Phi}_0$ as above, suppose the $\mathbf{W}_k$ for $k \in \{1, \ldots, k_{\max}\}$ are $n \times n$ matrices of known constants, with the entries of $\mathbf{W}_k$ relating to the kth group of 'neighbors'. Usually the $\mathbf{W}_k$ for $k > 0$ have zeros on the diagonal, and for $k \geq 0$ their non-zero entries are non-overlapping, i.e., $\mathbf{W}_k \circ \mathbf{W}_{k'} = 0$ for $0 \leq k < k'$, where $\circ$ denotes the Hadamard (or Schur) element-wise product. Suppose also that the $p \times p$ matrices $\mathbf{A}_k = (\alpha_{k,ii'})$, $k \geq 0$, contain parameters relating to the cross-dependences of the variables at a site with those in the kth group, with $\alpha_{0,ii} = -1$ for all i. Let $\mathbf{\Phi}_k = (\phi_{k,ii'})$ denote $\mathbf{H}\mathbf{A}_k$, and let $\mathbf{M}_{(\mathrm{s})}$ be an $np \times np$ matrix, with $\mathbf{P}_{(\mathrm{s})} = \mathbf{M}_{(\mathrm{s})}^\top\mathbf{R}_{(\mathrm{s})}\mathbf{M}_{(\mathrm{s})}$. It is assumed here that $\mathbf{M}_{(\mathrm{s})}$ is diagonal, $\mathbf{M}_{(\mathrm{s})} = \mathrm{diag}(\mathbf{M}_{(\mathrm{s}),1}, \ldots, \mathbf{M}_{(\mathrm{s}),n})$, and the $\mathbf{M}_{(\mathrm{s}),j}$ are diagonal $p \times p$ matrices of known constants relating to the variables at site j. Then, $\mathbf{R}_{(\mathrm{s})}$ is expressed as an additive model, viz.

$$\mathbf{R}_{(\mathrm{s})} = (\mathbf{I}_n \otimes \mathbf{H})\left\{-\sum_{k=0}^{k_{\max}}(\mathbf{W}_k \otimes \mathbf{A}_k)\right\} = -\sum_{k=0}^{k_{\max}}(\mathbf{W}_k \otimes \mathbf{\Phi}_k). \tag{2}$$

In many cases $\mathbf{M}_{(\mathrm{s})} = \mathbf{I}_{np}$, so that $\mathbf{P}_{(\mathrm{s})} = \mathbf{R}_{(\mathrm{s})}$. Symmetry of $\mathbf{P}_{(\mathrm{s})}$ is ensured if $\mathbf{\Phi}_k$ is symmetric whenever $\mathbf{W}_k$ is symmetric, and, if a $\mathbf{W}_k$ is asymmetric then the second sum also includes $\mathbf{W}_k^\top \otimes \mathbf{\Phi}_k^\top$, i.e., the first sum includes both $\mathbf{W}_k \otimes \mathbf{A}_k$ and $\mathbf{W}_k^\top \otimes \mathbf{H}^{-1}\mathbf{A}_k^\top\mathbf{H}$. If k_2 is the number of asymmetric pairs in Eq. (2), then the maximum number of 'free' parameters in $\mathbf{R}_{(\mathrm{s})}$ is $(1 + k_{\max})p(p+1)/2 - pk_2$.

Provided the $\mathbf{W}_k$, $k > 0$, have zeros on the diagonal, Eq. (2) implies that

$$\mathrm{E}(\mathbf{y}_j|y_{-j}) = -(w_{0,j}\mathbf{A}_0\mathbf{M}_{(\mathrm{s}),j})^{-1}\sum_{j'\neq j}\sum_{k>0} w_{k,jj'}\mathbf{A}_k\mathbf{M}_{(\mathrm{s}),j'}\mathbf{y}_{j'}$$

and

$$\mathrm{var}(\mathbf{y}_j|y_{-j}) = -(w_{0,j}\mathbf{M}_{(\mathrm{s}),j}\mathbf{H}\mathbf{A}_0\mathbf{M}_{(\mathrm{s}),j})^{-1} = -(w_{0,j}\mathbf{M}_{(\mathrm{s}),j}\mathbf{\Phi}_0\mathbf{M}_{(\mathrm{s}),j})^{-1}.$$

Thus the choice of $\mathbf{M}_{(\mathrm{s})}$ affects both the conditional means and the conditional variances.

For the uvCAR form, the off-diagonal elements of $\mathbf{A}_0$ are the within-site dependence parameters between the variables. If $\mathbf{M}_{(\mathrm{s})} = \mathbf{I}_{np}$ and $\mathbf{W}_0 = \mathbf{I}_n$, then $\mathbf{T}_j = \mathbf{T}$ and $\mathbf{H}_j = \mathbf{H}$ for all j, and the conditional variance $\mathrm{var}(\mathbf{y}_j|y_{-j})$ does not depend on the site. If $\mathbf{R}_{(\mathrm{s})}$ in Eq. (2) is separable with $\mathbf{A}_k = \alpha_k\mathbf{I}_p$ (where $\alpha_0 = -1$), then $\mathbf{R}_{(\mathrm{s})} = -\mathbf{F}_0 \otimes \mathbf{H}$, with $\mathbf{F}_0 = \sum_{k=0}^{k_{\max}} \alpha_k\mathbf{W}_k$. If any $\mathbf{W}_k$, $k > 0$, have non-zero diagonal elements, as for the model of [12]; see Section 2.4, or the reflective boundary conditions in Section 4.2.1, adjustments must be made to $\mathrm{E}(\mathbf{y}_j|y_{-j})$ and $\mathrm{var}(\mathbf{y}_j|y_{-j})$ using Eq. (1). Example B in Appendix A.1 gives a simple illustration.

If $\mathbf{A}_0 = -\mathbf{I}_p$, an alternative form of Eq. (2) has also been used in regional modeling, which makes symmetry arise more naturally, but also affects the interpretation of the parameters in the $\mathbf{A}_k$. This has $\mathbf{R}_{\mathrm{S},(\mathrm{s})}$, where

$$\mathbf{R}_{\mathrm{S},(\mathrm{s})} = (\mathbf{I}_n \otimes \mathbf{H}^{1/2})\left\{-\sum_{k=0}^{k_{\max}}(\mathbf{W}_k \otimes \mathbf{A}_k)\right\}\{\mathbf{I}_n \otimes (\mathbf{H}^{1/2})^\top\} = -\sum_{k=0}^{k_{\max}}(\mathbf{W}_k \otimes \mathbf{\Phi}_{\mathrm{S},k}) \tag{3}$$

where $\mathbf{H}^{1/2}$ is any 'square root' of $\mathbf{H}$ such that $\mathbf{H}^{1/2}(\mathbf{H}^{1/2})^\top = \mathbf{H}$ (unique and symmetric if $\mathbf{H} = \mathbf{H}_{\mathrm{D}}$), and where $\mathbf{\Phi}_{\mathrm{S},k}$ denotes $\mathbf{H}^{1/2}\mathbf{A}_k(\mathbf{H}^{1/2})^\top$. Symmetry of $\mathbf{R}_{\mathrm{S},(\mathrm{s})}$ is ensured if $\mathbf{A}_k$ is symmetric whenever $\mathbf{W}_k$ is symmetric, and, for each asymmetric $\mathbf{W}_k$ the sum also includes a term $\mathbf{W}_k^\top \otimes \mathbf{A}_k^\top$.

2.4. Applications of the additive model

On a regular rectangular $n_1 \times n_2$ lattice ($n = n_1n_2$), homogeneous (approximately stationary) models can be postulated using the mvCAR form with $\mathbf{T}_j = \mathbf{T}$ for all j. Then the usual model is Eq. (2) with $\mathbf{M}_{(\mathrm{s})} = \mathbf{I}_{np}$, $\mathbf{W}_0 = \mathbf{I}_n$, and for $k > 0$ the $\mathbf{W}_k$ are neighbor-incidence matrices for the kth group of lags (i.e., 0/1 matrices with (j, j') element 1 if site j' is separated from site j by a lag in the kth group); see Section 4.2.

Univariate CAR regional models have usually based dependence on some function of neighborhood — adjacency, length of common boundary, distance between sites or regional centres, etc. These assumptions are almost always arbitrary and unrealistic [17]. Multivariate CAR regional models have usually been based on $\mathbf{W}$, the symmetric regional neighbor-incidence matrix, i.e., the (j, j') element of $\mathbf{W}$ is 1 if regions j and j' are adjacent. Some recent generalizations have split $\mathbf{W}$ into two parts $\mathbf{W} = \mathbf{W}_{\mathrm{U}} + \mathbf{W}_{\mathrm{U}}^\top$, arbitrarily using its upper and lower triangular parts, $\mathbf{W}_{\mathrm{U}}$ and $\mathbf{W}_{\mathrm{L}} = \mathbf{W}_{\mathrm{U}}^\top$.

To allow the conditional variance to decrease with the number of neighbors, $\mathbf{W}_0$ is often taken as the diagonal matrix $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_n) = \mathrm{diag}(\mathbf{d})$, where $\mathbf{d} = \mathbf{W}\mathbf{1}_n$, i.e., d_j is the jth row sum of $\mathbf{W}$. Then $\mathbf{H}_j = d_j\mathbf{H}$ if $\mathbf{A}_0 = -\mathbf{I}_p$. A consequence is that the conditional correlations between neighboring sites j, j' are approximately proportional to $(d_jd_{j'})^{-1/2}$. Note that those given by [11] only consider the conditional distributions of the two relevant variables, and not that of the $2p$ variables at the two sites.

An alternative to taking $\mathbf{W}_0 = \mathbf{D}$ is to use $\mathbf{M}_{(\mathrm{s})}$ with $\mathbf{W}_0 = \mathbf{I}_n$ and the $\mathbf{M}_{(\mathrm{s}),j}$ being precision measures for site j; see Section 2.3. A possible advantage is that the neighbor conditional correlations are then constant over the sites; see [21, 22]. A special case for $\mathbf{M}_{(\mathrm{s})}$ is the Kronecker product $\mathbf{M}_{\mathrm{s}} \otimes \mathbf{M}_{\mathrm{v}}$ of a diagonal $n \times n$ matrix $\mathbf{M}_{\mathrm{s}}$ of site constants and a diagonal $p \times p$ matrix $\mathbf{M}_{\mathrm{v}}$ of constants relating to the variables, possibly with $\mathbf{M}_{\mathrm{s}} = \mathbf{I}_n$, or $\mathbf{M}_{\mathrm{v}} = \mathbf{I}_p$.

Some CAR models that have been proposed for regional data are discussed in Appendix A.2. Note that in the original proposals a wide range of different notation has been used, and protracted derivations given for relatively simple and similar models. To allow comparisons between them, the models have been reparametrized. Henceforth, except where otherwise stated, it is assumed that if $\mathbf{M}_{(\mathrm{s})} \neq \mathbf{I}_{np}$, then $\mathbf{M}_{(\mathrm{s})} = \mathbf{M}_{\mathrm{s}} \otimes \mathbf{M}_{\mathrm{v}}$. Then $\mathbf{M}_{(\mathrm{s})}$ can, if desired, be assimilated into the additive Eqs. (2)–(3) by suitably redefining the $\mathbf{\Phi}_k$ or the $\mathbf{A}_k$, or the $\mathbf{W}_k$.

Note that the marginal distribution of each variable in a separable multivariate CAR is a univariate CAR, but it is clear from Eqs. (2)–(3) that for a general multivariate CAR the marginal distributions are not CARs. On a regular rectangular lattice they are Rational Spectral Density models; see [9].

2.5. Ordering observations by variables

Although the mvCAR form is defined by site ordering, it is often more convenient to consider the model for the observations ordered by variables. This ordering clearly shows the relevant contribution of each variable and the joint contributions of pairs of variables. Most of the models for regional data have been given this way. The two orderings are connected by the $np \times np$ permutation matrix $\mathbf{Z}$, where $\mathbf{Z}^\top = \mathbf{Z}^{-1} = (\mathbf{I}_n \otimes \mathbf{e}_1 \cdots \mathbf{I}_n \otimes \mathbf{e}_p)$, and $\mathbf{e}_i$ is a p-vector with 1 in the ith position and zero elsewhere.

Then $\mathbf{y}_{(\mathrm{v})} = \mathbf{Z}\mathbf{y}_{(\mathrm{s})}$, and $\mathbf{y}_{(\mathrm{s})} = \mathbf{Z}^\top\mathbf{y}_{(\mathrm{v})}$. Hence $\mathbf{V}_{(\mathrm{v})} = \mathbf{Z}\mathbf{V}_{(\mathrm{s})}\mathbf{Z}^\top$, $\mathbf{V}_{(\mathrm{s})} = \mathbf{Z}^\top\mathbf{V}_{(\mathrm{v})}\mathbf{Z}$, $\mathbf{P}_{(\mathrm{v})} = \mathbf{Z}\mathbf{P}_{(\mathrm{s})}\mathbf{Z}^\top$, and $\mathbf{P}_{(\mathrm{s})} = \mathbf{Z}^\top\mathbf{P}_{(\mathrm{v})}\mathbf{Z}$. Further, if $\mathbf{B}_1$ is $p\times p$ and $\mathbf{B}_2$ is $n\times n$, then $\mathbf{Z}^\top(\mathbf{B}_1\otimes\mathbf{B}_2)\mathbf{Z} = \mathbf{B}_2\otimes\mathbf{B}_1$. Hence if $\mathbf{V}_{(\mathrm{s})}$ (or $\mathbf{P}_{(\mathrm{s})}$) is expressed as $\sum a_k\mathbf{B}_{2,k}\otimes\mathbf{B}_{1,k}$ for scalar a_k with all $\mathbf{B}_{1,k}$ being $p \times p$, and all $\mathbf{B}_{2,k}$ being $n \times n$, then $\mathbf{V}_{(\mathrm{v})}$ (or $\mathbf{P}_{(\mathrm{v})}$) $= \sum a_k\mathbf{B}_{1,k} \otimes \mathbf{B}_{2,k}$. In particular, the separable model has $\mathbf{P}_{(\mathrm{v})} = -\mathbf{H} \otimes \mathbf{F}_0$. Henceforth, variable ordering is assumed, and the suffix (v) is dropped.

If $\mathbf{T}_j = w_{0,j}^{-1}\mathbf{T}$ for all j, the general additive model (2) with the restriction on $\mathbf{M}_{(\mathrm{s})}$ becomes

$$\mathbf{P} = (\mathbf{M}_\mathrm{v} \otimes \mathbf{M}_\mathrm{s})\mathbf{R}(\mathbf{M}_\mathrm{v} \otimes \mathbf{M}_\mathrm{s}),$$

with

$$\mathbf{R} = (\mathbf{H} \otimes \mathbf{I}_n)\left\{-\sum_{k=0}^{k_{\max}}(\mathbf{A}_k \otimes \mathbf{W}_k)\right\} = -\sum_{k=0}^{k_{\max}}(\mathbf{\Phi}_k \otimes \mathbf{W}_k). \tag{4}$$

The alternative form (3) uses $\mathbf{R}_\mathrm{S}$, with

$$\mathbf{R}_\mathrm{S} = (\mathbf{H}^{1/2} \otimes \mathbf{I}_n)\left\{-\sum_{k=0}^{k_{\max}}(\mathbf{A}_k \otimes \mathbf{W}_k)\right\}\left\{(\mathbf{H}^{1/2})^\top \otimes \mathbf{I}_n\right\} = -\sum_{k=0}^{k_{\max}}(\mathbf{\Phi}_{\mathrm{S},k} \otimes \mathbf{W}_k). \tag{5}$$

Then the $n \times n$ submatrices of $\mathbf{R}$ are $\mathbf{R}_{ii'} = -\sum_{k=0}^{k_{\max}}\phi_{k,ii'}\mathbf{W}_k$, and those of $\mathbf{R}_\mathrm{S}$ are $\mathbf{R}_{\mathrm{S},ii'} = -\sum_{k=0}^{k_{\max}}\phi_{\mathrm{S},k,ii'}\mathbf{W}_k$ for all $i, i' \in \{1, \ldots, p\}$.

2.6. Additive model with symmetric components

Although some models (see Appendix A.2) do use asymmetric components $\mathbf{W}_k$ in Eqs. (2)–(3) and (4)–(5), many models in common use have all the $\mathbf{W}_k$ symmetric. It is assumed here henceforth, apart from the torus boundary assumption in Section 4.2, that this holds.

Then with Eq. (4), symmetry of $\mathbf{P}$ is ensured if $\mathbf{\Phi}_k = \mathbf{H}\mathbf{A}_k$ is symmetric for all k, i.e., if $\mathbf{A}_k^\top\mathbf{H} = \mathbf{H}\mathbf{A}_k$, or $\mathbf{T}\mathbf{A}_k^\top = \mathbf{A}_k\mathbf{T}$. There are many ways to ensure this holds, essentially specifying a symmetric matrix $\mathbf{H}_{\mathrm{A},k}$ and setting $\mathbf{A}_k = \mathbf{H}^{-1}\mathbf{H}_{\mathrm{A},k}$. With an initial $\mathbf{A}_k^*$, one possibility is to modify the lower triangular part of $\mathbf{A}_k^*$ to ensure that $\mathbf{H}\mathbf{A}_k^*$ is symmetric. Another possibility is to use $\mathbf{H}_{\mathrm{A},k} = (\mathbf{H}\mathbf{A}_k^* + \mathbf{A}_k^{*\top}\mathbf{H})/2$. A method used in simple regional models has a symmetric $\mathbf{A}_k^*$, and $\mathbf{H}_{\mathrm{A},k} = \mathbf{H} \circ \mathbf{A}_k^*$.

The resulting $\mathbf{A}_k$ may be difficult to interpret. Although it does not directly give the conditional means, it may be simpler to specify the $\mathbf{\Phi}_k$. Then $\mathbf{\Phi}_0 = \mathbf{H}\mathbf{A}_0$ gives $\mathbf{H}$ as $-\mathbf{\Phi}_0$ in the mvCAR form, and $-\mathrm{diag}(\phi_{0,11}, \ldots, \phi_{0,pp})$ in the uvCAR form, with $\mathbf{A}_k = \mathbf{H}^{-1}\mathbf{\Phi}_k$ for $k \geq 0$.

With Eq. (5), symmetry of $\mathbf{P}$ only needs the $\mathbf{A}_k$ to be symmetric, but these may be difficult to interpret as the conditional mean $\mathrm{E}(\mathbf{y}_j|y_{-j})$ now involves terms in $\{\mathbf{A}_0(\mathbf{H}^{1/2})^\top\}^{-1}\mathbf{A}_k(\mathbf{H}^{1/2})^\top\mathbf{y}_{j'}$. Again, it may be simpler to specify the $\mathbf{\Phi}_{\mathrm{S},k}$, with $\mathbf{H}$ obtained as above, but using $\mathbf{\Phi}_{\mathrm{S},0}$ instead of $\mathbf{\Phi}_0$, and $\mathbf{A}_k = \mathbf{H}^{-1/2}\mathbf{\Phi}_{\mathrm{S},k}(\mathbf{H}^{-1/2})^\top$ for $k \geq 0$.

3. Calculating the Gaussian likelihood and checking that P is positive definite

This section is concerned with Gaussian maximum likelihood estimation of multivariate CAR models. If $\mathbf{P}$ is expressed as in Eqs. (4)–(5) in terms of the $\mathbf{\Phi}_k$ or $\mathbf{\Phi}_{\mathrm{S},k}$, then the usual result for CARs holds; maximum likelihood estimation equates sample and population correlations for each $k > 0$, although this result is rarely useful in practice. The standard method requires minimization of $\mathbf{y}^\top\mathbf{P}\mathbf{y} - \ln(|\mathbf{P}|)$, which has two main difficulties. First, the minimization must be over the valid range of values for the parameters (i.e., the estimated $\mathbf{P}$ must be pd). Second, the evaluation of $|\mathbf{P}|$ and/or of $\mathbf{y}^\top\mathbf{P}\mathbf{y}$, can be very slow. Note that it may be simpler to estimate $\mathbf{P}$ in terms of the $\mathbf{\Phi}_k$ or $\mathbf{\Phi}_{\mathrm{S},k}$, and then transform back to $\mathbf{H}$ and the $\mathbf{A}_k$; see Section 2.6.

Suppose that $\mathbf{R}_1$ and $\mathbf{R}_2$ are $np \times np$ matrices, and that $\mathbf{P} = \mathbf{R}_1\mathbf{R}_c\mathbf{R}_2$. Then $|\mathbf{P}| = |\mathbf{R}_1| \times |\mathbf{R}_2| \times |\mathbf{R}_c|$, and the pd conditions are those on $\mathbf{R}_c$ together with those on $\mathbf{R}_1$ and $\mathbf{R}_2$. If $\mathbf{u}_1 = \mathbf{R}_1^\top\mathbf{y}$ and $\mathbf{u}_2 = \mathbf{R}_2\mathbf{y}$ then $\mathbf{y}^\top\mathbf{P}\mathbf{y} = \mathbf{u}_1^\top\mathbf{R}_c\mathbf{u}_2$. If $\mathbf{R}_c$ is partitioned, then using standard results may be quicker for obtaining $|\mathbf{R}_c|$. For example, if $\mathbf{R}_c$ is partitioned into 2×2 components by a and b, as in Section 2.1, then $|\mathbf{R}_c| = |\mathbf{R}_{c,a}| \times |\mathbf{R}_{c,bb.a}|$, where $\mathbf{R}_{c,bb.a}$ denotes $\mathbf{R}_{c,bb} - \mathbf{R}_{c,ba}\mathbf{R}_{c,aa}^{-1}\mathbf{R}_{c,ab}$, and $\mathbf{R}_c$ is pd if and only if both $\mathbf{R}_{c,aa}$ and $\mathbf{R}_{c,bb.a}$ are pd. If the partitioned parts are conformable, as when the $p \times p$ $\mathbf{R}_{c,ii'}$ with $i, i' \in \{1, \ldots, p\}$, are used, then $|\mathbf{R}_c| = |\mathbf{R}_{c,11}\mathbf{R}_{c,22.1}|$ for $p = 2$, etc. The relevant determinants can be evaluated directly, or as the square of the product of the diagonal elements of the Cholesky decomposition (usually quicker), or as the product of the eigenvalues (usually slowest if the eigenvalues are computed, quickest if the eigenvalues are known). If further the components all commute, then the usual determinant formula applies to the $\mathbf{R}_{c,ii'}$, e.g., if $p = 2$, $|\mathbf{R}_c| = |\mathbf{R}_{c,11}\mathbf{R}_{c,22} - \mathbf{R}_{c,12}\mathbf{R}_{c,21}|$. However, in this latter case, there is a much quicker method.

Lemma 1 below, and Corollaries 1 and 2, can be used to show that if all the $\mathbf{R}_{c,ii'}$ commute, there is a direct way to obtain the conditions for $\mathbf{R}_c$ to be pd, and a very fast way to obtain $|\mathbf{R}_c|$, and to evaluate $\mathbf{y}^\top\mathbf{P}\mathbf{y}$ when $\mathbf{R}_1 = \mathbf{R}_2^\top$, with $\mathbf{R}_1$ fixed.

Lemma 1. *Suppose the real $np \times np$ matrix $\mathbf{Q}$ is partitioned into $n \times n$ blocks $\mathbf{Q}_{ii'}$ with $i, i' \in \{1, \ldots, p\}$, and that the $\mathbf{Q}_{ii'}$ commute. Let the unitary $\mathbf{S} = (\mathbf{s}_1 \ldots \mathbf{s}_n)$ be such that $\mathbf{Q}_{ii'} = \mathbf{S}\mathbf{\Theta}_{ii'}\mathbf{S}^{-1}$, where the $\mathbf{\Theta}_{ii'}$ can be taken as upper triangular with the eigenvalues $\theta_{ii',j}$ on the diagonal, matched appropriately. Then the eigenvalues of $\mathbf{Q}$ are those of the $p \times p$ matrices $\mathbf{C}_1, \ldots, \mathbf{C}_n$, where $\mathbf{C}_j$ has (i, i')-element $\theta_{ii',j}$.*

Proof. Since $\mathbf{Q}_{ii'} = \mathbf{S}\mathbf{\Theta}_{ii'}\mathbf{S}^{-1}$, it follows that $\mathbf{Q} = (\mathbf{I}_p \otimes \mathbf{S})\mathbf{\Theta}(\mathbf{I}_p \otimes \mathbf{S}^{-1})$, where $\mathbf{\Theta}$ is the $np \times np$ matrix with blocks $\mathbf{\Theta}_{ii'}$. Thus $\mathbf{Q}$ and $\mathbf{\Theta}$ are similar (conjugate), their eigenvalues are the same, and $|\mathbf{Q}| = |\mathbf{\Theta}|$. But, since the $\mathbf{\Theta}_{ii'}$ are upper triangular, $\mathbf{\Theta}$ can be permuted into $\mathbf{Z}\mathbf{\Theta}\mathbf{Z}^\top$, which is block upper triangular with the blocks $\mathbf{C}_j$ on the diagonal. □

Remark. In most applications in Section 4, the $\mathbf{Q}_{ii'}$ are all symmetric, or if not symmetric they do not have repeated eigenvalues. Then the $\mathbf{Q}_{ii'}$ are simultaneously diagonalizable, so that $\mathbf{s}_j$ is the right eigenvector of $\mathbf{Q}_{ii'}$ with corresponding eigenvalue $\theta_{ii',j}$, i.e., $\mathbf{Q}_{ii'}\mathbf{s}_j = \theta_{ii',j}\mathbf{s}_j$. Then $\mathbf{Q}_{ii'} = \mathbf{S}\mathbf{\Theta}_{ii'}\mathbf{S}^{-1}$, where $\mathbf{\Theta}_{ii'} = \mathrm{diag}(\theta_{ii',1}, \ldots, \theta_{ii',n})$.

Corollary 1. *$\mathbf{Q}$ in Lemma 1 is pd if and only if all the $\mathbf{C}_j$ are pd, and $|\mathbf{Q}| = \prod_j |\mathbf{C}_j|$, so that only the $p \times p$ matrices $\mathbf{C}_j$ need to be considered. Each $\mathbf{C}_j$ is pd if and only if all its eigenvalues are positive, or equivalently if and only if all its principal minors are positive.*

Note that the eigenvalues of the $\mathbf{C}_j$ are not needed if only $|\mathbf{Q}|$ is required. Thus when $p = 2$, $|\mathbf{Q}| = \prod_j(\theta_{11,j}\theta_{22,j} - \theta_{12,j}\theta_{21,j})$, and the conditions on the minors of the $\mathbf{C}_j$s mean that $\mathbf{Q}$ is pd if and only if $\theta_{11,j} > 0, \theta_{22,j} > 0$, and $\theta_{11,j}\theta_{22,j} > \theta_{12,j}\theta_{21,j}$ for all $j \in \{1, \ldots, n\}$.

Corollary 2. *If $\mathbf{Q}$ in Lemma 1 is symmetric, the quadratic form $\mathbf{q}^\top\mathbf{Q}\mathbf{q}$ can be evaluated as $\sum_j \mathbf{c}_j^\top\mathbf{C}_j\mathbf{c}_j$, where $\mathbf{c} = (\mathbf{c}_1^\top \cdots \mathbf{c}_n^\top)^\top = \mathbf{Z}(\mathbf{I}_p \otimes \mathbf{S}^\top)\mathbf{q} = \mathbf{Z}(\mathbf{q}_1^\top\mathbf{S} \cdots \mathbf{q}_n^\top\mathbf{S})^\top$.*

This result is useful if $\mathbf{R}_1 = \mathbf{R}_2^\top$, and $\mathbf{u} = \mathbf{R}_2\mathbf{y}$ so that $\mathbf{y}^\top\mathbf{P}\mathbf{y} = \mathbf{u}^\top\mathbf{R}_c\mathbf{u}$, and $\mathbf{R}_1$ is fixed.

4. Applications of Lemma 1 to multivariate CARs

Applications of Lemma 1, and Corollaries 1, 2, are considered for some regional models in Section 4.1, and for some regular rectangular lattice models in Section 4.2. The $\mathbf{R}_{ii'}$ or $\mathbf{R}_{S,ii'}$ in Section 2.5 commute if the $\mathbf{W}_k$ commute, $k \geq 0$. If $\mathbf{W}_0 \neq \mathbf{I}_n$ and the $\mathbf{W}_0^{-1}\mathbf{W}_k$ commute, the result can also be used on

$$\mathbf{R}_T = (\mathbf{I}_p \otimes \mathbf{W}_0^{-1})\mathbf{P} = -\sum_{k=0}^{k_{\max}} \mathbf{\Phi}_k \otimes (\mathbf{W}_0^{-1}\mathbf{W}_k)$$

or similarly, if the $\mathbf{W}_0^{-1/2}\mathbf{W}_k\mathbf{W}_0^{-1/2}$ commute on

$$\mathbf{R}_L = (\mathbf{I}_p \otimes \mathbf{W}_0^{-1/2})\mathbf{P}(\mathbf{I}_p \otimes \mathbf{W}_0^{-1/2}) = -\sum_{k=0}^{k_{\max}} \mathbf{\Phi}_k \otimes (\mathbf{W}_0^{-1/2}\mathbf{W}_k\mathbf{W}_0^{-1/2}).$$

In the latter two cases, the submatrices $\mathbf{R}_{\mathrm{T},ii'} = -\sum_{k=0}^{k_{\max}} \phi_{\mathrm{k},ii'} \mathbf{W}_0^{-1}\mathbf{W}_k$ or $\mathbf{R}_{\mathrm{L},ii'} = -\sum_{k=0}^{k_{\max}} \phi_{\mathrm{k},ii'} \mathbf{W}_0^{-1/2}\mathbf{W}_k\mathbf{W}_0^{-1/2}$ commute. Then the appropriate $\mathbf{W}_k$ or $\mathbf{W}_0^{-1}\mathbf{W}_k$ or $\mathbf{W}_0^{-1/2}\mathbf{W}_k\mathbf{W}_0^{-1/2}$ can be upper triangularized, with corresponding eigenvalues ν_{kj}, and the $\mathbf{C}_j$ of Lemma 1 are $\mathbf{C}_j = -\sum_{k=0}^{k_{\max}} \nu_{kj}\mathbf{\Phi}_k$. This always applies if $k_{\max} = 1$ in Eqs. (4)–(5), as for the simple regional models; see Section 4.1. Then $|\mathbf{P}| = |\mathbf{W}_0|^p \times |\mathbf{R}_{\mathrm{T}}| = |\mathbf{W}_0|^p \times |\mathbf{R}_{\mathrm{L}}|$.

Note that if, using Lemma 1, the eigenvalues of $\mathbf{R}$ are $\{\theta_{ii,j}\}$, then the necessary conditions $\theta_{ii,j} > 0$ for all $i \in \{1, \ldots, p\}$ are precisely the necessary and sufficient conditions for the univariate CARs defined by the $\mathbf{R}_{ii}$, and the conditions $\theta_{ii,j}\theta_{i'i',j} > \theta_{ii',j}\theta_{i'i,j}$ are the additional ones for the bivariate CARs defined by the $\mathbf{R}_{ii'}$, $i \neq i'$, etc., and similarly for $\mathbf{R}_{\mathrm{S}}$ or $\mathbf{R}_{\mathrm{T}}$ or $\mathbf{R}_{\mathrm{L}}$.

4.1. Irregular sites and regions

Some simple models for regional data have $k_{\max} = 1$ with $\mathbf{W}_1$ and the $\mathbf{\Phi}_0$, $\mathbf{\Phi}_1$ symmetric. The resulting model $\mathbf{P} = -(\mathbf{M}_{\mathrm{v}} \otimes \mathbf{M}_{\mathrm{s}})(\mathbf{\Phi}_0 \otimes \mathbf{W}_0 + \mathbf{\Phi}_1 \otimes \mathbf{W}_1)(\mathbf{M}_{\mathrm{v}} \otimes \mathbf{M}_{\mathrm{s}})$ can be represented in several ways by taking out $\mathbf{H}$ and/or $\mathbf{W}_0$ (one side) or $\mathbf{H}^{1/2}$ and/or $\mathbf{W}_0^{1/2}$ (both sides). Let $\mathbf{R}_{\mathrm{R}} = -(\mathbf{\Psi}_0 \otimes \mathbf{I}_n + \mathbf{\Psi}_1 \otimes \mathbf{E}_2)$, and first consider the symmetric form

$$\mathbf{P} = (\mathbf{E}_0 \otimes \mathbf{E}_1)\mathbf{R}_{\mathrm{R}}(\mathbf{E}_0^\top \otimes \mathbf{E}_1^\top) \tag{6}$$

for a $p \times p$ pd $\mathbf{E}_0$ and $p \times p$ symmetric $\mathbf{\Psi}_k$ with $\mathbf{E}_0^\top\mathbf{\Psi}_k\mathbf{E}_0 = \mathbf{M}_{\mathrm{v}}\mathbf{\Phi}_k\mathbf{M}_{\mathrm{v}}$; and for $n \times n$ $\mathbf{E}_1$ and $\mathbf{E}_2$, with $\mathbf{E}_1\mathbf{E}_1^\top = \mathbf{M}_{\mathrm{s}}\mathbf{W}_0\mathbf{M}_{\mathrm{s}}$, $\mathbf{E}_1\mathbf{E}_2\mathbf{E}_1^\top = \mathbf{M}_{\mathrm{s}}\mathbf{W}_1\mathbf{M}_{\mathrm{s}}$. Then $|\mathbf{P}| = |\mathbf{E}_0|^{2n} \times |\mathbf{M}_{\mathrm{s}}|^{2p} \times |\mathbf{W}_0|^p \times |\mathbf{R}_{\mathrm{R}}|$, and, using Corollary 2, $\mathbf{y}^\top\mathbf{P}\mathbf{y}$ can be obtained as $\mathbf{u}^\top\mathbf{R}_{\mathrm{R}}\mathbf{u}$ with $\mathbf{u} = (\mathbf{E}_0^\top \otimes \mathbf{E}_1^\top)\mathbf{y}$.

A more general representation is

$$\mathbf{P} = (\mathbf{E}_3 \otimes \mathbf{E}_4)\mathbf{R}_{\mathrm{R}}(\mathbf{E}_5 \otimes \mathbf{E}_6) \tag{7}$$

where now the $\mathbf{\Psi}_k$ need not be symmetric, and $\mathbf{E}_3\mathbf{\Psi}_k\mathbf{E}_5 = \mathbf{M}_{\mathrm{v}}\mathbf{\Phi}_k\mathbf{M}_{\mathrm{v}}$, $\mathbf{E}_4\mathbf{E}_6 = \mathbf{M}_{\mathrm{s}}\mathbf{W}_0\mathbf{M}_{\mathrm{s}}$, and $\mathbf{E}_4\mathbf{E}_2\mathbf{E}_6 = \mathbf{M}_{\mathrm{s}}\mathbf{W}_1\mathbf{M}_{\mathrm{s}}$. Then $|\mathbf{P}| = |\mathbf{E}_3\mathbf{E}_5|^n \times |\mathbf{M}_{\mathrm{s}}|^{2p} \times |\mathbf{W}_0|^p \times |\mathbf{R}_{\mathrm{R}}|$.

In both representations, the pd conditions on $\mathbf{R}_{\mathrm{R}}$ usually give the pd conditions on $\mathbf{P}$, possibly after taking into account the conditions on $\mathbf{H}$.

Result 1. If the eigenvalues of $\mathbf{E}_2$ are ω_j, then by Lemma 1, the eigenvalues of $\mathbf{R}_{\mathrm{R}} = -(\mathbf{\Psi}_0 \otimes \mathbf{I}_n + \mathbf{\Psi}_1 \otimes \mathbf{E}_2)$ are those of the $\mathbf{C}_j = -(\mathbf{\Psi}_0 + \omega_j\mathbf{\Psi}_1) = -\mathbf{\Psi}_0(\mathbf{I}_p + \omega_j\mathbf{\Psi}_0^{-1}\mathbf{\Psi}_1)$ with $j \in \{1, \ldots, n\}$, i.e., $c_{j,ii'} = -(\psi_{0,ii'} + \omega_j\psi_{1,ii'})$.

4.1.1. The model of Jin et al. [11]

The model of Jin et al. [11] has $\mathbf{P}_{ii'} = \eta_{ii'}(\mathbf{D} - \alpha_{ii'}\mathbf{W})$, i.e., $\mathbf{M}_{\mathrm{s}} = \mathbf{I}_{np}$ and $\mathbf{P} = \mathbf{H} \otimes \mathbf{D} - (\mathbf{H} \circ \mathbf{A}) \otimes \mathbf{W}$. When first suggested it was considered "difficult to check the conditions guaranteeing positive definiteness" [12]. However, this $\mathbf{P}$ was obtained in [11] by taking a linear transform of a model which has $\mathbf{H} = \mathbf{I}_p$, but their pd conditions do not directly apply to $\mathbf{A} = (\alpha_{ii'})$.

The model of [11] was also derived in [18] using a transformation from independent univariate CARs. These derivations are useful if the model is to be used as a prior parameter model. However, if the model is being used directly on continuous regional data, it may be easier and more natural to consider the final model. It is shown here that using Corollary 1 gives a direct derivation of $|\mathbf{P}|$, and suggests a different, more meaningful, method for checking that the parameters are allowable. Moreover, Corollary 2 gives a dramatic reduction in computer time when fitting the model.

Following Eq. (6), $\mathbf{P}$ can be expressed as

$$\mathbf{P} = (\mathbf{I}_p \otimes \mathbf{E}_1)\{\mathbf{H} \otimes \mathbf{I}_n - (\mathbf{H} \circ \mathbf{A}) \otimes \mathbf{E}_2\}(\mathbf{I}_p \otimes \mathbf{E}_1),$$

with $\mathbf{R}_{\mathrm{R},ii'} = \eta_{ii'}(\mathbf{I}_n - \alpha_{ii'}\mathbf{E}_2)$, where $\mathbf{E}_1 = \mathbf{D}^{1/2}$, $\mathbf{E}_2 = \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$. The form of $\mathbf{E}_2$ implies that $\omega_{\max} = \max(\omega_j) = 1$ and $-1 \leq \omega_{\min} < 0$, where $\omega_{\min} = \min(\omega_j)$. If $\mathbf{H}_{\mathrm{A}} = \mathbf{H}^{-1}(\mathbf{H} \circ \mathbf{A})$, then Result 1 can be used with $\mathbf{C}_j = \mathbf{H} - \omega_j\mathbf{H} \circ \mathbf{A} = \mathbf{H}(\mathbf{I}_p - \omega_j\mathbf{H}_{\mathrm{A}}) = \mathbf{H} \circ (\mathbf{J}_p - \omega_j\mathbf{A})$ for all $j \in \{1, \ldots, n\}$, i.e., $c_{j,ii'} = \eta_{ii'}(1 - \omega_j\alpha_{ii'})$.

Another expression for $\mathbf{P}$ in the form of Eq. (6), essentially the one used in [11], is

$$\mathbf{P} = (\mathbf{E}_0 \otimes \mathbf{E}_1)(\mathbf{I}_{np} - \mathbf{H}_{\mathrm{A,S}} \otimes \mathbf{E}_2)(\mathbf{E}_0^\top \otimes \mathbf{E}_1),$$

where $\mathbf{E}_0 = \mathbf{H}^{1/2}$, and $\mathbf{H}_{\mathrm{A,S}}$ denotes $(\mathbf{H}^{1/2})^{-1}(\mathbf{H} \circ \mathbf{A})\{(\mathbf{H}^{1/2})^\top\}^{-1}$.

Using either of these two forms for $\mathbf{P}$ allows $\mathbf{y}^\top\mathbf{P}\mathbf{y}$, for differing $\mathbf{P}$, to be obtained quicker as $\mathbf{u}^\top\mathbf{R}_{\mathrm{R}}\mathbf{u}$; see the comment below Eq. (6).

4.1.2. Checking that P for the model of [11] is positive-definite

There are several ways to check if $\mathbf{P}$ is pd, and to evaluate $|\mathbf{P}|$. The first way uses $\mathbf{C}_j = \mathbf{H}(\mathbf{I}_p - \omega_j \mathbf{H}_\mathrm{A})$ and the eigenvalues $\mu_1, \ldots, \mu_p$ of $\mathbf{H}_\mathrm{A}$. Then the eigenvalues of $\mathbf{I}_p - \omega_j \mathbf{H}_\mathrm{A}$ are $1 - \omega_j \mu_i$, so that $|\mathbf{P}| = |\mathbf{H}|^n \times |\mathbf{D}|^p \times \prod_{i,j}(1 - \omega_j \mu_i)$. $\mathbf{P}$ is pd if and only if all the $\mathbf{C}_j$ are pd, which is equivalent to $\mathbf{I}_p - \omega_j \mathbf{H}_\mathrm{A}$ being pd for all j. Thus $\mathbf{P}$ is pd if and only if $\omega_j \mu_i < 1$ for all i and j. This implies that $1/\omega_{\min} < \mu_i < 1/\omega_{\max}$ for all i, i.e., that $\min(\mu_i) > 1/\omega_{\min}$, and $\max(\mu_i) < 1/\omega_{\max}$.

The conditions on the ω_js in terms of the μ_is are less useful, and slightly more complicated as it is possible for $\min(\mu_i)$ to be positive, or $\max(\mu_i)$ to be negative; this occurs in Example A in Appendix A.4. Thus they are usually $\omega_{\min} > 1/\min(\mu_i)$ and $\omega_{\max} < 1/\max(\mu_i)$, but if $\min(\mu_i) > 0$ then it is only $\omega_{\max} < 1/\max(\mu_i)$, and if $\max(\mu_i) < 0$ then it is only $\omega_{\min} > 1/\min(\mu_i)$. Since these pd conditions involve the eigenvalues of $\mathbf{H}_\mathrm{A}$, it is not straightforward to obtain the conditions on the $\alpha_{ii'}$s.

The form $\mathbf{R}_\mathrm{R} = \mathbf{I}_{np} - \mathbf{H}_{\mathrm{A,S}} \otimes \mathbf{E}_2$ was essentially used in [11]. Then $\mathbf{P}$ is pd provided $\mathbf{I}_{np} - \mathbf{H}_{\mathrm{A,S}} \otimes \mathbf{E}_2$ is pd. Since the eigenvalues of $\mathbf{H}_{\mathrm{A,S}}$ are identical to those of $\mathbf{H}_\mathrm{A}$, the conditions on $\mathbf{P}$ and $|\mathbf{P}|$ are the same as those above. This implies that the alternative constraint in Section 3.3 of [11] is in general incorrect.

The method of [18] expresses $\mathbf{P}$ as $(\mathbf{H}^{1/2} \otimes \mathbf{E}_1)(\mathbf{I}_{np} - \mathbf{A}^* \otimes \mathbf{E}_2)\{(\mathbf{H}^{1/2})^\top \otimes \mathbf{E}_1\}$, where $\mathbf{A}^*$ is diagonal, and the diagonal elements are allowable parameters for one-dimensional CARs. It is shown in Appendix A.3 how this assumption fits into the present framework. This approach is compared with the direct and the approaches of [11] in Example C in Appendix A.4.

A different way to check $\mathbf{P}$ is pd uses Lemma 1 and the $\mathbf{C}_j$ directly. Then $|\mathbf{P}| = |\mathbf{D}|^p \times \prod_j |\mathbf{C}_j|$, and $\mathbf{P}$ is pd if and only if all the principal minors of $\mathbf{C}_j$ are positive for all $j \in \{1, \ldots, n\}$. In particular, it is necessary, for all $i \in \{1, \ldots, p\}$, that $\alpha_{ii}\omega_j < 1$, for each ω_j, i.e., that $1/\omega_{\min} < \alpha_{ii} < 1/\omega_{\max}$ (the conditions for the univariate CARs); that all 2×2 minors are positive, i.e., that $(1 - \alpha_{ii}\omega_j)(1 - \alpha_{i'i'}\omega_j) > r_\eta^2 (1 - \alpha_{ii'}\omega_j)^2$ for all $i \neq j$, where r_η^2 denotes $\eta_{12}^2/(\eta_{11}\eta_{22})$ with $0 < r_\eta^2 < 1$ (the conditions for the bivariate CARs); and, if $p > 2$, that $|\mathbf{C}_j| > 0$. An advantage of this way is that it is easy to see theoretically what the restrictions are on the parameters, and that this can be considered sequentially for the univariate CARs, the bivariate CARs, the trivariate CARs, etc.

Note that in general, the μ_i satisfy $|\mathbf{H} - (\mathbf{H} \circ \mathbf{A})\mu_i^{-1}| = 0$, together with $\mu_i \omega_j < 1$, whereas the condition $|\mathbf{C}_j| > 0$ is $|\mathbf{H} - (\mathbf{H} \circ \mathbf{A})\omega_j| > 0$. Thus the end points of the intervals of ω_j for which $|\mathbf{C}_j| > 0$ are given by the $1/\mu_i$.

In particular, if $p = 2$, then

$$|\mathbf{C}_j| = (\eta_{11}\eta_{22})\{(1 - \alpha_{11}\omega_j)(1 - \alpha_{22}\omega_j) - r_\eta^2 (1 - \alpha_{12}\omega_j)^2\} = |\mathbf{H}| - \mathrm{tr}(\mathbf{H}_\mathrm{A})\omega_j + |(\mathbf{H} \circ \mathbf{A})|\omega_j^2.$$

If $\mathbf{H}$ is given, then the pd conditions for the α_{ij}s are

$$(1 - \alpha_{11}\omega_j)(1 - \alpha_{22}\omega_j) - r_\eta^2 (1 - \alpha_{12}\omega_j)^2 > 0 \quad \text{for all } \omega_j,\, j \in \{1, \ldots, n\}.$$

If also the α_{ii} are given and satisfy $1/\omega_{\min} < \alpha_{ii} < 1/\omega_{\max}$ for $i \in \{1, 2\}$, the conditions for α_{12} can be written

$$\{(1 - \alpha_{11}\omega_j)(1 - \alpha_{22}\omega_j)/r_\eta^2 - 1\} + 2\omega_j\alpha_{12} - \omega_j^2\alpha_{12}^2 > 0 \quad \text{for all } \omega_j,\, j \in \{1, \ldots, n\}.$$

If the $\alpha_{ii}\omega_{\max}$ are not too far from 1, the allowable region for α_{12} may result from using just $\omega_j = \omega_{\max}$. Otherwise, additional constraints may arise from $\omega_{\min}$. For the usual regional model with $\omega_{\max} = 1$, it is necessary that $\{(1 - \alpha_{11})(1 - \alpha_{22})/r_\eta^2 - 1\} + 2\alpha_{12} - \alpha_{12}^2 > 0$. If $1 - \alpha_{11}$ and $1 - \alpha_{22}$ are not 'large', this condition may also be sufficient, as in Example E in Appendix A.4. If $\omega_{\min} = -1$, then it is necessary that $\{(1 + \alpha_{11})(1 + \alpha_{22})/r_\eta^2 - 1\} - 2\alpha_{12} - \alpha_{12}^2 > 0$.

The theory is demonstrated in Examples B and C in Appendix A.4.

4.1.3. Other simple regional models

Some other simple models with $k_{\max} = 1$ are considered here. If $\mathbf{A}$ is symmetric in the CAMCAR model of [21] (see Appendix A.2), and if $\mathbf{M}_{(\mathrm{s})} = \mathbf{M}_\mathrm{s} \otimes \mathbf{M}_\mathrm{v}$, then $\mathbf{P}$ can be expressed in the form of Eq. (6) as

$$\mathbf{P} = (\mathbf{M}_\mathrm{v}\mathbf{H}^{1/2} \otimes \mathbf{M}_\mathrm{s})(\mathbf{I}_{np} - \mathbf{A} \otimes \mathbf{W})(\mathbf{H}^{1/2}\mathbf{M}_\mathrm{v} \otimes \mathbf{M}_\mathrm{s}).$$

Using Result 1, the eigenvalues of $\mathbf{R}_\mathrm{R} = \mathbf{I}_{np} - \mathbf{A} \otimes \mathbf{W}$ are those of the $\mathbf{C}_j = \mathbf{I}_p - \omega_j \mathbf{A}$, where $\omega_1, \ldots, \omega_n$ are the eigenvalues of $\mathbf{W}$.

Similarly, if $\mathbf{A}_1$ is symmetric in the model of [22] (see Appendix A.2), then $\mathbf{P}$ can be expressed in the form of Eq. (6) as

$$\mathbf{P} = -(\mathbf{H}_{\mathrm{D}}^{1/2} \otimes \mathbf{I}_n)(\mathbf{A}_0 \otimes \mathbf{I}_n + \mathbf{A}_1 \otimes \mathbf{W})(\mathbf{H}_{\mathrm{D}}^{1/2} \otimes \mathbf{I}_n),$$

with $\mathbf{R}_{\mathrm{R}} = -(\mathbf{A}_0 \otimes \mathbf{I}_n + \mathbf{A}_1 \otimes \mathbf{W})$ and $\mathbf{C}_j = -(\mathbf{A}_0 + \omega_j \mathbf{A}_1)$. For both of these simplified models, $\mathbf{y}^\top \mathbf{P} \mathbf{y}$ can be obtained as $\mathbf{u}^\top \mathbf{R}_{\mathrm{R}} \mathbf{u}$ – see the comment below Eq. (6).

The model of [10] (see Appendix A.2) expressed in the form of Eq. (7) has

$$\mathbf{P} = -(\mathbf{H}_{\mathrm{D}} \otimes \mathbf{M}_{\mathrm{s}})(\mathbf{A}_0 \otimes \mathbf{I}_n + \mathbf{A}_1 \otimes \mathbf{W})(\mathbf{I}_p \otimes \mathbf{M}_{\mathrm{s}}).$$

This has $\mathbf{R}_{\mathrm{R}}$ and the $\mathbf{C}_j$ equal to those for the model of [22] above. An illustration is given in Example F in Appendix A.4.

The model in uvCAR form with

$$\mathbf{P} = -(\mathbf{H}_{\mathrm{D}} \otimes \mathbf{I}_n)(\mathbf{A}_0 \otimes \mathbf{D} + \mathbf{A}_1 \otimes \mathbf{W}) = -(\mathbf{H}_{\mathrm{D}}\mathbf{A}_0 \otimes \mathbf{D} + \mathbf{H}_{\mathrm{D}}\mathbf{A}_1 \otimes \mathbf{W})$$

is equivalent to the model of [11], but more natural to specify and simpler to interpret. This $\mathbf{P}$ can be written in the form of Eq. (6) as

$$\mathbf{P} = -(\mathbf{I}_p \otimes \mathbf{E}_1)(\mathbf{H}_{\mathrm{D}}\mathbf{A}_0 \otimes \mathbf{I}_n + \mathbf{H}_{\mathrm{D}}\mathbf{A}_1 \otimes \mathbf{E}_2)(\mathbf{I}_p \otimes \mathbf{E}_1),$$

with $\mathbf{R}_{\mathrm{R}} = -(\mathbf{H}_{\mathrm{D}}\mathbf{A}_0 \otimes \mathbf{I}_n + \mathbf{H}_{\mathrm{D}}\mathbf{A}_1 \otimes \mathbf{E}_2)$, where $\mathbf{E}_1$ and $\mathbf{E}_2$ are as in Section 4.1.1. Thus $\mathbf{H}$ and $\mathbf{H} \circ \mathbf{A}$ in the model of [11] are replaced by $-\mathbf{H}_{\mathrm{D}}\mathbf{A}_0$ and $\mathbf{H}_{\mathrm{D}}\mathbf{A}_1$, respectively. The eigenvalues of $\mathbf{R}_{\mathrm{R}}$ are those of the $\mathbf{C}_j = -(\mathbf{H}_{\mathrm{D}}\mathbf{A}_0 + \omega_j \mathbf{H}_{\mathrm{D}}\mathbf{A}_1)$, where the ω_js are the eigenvalues of $\mathbf{E}_2$, and $\mathbf{y}^\top \mathbf{P} \mathbf{y}$ can be obtained as $\mathbf{u}^\top \mathbf{R}_{\mathrm{R}} \mathbf{u}$. However, parameter estimation appears to perform better when the $\mathbf{H}_{\mathrm{D}}\mathbf{A}_k$ in $\mathbf{R}_{\mathrm{R}}$ are replaced by $\mathbf{\Phi}_{\mathrm{S},k}$, and the $\phi_{\mathrm{S},k,ii'}$s are estimated initially; see Section 2.6. This method may give a better way to fit the model of [11].

Lemma 1 cannot in general be used for models with asymmetric $\mathbf{\Phi}_k \otimes \mathbf{W}_k$ terms in Eq. (4), or asymmetric $\mathbf{A}_k \otimes \mathbf{W}_k$ terms in Eq. (5). It also cannot in general be used for the model of [13] (unless $\mathbf{D}$ is proportional to $\mathbf{I}_n$), as the diagonal blocks $\mathbf{P}_{ii}$ are proportional to $2\mathbf{D} + \mathbf{I}_n - \alpha_{ii}\mathbf{W}$, but the off-diagonal blocks $\mathbf{P}_{ii'}$, $i \neq i'$, are proportional to $\mathbf{I}_n - \alpha_{ii'}\mathbf{W}$.

4.2. Sites on a regular rectangular lattice

Suppose the sites are on a regular rectangular $n_1 \times n_2$ lattice (n_1 rows and n_2 columns), and that they are ordered lexicographically. The order of a CAR is defined by its neighbor sets. For $p = 1$, the conditional mean of a first-order CAR(1) uses horizontal and vertical neighbors (with one or two parameters), a second-order CAR(2) also has diagonal neighbors (one or two parameters), a third-order CAR(3) includes lag (2,0) and (0,2) neighbors (one or two parameters), etc. For $p > 1$, the order is defined here as the maximum neighbor order among and between the variables.

The usual multivariate CAR models assume $\mathbf{T}_j = \mathbf{T}$ for all j, and $\mathbf{M}_{(\mathrm{s})} = \mathbf{I}_{np}$, and are given by Eq. (4), where $\mathbf{W}_0 = \mathbf{I}_n$, and for $k > 0$ each $\mathbf{W}_k$ is a Kronecker product of a within-column-lag incidence matrix and a within-row-lag incidence matrix (with the identity matrix $\mathbf{I}$ for lag 0), or a sum of such products; see Section 4.2.2. In the mvCAR form, symmetry of $\mathbf{P}$ requires that $\mathbf{H}\mathbf{A}_k = \mathbf{\Phi}_k$ is symmetric for all k, i.e., $\sum_r \eta_{ir}\alpha_{k,ri'} = \sum_r \eta_{i'r}\alpha_{k,ri}$ for all $i \neq i'$, and all k. The uvCAR form is possibly simpler, as then symmetry of $\mathbf{P}$ merely requires that $\eta_{ii}\alpha_{k,ii'} = \eta_{i'i'}\alpha_{k,i'i}$ for all $i \neq i'$, and all k.

A multivariate CAR is reversible [15] if the between-variable dependence from $\mathbf{A}_k$ at lag (g_1, g_2) is equal to that at lag $(-g_1, -g_2)$ for all lags $g_m \in \{0, \ldots, n_m - 1\}$, $m \in \{1, 2\}$, and all pairs of variables $i, i' \in \{1, \ldots, p\}$. This implies that the $\mathbf{W}_k$ in Eq. (4) are symmetric. Extending the univariate definition, a multivariate CAR with order larger than 1 is a reflection-symmetric CAR (RS-CAR) if it is reversible, and additionally if the between-variable dependence from $\mathbf{A}_k$ at lag (g_1, g_2) is equal to that at lag $(g_1, -g_2)$ for all lags $g_m \in \{0, \ldots, n_m - 1\}$, $m \in \{1, 2\}$, and all pairs of variables $i, i' \in \{1, \ldots, p\}$. This implies symmetry of both components of the Kronecker products involved in any $\mathbf{W}_k$.

Even for univariate CARs on a finite planar lattice, it is only feasible to use exact stationarity, in some restricted special cases, such as separability of the correlation function, or having a finite unilateral representation. Eq. (1) specifies the stationary form of $\mathbf{P}$ except at border sites, where the conditional mean for the infinite-lattice process involves sites outside the finite lattice. The correct conditional distributions for border sites, and also the correlations, usually cannot be simply specified. Thus the stationary $\mathbf{P}$ is difficult to obtain accurately. Also, stationarity conditions

can be very complicated if the CAR is not reflection-symmetric [14]. In practice, a non-stationary form of $\mathbf{P}$ is specified using neighbor assumptions on the $\mathbf{W}_k$, or equivalently assumptions on the conditional distributions of border sites using observed values at other sites (or 0) as proxies for unobserved ones; see Section 4.2.1. The conditions for $\mathbf{P}$ to be pd may then be slightly more relaxed than the stationarity conditions, but can also be complicated [14]. Stationarity conditions for multivariate CARs are considered in [15].

4.2.1. Boundary assumptions

Several possible boundary assumptions have been proposed for CAR models, each with some advantages and some disadvantages. The torus assumption (TBA) essentially wraps the lattice so that the first and last rows are adjacent, and similarly with the columns. The fixed-value assumption (FBA) (often called the Dirichlet boundary condition) sets the values outside the lattice to 0 (the mean), i.e., the corresponding terms are removed from the conditional mean. Balram and Moura [1] discuss two assumptions using reflections about the boundaries: the reflective (or variational or asymmetric Neumann) assumption RBA, and the symmetric Neumann condition (SBA). Besag and Kooperberg [3] present two other possibilities for a one-parameter CAR(1). A negative-reflection (NBA) possibility is presented in [19]. These boundary assumptions are discussed further and compared in Appendix A.5.

4.2.2. Using common eigenvectors for the $\mathbf{W}_k$

Under certain boundary and model assumptions, the same eigenvectors can be used for all the $\mathbf{W}_k$. Recall that the eigenvalues of a Kronecker product are all the products of the eigenvalues of the constituent matrices. For $g_2 \in \{0, \ldots, n_2 - 1\}$, and a particular boundary assumption, let $\mathbf{U}_{n_2,g_2}$ be the boundary-corrected within-row right-adjacency matrix for lag g_2, i.e., the (j_2, j_2') element of $\mathbf{U}_{n_2,g_2}$ is 1 if, in a lattice row, j_2' is a right neighbor of j_2. The standard planar neighbors have only $j_2' = j_2 + g_2$ if $j_2 + g_2 \le n_2$, which corresponds to the FBA, with $\mathbf{U}_{n_2,0} = \mathbf{I}_{n_2}$. Other boundary assumptions include other neighbors. The lag g_2 within-row left-adjacency matrix is the half-turn rotation of $\mathbf{U}_{n_2,g_2}$, i.e., $\mathbf{Z}_{n_2}\mathbf{U}_{n_2,g_2}\mathbf{Z}_{n_2}$ where the $n_m \times n_m$ reflection matrix $\mathbf{Z}_{n_m}$ has ones on the NE-SW diagonal. Similarly, let $\mathbf{U}_{n_1,g_1}$ be the within-column down-adjacency matrix for lag $g_1 \in \{0, \ldots, n_1 - 1\}$, with $\mathbf{U}_{n_1,0} = \mathbf{I}_{n_1}$, and $\mathbf{Z}_{n_1}\mathbf{U}_{n_1,g_1}\mathbf{Z}_{n_1}$ the up-column adjacency matrix. For most boundary assumptions, but not for the RBA, $\mathbf{Z}_{n_m}\mathbf{U}_{n_m,g_m}\mathbf{Z}_{n_m} = \mathbf{U}_{n_m,g_m}^\top$. Then, for $g_2 > 0$, $\mathbf{N}_{n_2,g_2} = \mathbf{U}_{n_2,g_2} + \mathbf{Z}_{n_2}\mathbf{U}_{n_2,g_2}\mathbf{Z}_{n_2}$ is the lag g_2 row-neighbor adjacency matrix, and, for $g_1 > 0$, $\mathbf{N}_{n_1,g_1} = \mathbf{U}_{n_1,g_1} + \mathbf{Z}_{n_1}\mathbf{U}_{n_1,g_1}\mathbf{Z}_{n_1}$ is the lag g_1 column-neighbor adjacency matrix. Let $\mathbf{N}_{n_m,0} = \mathbf{I}_{n_m}$ for $m \in \{1, 2\}$.

For the TBA, $\mathbf{U}_{n_m,g_m}$ includes ones in the $(n_m - g_m + j_m, j_m)$ positions, $j_m \in \{1, \ldots, g_m\}$, and the $\mathbf{U}_{n_m,g_m}$ are circulant matrices, with $\mathbf{Z}_{n_m}\mathbf{U}_{n_m,g_m}\mathbf{Z}_{n_m} = \mathbf{U}_{n_m,g_m}^\top = \mathbf{U}_{n_m,g_m}^{-1}$, and $\mathbf{U}_{n_m,g_m} = (\mathbf{U}_{n_m,1})^{g_m}$. Since the $\mathbf{U}_{n_m,g_m}$ can use the same eigenvectors, with corresponding eigenvalues $\exp(2\pi i g_m j_m/n_m)$ for $j_m \in \{0, \ldots, n_m - 1\}$ (where here $i = \sqrt{-1}$), the CAR can be completely general with the $\mathbf{W}_k$ using any of $\mathbf{U}_{n_1,g_1} \otimes \mathbf{U}_{n_2,g_2}$, $\mathbf{U}_{n_1,g_1} \otimes \mathbf{U}_{n_2,g_2}^\top$, $\mathbf{U}_{n_1,g_1}^\top \otimes \mathbf{U}_{n_2,g_2}$, $\mathbf{U}_{n_1,g_1}^\top \otimes \mathbf{U}_{n_2,g_2}^\top$. For any reversible CAR, the asymmetric $\mathbf{W}_k$ occur in pairs giving terms $\mathbf{U}_{n_1,g_1} \otimes \mathbf{U}_{n_2,g_2} + \mathbf{U}_{n_1,g_1}^\top \otimes \mathbf{U}_{n_2,g_2}^\top$, or $\mathbf{U}_{n_1,g_1} \otimes \mathbf{U}_{n_2,g_2}^\top + \mathbf{U}_{n_1,g_1}^\top \otimes \mathbf{U}_{n_2,g_2}$ with $g_m \in \{1, \ldots, n_m\}$. The corresponding eigenvalues are then $2\cos\{2\pi(g_1 j_1/n_1 + g_2 j_2/n_2)\}$ or $2\cos\{2\pi(g_1 j_1/n_1 - g_2 j_2/n_2)\}$, respectively, for $j_m \in \{0, \ldots, n_m - 1\}$. This corresponds to using the discrete Fourier transform (DFT).

Common eigenvectors can be used under many other boundary assumptions for first-order neighbors if the row neighbor adjacency matrix is $\mathbf{W}_1 = \mathbf{I}_{n_1} \otimes \mathbf{N}_{n_2,1}$, and the column one is $\mathbf{W}_2 = \mathbf{N}_{n_1,1} \otimes \mathbf{I}_{n_2}$. This does not apply to the 'average' and initial 'new strategy' of [3]. If second-order RS-neighbors are included, and the adjacency matrix is taken as $\mathbf{W}_3 = \mathbf{W}_1 \otimes \mathbf{W}_2 = \mathbf{N}_{n_1,1} \otimes \mathbf{N}_{n_2,1}$, then all three $\mathbf{W}_k$ matrices commute and can use the same eigenvectors. For the FBA, the eigenvalues of $\mathbf{N}_{n_m,1}$ are $2\cos\{\pi(j_m + 1)/(n_m + 1)\}$ for $j_m \in \{0, \ldots, n_m - 1\}$ with $m \in \{1, 2\}$. Another disadvantage of the SBA (see Appendix A.5) is that there is in general no closed formula for the eigenvalues; see Lemma 3 of [1].

Under the RBA, $\mathbf{U}_{n_m,g_m}$, $g_m > 0$, includes 1 in the $(n_m + 1 - j_m, n_m + j_m - g_m)$ positions, with $j_m \in \{1, \ldots, g_m\}$. For the NBA, $\mathbf{U}_{n_m,g_m}$, $g_m > 1$, includes -1 in the $(n_m + 1 - j_m, n_m + 1 + j_m - g_m)$ positions, for $j_m \in \{1, \ldots, g_m - 1\}$. In both cases, the $\mathbf{W}_k$ are of the form $\mathbf{N}_{n_1,g_1} \otimes \mathbf{N}_{n_2,g_2}$, and can use the same eigenvectors. For the RBA, the eigenvalues of $\mathbf{N}_{n_m,g_m}$ are $2\cos(\pi g_m j_m/n_m)$, for $j_m \in \{0, \ldots, n_m - 1\}$ and $m \in \{1, 2\}$, which corresponds to using the discrete cosine transform (DCT); see [4, 5, 19]. For the NBA, the eigenvalues of $\mathbf{N}_{n_m,g_m}$ are $2\cos\{\pi g_m j_m/(n_m+1)\}$, $j_m \in \{0, \ldots, n_m-1\}$, $m \in \{1, 2\}$, which corresponds to using the discrete sine transform (DST); see [19].

With any CAR and boundary assumptions which result in common eigenvectors for the $\mathbf{W}_k$, Lemma 1 can be used to find the conditions for $\mathbf{P}$ to be pd, theoretically for a model with few parameters, or computationally with more complicated models. This is illustrated in Example 1 below.

Example 1. Consider a RS-CAR(2). Then,

$$\mathbf{P}_{ii'} = -\phi_{0,ii'}\mathbf{I}_n - \phi_{1,ii'}\mathbf{I}_{n_1} \otimes \mathbf{N}_{n_2,1} - \phi_{2,ii'}\mathbf{N}_{n_1,1} \otimes \mathbf{I}_{n_2} - \phi_{3,ii'}\mathbf{N}_{n_1,1} \otimes \mathbf{N}_{n_2,1}.$$

If the eigenvalues of $\mathbf{N}_{n_m,1}$ are λ_{m,j_m} for all $j_m \in \{1, \ldots, n_m\}$, then, using Lemma 1, the (i, i') element of $\mathbf{C}_{j_1,j_2}$, for $j_m \in \{1, \ldots, n_m\}$, is $-\phi_{0,ii'} - \phi_{1,ii'}\lambda_{2,j_2} - \phi_{2,ii'}\lambda_{1,j_1} - \phi_{3,ii'}\lambda_{1,j_1}\lambda_{2,j_2}$. Then the eigenvalues of $\mathbf{P}$ are those of the $\mathbf{C}_{j_1,j_2}$, $\mathbf{P}$ is pd provided that all $\mathbf{C}_{j_1,j_2}$ are pd, and $|\mathbf{P}| = \prod |\mathbf{C}_{j_1,j_2}|$. □

Computer timing evaluations using Matlab [24] suggest that using the result $|\mathbf{P}| = \prod_j |\mathbf{C}_j|$ substantially reduces the time to evaluate $|\mathbf{P}|$. For example, when $p = 2$ and $n_1 = n_2 = 20$, and calculating the eigenvalues of the $\mathbf{W}_k$, the reduction factor is around 300 for $|\mathbf{P}|$, around 200 for the product of the diagonal elements of the Cholesky, around 900 for the product of the calculated eigenvalues of $\mathbf{P}$, and around 200, 400, 1200, respectively, using the 2×2 partitioned matrices. Similar reductions occur for $p = 3$ and $n_1 = n_2 = 16$. Since these n_1 and n_2 are small, there is only a small extra reduction factor of around 1.2 when the known eigenvalues of the $\mathbf{W}_k$ are used. The number of operations is of the same order for each of the DFT, DCT and DST, but the RBA and NBA do not appear to have been widely used for fitting higher-order RS-CARs.

5. Applications to real data

In this section a brief discussion of two applications is given. Both data sets are large and exhibit complicated behavior, requiring very detailed analyses. The aim here is to obtain some preliminary insight by fitting multivariate CARs.

5.1. Per capita income and unemployment rates in Italy

The first data set has regional data from 2011 for 8078 municipalities in Italy, and uses per capita income (PCI) in euros, from the Italian Ministry of Finances (`http://www1.finanze.gov.it/finanze2/pagina_dichiarazioni/dichiarazioni.php`), and unemployment rate (UR) from the 2011 census, from the Italian Institute of Statistics (ISTAT) (`http://ottomilacensus.istat.it/download-dati/`). The PCI ranges from 11,998 to 53,589, with mean 20,416 and sd 3044, while UR ranges from 0 to 42.2, with mean 10.13 and sd 6.3. These are both macroeconomic factors that can be used to gauge the state of an economy. An increase in demand leads to an increase in gross domestic product (GDP), which must be accompanied by a corresponding increase in productivity, and a consequent decrease in unemployment levels. The inverse relationship between UR and GDP is commonly referred to as Okun's law, and the bivariate CAR allows this to be studied conditionally given the values at neighboring locations. At municipality area level in Italy, only the PCI is available, which is used here as a proxy for GDP. The data are represented in Figure 1. The correlation between PCI and UR is −0.441. Between neighboring regions, both variables are correlated — Moran's coefficient being 0.763 (PCI) and 0.638 (UR), with the negative −0.404 between PCI and UR.

Both variables are positively skewed, and were log-transformed (after adding 0.5 to UR). Both variables also have a strong trend from North-West to South-East, and a planar trend (using the centroids of each region) was then removed. The detrended ln(PCI) ranges from 9.39 to 10.89, with mean 9.91 and sd 0.145, while the detrended ln(UR + 0.5) ranges from −0.69 to 3.75, with mean 2.20 and sd 0.591. The four relevant correlations only slightly differ from those for the original data.

For the CAR modeling of the detrended log variables, constant means are assumed, and the usual neighborhood adjacency matrix $\mathbf{W}$, with the associated $\mathbf{D}$ (see Section 2.4), are used. The number of neighbors ranges from 1 to 29, with median 6.

The full 6-parameter model $\mathbf{P} = -(\mathbf{H}_\mathrm{D}\mathbf{A}_0 \otimes \mathbf{D} + \mathbf{H}_\mathrm{D}\mathbf{A}_1 \otimes \mathbf{W})$ described in Section 4.1.3 was fitted. The resulting estimates (estimated standard errors) are: $\hat{\tau}_{11} = 0.0425$ (0.0007), $\hat{\tau}_{22} = 0.502$ (0.0082), $\hat{\alpha}_{0,12} = -0.0037$ (0.0033), $\hat{\alpha}_{1,11} = 0.987$ (0.0031), $\hat{\alpha}_{1,12} = 0.0052$ (0.0036), $\hat{\alpha}_{1,22} = 0.958$ (0.0064).

Using Corollary 2 to obtain $\mathbf{y}^\top\mathbf{P}\mathbf{y}$ as $\mathbf{u}^\top\mathbf{R}_\mathrm{R}\mathbf{u}$ with $\mathbf{u} = (\mathbf{I}_p \otimes \mathbf{D}^{1/2})\mathbf{y}$ reduced the computer time by a factor of around 250, and fitting using $\mathbf{H}_\mathrm{D}\mathbf{A}_k = \mathbf{\Phi}_{\mathrm{S},k}$ for $k \in \{0, 1\}$ (see Section 4.1.3) performed better than directly estimating $\mathbf{H}_\mathrm{D}$ and the $\mathbf{A}_k$s.

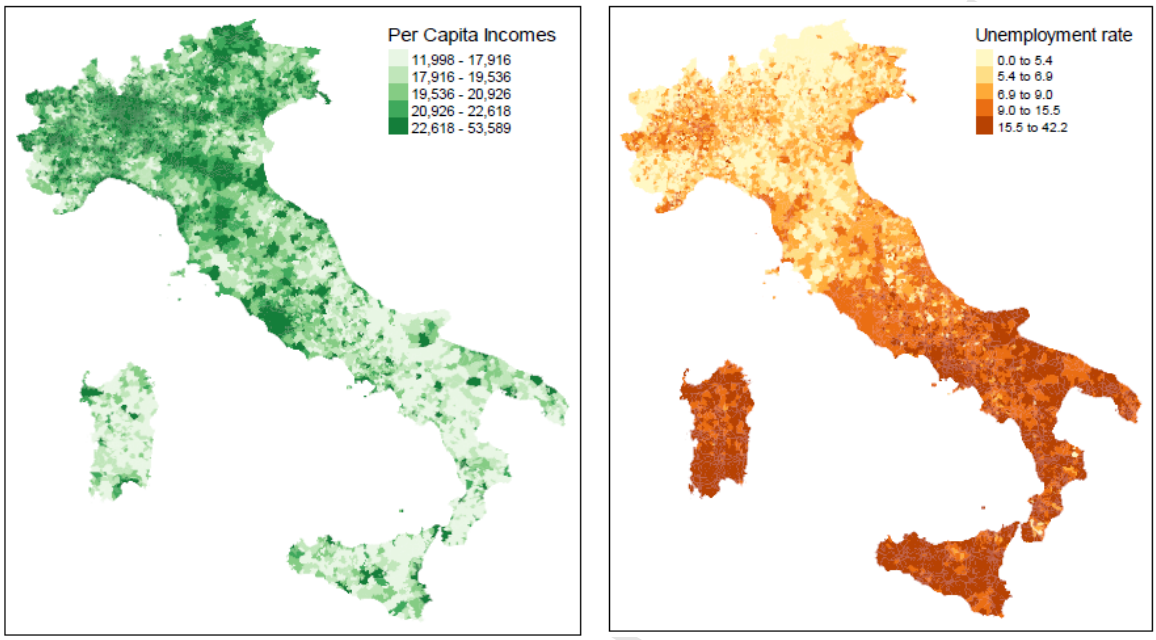

Figure 1: Maps showing the per capita income (PCI) and unemployment rate (UR) in Italy in 2011.

Although the original sample statistics show reasonable negative correlations between the variables within regions and between neighboring regions, both between-variable estimates $\hat{\alpha}_{0,12}$ and $\hat{\alpha}_{1,12}$ are relatively small, and the fit with $\alpha_{0,12} = \alpha_{1,12} = 0$ is not significantly different (and the remaining estimates are essentially unchanged). This fit suggests that Okun's law may not hold when conditioning on neighboring regions.

5.2. NO_2 and O_3 concentrations in Northern Italy

The second data set uses modeled average daily measurements of nitrogen dioxide (NO_2) and ozone (O_3) for July 7, 2016 on an 82×128 regular grid (n = 10,496), with a spatial resolution of 5×5 km^2. The area covers Northern Italy, and includes some parts of France, Switzerland and Slovenia, plus some of the adjacent Ligurian and Adriatic seas, ranging approximately from Grenoble in the West to Trieste in the East, and from Arezzo in the South to Lausanne in the North. One of the main problems caused by air pollution in urban areas is photochemical oxidants. Among these, NO_2 and O_3 are particularly important because they can affect human health. The data came from the NINFA2015 (Northern Italy Network to Forecast Photochemical and Aerosol pollution) model [23].

The NO_2 values (in μg/m^3) range from 0.144 to 78.063 with mean 5.37, and the O_3 values (in μg/m^3) range from 86.04 to 213.07, with mean 143.75. Their within-location correlation is 0.400. Both variables are positively skewed, and were log-transformed. The transformed variables are shown in Figure 2, which suggests strong complicated trends in the variables.

To remove the large scale spatial trends, the log data were corrected by first-order spatial differencing with equal weights. The detrended ln(NO_2) ranges from -0.950 to 1.836, with sd 0.211, while the detrended ln(O_3) ranges from -0.287 to 2.634, with sd 0.247. Their neighbor correlations are 0.200 (N-S) and 0.214 (E-W), and 0.703 (N-S) and 0.794 (E-W), respectively. Their within-site correlation is now 0.349.

An 18-parameter RS-CAR(3) with the RBA was fitted. Although some of the estimated parameters appear small, no model simplification was possible, e.g., the GLRT statistic for reducing to the RS-CAR(2) is 224 on 6 degrees of freedom. Using Corollary 1 reduced the computer time for fitting by a factor of around 1700 compared with the best alternative. Fitting higher-order models with many more parameters takes significantly longer, and has convergence problems.

The RS-CAR(3) with the TBA has a much higher deviance than the RBA fit. Similarly, the second order RS-CAR fits with the FBA and TBA have much higher deviances than the RBA fit.

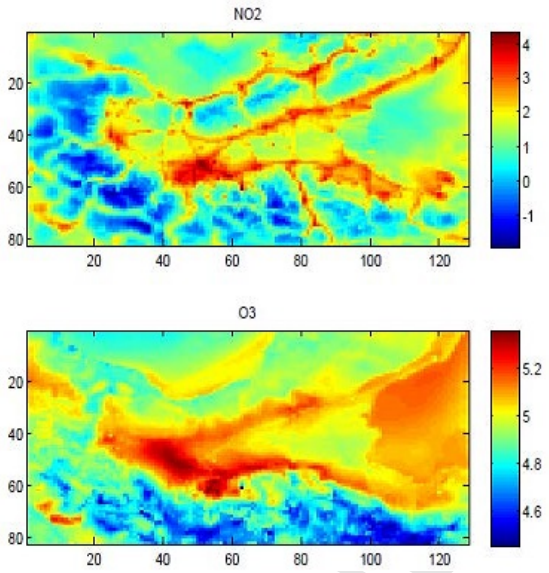

Figure 2: Maps of the log-transformed NO_2 and O_3 on 7th July 2016.

Table 1: Estimated conditional correlations between the detrended ln(NO_2) and ln(O_3) for pairs of pixels at lags (0,0), (0,1), (1,0), (1,1), (0,2), and (2,0), where lag (g_1, g_2) is a vertical (N-S) lag of g_1 and a horizontal (E-W) lag of g_2.

		(0,0)		**(0,1)**		**(0,2)**	
		NO_2	**O_3**	**NO_2**	**O_3**	**NO_2**	**O_3**
(0,0)	**NO_2**	1	0.197	0.144	0.139	−0.130	−0.036
	O_3	0.197	1	0.139	0.546	−0.036	−0.087
(1,0)	**NO_2**	−0.003	−0.020	0.006	0.014		
	O_3	−0.020	−0.060	0.014	0.048		
(2,0)	**NO_2**	−0.006	0.004				
	O_3	0.004	−0.003				

The estimated conditional correlations from the third order RS-CAR with the RBA are shown in Table 1. There is a positive conditional correlation between the detrended ln(NO_2) and ln(O_3) within a location, and at lag (0,1) — the horizontal (E-W) neighbor. Also, there is some univariate conditional correlation at lag (0,1), quite strong for O_3, plus smaller negative ones at lag (0,2). Other conditional correlations are small.

Note that in general, chemical reactions between air pollutants are very complex, and the relationships between them may differ from one location to another. The region considered here covers a very large area, containing many factors that may influence relationships, e.g., sea and mountains.

6. Discussion

Connections between different definitions of multivariate CARs and different data orderings have been given in general. A simplified additive model encompasses many of the models commonly proposed. If this model has symmetric components, then a natural check can be made on the permissible region of the parameters in simple regional models, including other forms for **W**, and in some common cases on regular rectangular lattices. Gaussian maximum likelihood is feasible on much larger data sets than was possible previously.

Usage for the rectangular lattice may appear to be limited as Lemma 1 only fully applies to general CARs with the TBA. However, any RS-CAR can be fitted under the RBA or NBA, and some other boundary assumptions can be used for RS-CARs up to second-order dependence. Low-order CAR models are frequently used as priors in hierarchical modeling, and can often capture the main features of a real data set. With large lattices and higher-order models, it is common to use the TBA, either in itself, or as a first approximation to fitting a planar model. If n_1 and n_2 are large, the estimates obtained using the TBA may be very close to those using a different, more reasonable, boundary assumption, which could then be used with those estimates for prediction, etc. Inadequacies detected in a fitted model can then suggest other, harder to fit, models. Similarly, for smaller lattices, fitting an RS-CAR with the RBA or NBA may provide useful insight into more appropriate models.

These results can be extended to other regular lattices, including higher-dimensional rectangular lattices. The results have been presented in terms of CAR (GMRF) models, but Lemma 1 can still be used if there is additive independent noise (or measurement error) — an errors-in-variables CAR [2], and for the more general Rational Spectral Density models [9] provided the Direct Covariance component uses the same boundary and symmetry assumptions as the CAR component.

Simple extensions would allow the matrix $\mathbf{M}$ in Section 2.3 to be non-diagonal, and possibly also dependent on parameters.

If the data have measurement error, so that the (noisy) observed data are $\mathbf{z} = \mathbf{y} + \boldsymbol{\varepsilon}$, and $\boldsymbol{\varepsilon}$ is independent of the original data $\mathbf{y}$ (the signal) with $\text{var}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}_{np}$, $\text{var}(\mathbf{y}) = \mathbf{P}^{-1}$, and interest is in $\mathbf{y}$, then $\mathbf{y}$ is predicted by $\hat{\mathbf{y}} = \text{E}(\mathbf{y}|\mathbf{z}) = (\mathbf{I}_{np} + \sigma^2\mathbf{P})^{-1}\mathbf{z}$, with mean-square error $\text{E}\{(\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})^\top\} = \sigma^2(\mathbf{I}_{np} + \sigma^2\mathbf{P})^{-1}$. Fitting can be carried out directly on the noisy data $\mathbf{z}$ as an errors-in-variables CAR to then obtain $\hat{\mathbf{y}}$. Alternatively, an iteration could use given estimates of σ^2 and the CAR parameters to estimate $\hat{\mathbf{y}}$, and then refit the CAR to the estimated $\hat{\mathbf{y}}$ to get updated parameter estimates.

If there are missing values, or outlying values that are removed, the methods in [16] can be used to obtain the maximum likelihood estimates, and if required, estimates of these values. Further work is in progress on Gaussian maximum likelihood fitting of general multivariate CARs on a regular rectangular lattice, and investigating the use of the Kalman filter on a unilateral representation of the CAR as another method for fitting data with measurement error.

Appendix

A.1. Examples for Section 2

Example A Suppose $p = 2$, with n sites on a line, and that conditional variances and covariances do not depend on the site, so that $\mathbf{T}_j = \mathbf{T} = (\tau_{ii'})$ and $\mathbf{H}_j = \mathbf{H} = (\eta_{ii'})$ for all j. Also, for j an interior site, $1 < j < n$ (boundary assumptions are discussed in Section 4.2.1), suppose dependences are only on adjacent (first-order) neighbors. Then, the mvCAR and uvCAR forms can both be written, viz.

$$\begin{aligned}
&\text{E}(y_{1,j}|y_{-*}) = \alpha_{0,12}y_{2,j} + \alpha_{1,11}y_{1,j-1} + \alpha_{2,11}y_{1,j+1} + \alpha_{1,12}y_{2,j-1} + \alpha_{2,12}y_{2,j+1}, \quad \text{var}(y_{1,j}|y_{-*}) = \tau_{11}, \quad \text{and}\\
&\text{E}(y_{2,j}|y_{-*}) = \alpha_{0,21}y_{1,j} + \alpha_{1,22}y_{2,j-1} + \alpha_{2,22}y_{2,j+1} + \alpha_{1,21}y_{1,j-1} + \alpha_{2,21}y_{1,j+1}, \quad \text{var}(y_{2,j}|y_{-*}) = \tau_{22},
\end{aligned}$$

with $\text{cov}(y_{1,j}, y_{2,j}|y_{-*}) = \tau_{12}$. The mvCAR form has * equal to j, and $\alpha_{0,12} = \alpha_{0,21} = 0$. The uvCAR form has * equal to (i, j) for $(y_{i,j}|y_{-*})$, with $\tau_{12} = 0$.

For simplicity, assume now that the cross-dependences are equal, so that $\alpha_{1,ii'} = \alpha_{2,ii'}$ for all i, i'. Symmetry of $\mathbf{P}_{(\text{s})}$ requires that $\eta_{12}\alpha_{1,11} + \eta_{22}\alpha_{1,21} = \eta_{11}\alpha_{1,12} + \eta_{12}\alpha_{1,22}$ for the mvCAR form; and that $\eta_{11}\alpha_{0,12} = \eta_{22}\alpha_{0,21}$ and $\eta_{11}\alpha_{1,12} = \eta_{22}\alpha_{1,21}$ for the uvCAR form. In both cases there are 6 'free' parameters (subject to pd conditions).

Given the mvCAR form, the uvCAR form is

$$\begin{aligned}
&\text{E}(y_{1,j}|y_{-(i,j)}) = \{-\eta_{12}y_{2,j} + (\eta_{11}\alpha_{1,11} + \eta_{12}\alpha_{1,21})(y_{1,j-1} + y_{1,j+1}) + (\eta_{11}\alpha_{1,12} + \eta_{12}\alpha_{1,22})(y_{2,j-1} + y_{2,j+1})\}/\eta_{11},\\
&\text{var}(y_{1,j}|y_{-(i,j)}) = 1/\eta_{11};
\end{aligned}$$

and similarly for $y_{2,j}$. Note the within-site cross-dependence term $-(\eta_{12}/\eta_{11})y_{2,j}$. Given the uvCAR form, the mvCAR form is $\text{E}(\mathbf{y}_j|y_{-j}) = -\mathbf{A}_0^{-1}\mathbf{A}_1(\mathbf{y}_{j-1} + \mathbf{y}_{j+1})$ and $\text{var}(\mathbf{y}_j|y_{-j}) = -\mathbf{A}_0^{-1} \times \text{diag}(\tau_{11}, \tau_{22})$, where

$$\mathbf{A}_0 = \begin{bmatrix} -1 & \alpha_{0,12} \\ \alpha_{0,21} & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{A}_1 = \begin{bmatrix} \alpha_{1,11} & \alpha_{1,12} \\ \alpha_{1,21} & \alpha_{1,22} \end{bmatrix}$$

with $\eta_{11}\alpha_{0,12} = \eta_{22}\alpha_{0,21}$ and $\eta_{11}\alpha_{1,12} = \eta_{22}\alpha_{1,21}$ (and $\alpha_{0,12}\alpha_{0,21} < 1$). Note that the off-diagonal elements of $\text{var}(\mathbf{y}_j|y_{-j})$ involve $\alpha_{0,12}$, with those of $\{\text{var}(\mathbf{y}_j|y_{-j})\}^{-1}$ being $-\eta_{11}\alpha_{0,12}$. □

Example B. Assuming $\mathbf{y}_0 = \mathbf{y}_{n+1} = \mathbf{0}$ (see the FBA in Section 4.2.1), the original model in Example A results if $\mathbf{M}_{(s)} = \mathbf{I}_{np}$, and in Eq. (2), $k_{\max} = 2$, $\mathbf{W}_0 = \mathbf{I}_n$, with $\mathbf{W}_2 = \mathbf{U}_{n,1}$, the adjacent (lag 1) right neighbor incidence matrix having 1 in the $(j, j+1)$ position for $j \in \{1, \ldots, n-1\}$, and $\mathbf{W}_1 = \mathbf{W}_2^\top$, the incidence matrix for adjacent left neighbors. The mvCAR form has $\mathbf{A}_0 = -\mathbf{I}_2$ (i.e., $\alpha_{0,12} = \alpha_{0,21} = 0$), and a general $\mathbf{H}$. The uvCAR form has $\mathbf{T}$ and $\mathbf{H}$ diagonal, $\mathbf{H} = \mathbf{H}_\text{D}$ (i.e., $\tau_{12} = \eta_{12} = 0$). Symmetry of $\mathbf{P}_{(s)}$ requires that $\mathbf{A}_2 = \mathbf{H}^{-1}\mathbf{A}_1^\top\mathbf{H}$. If, as in most of Example A, $\alpha_{1,ii'} = \alpha_{2,ii'}$ for all i, i', then $k_{\max} = 1$, with $\mathbf{W}_1 = \mathbf{U}_{n,1} + \mathbf{U}_{n,1}^\top = \mathbf{N}_{n,1}$, the neighbor adjacency matrix, and symmetry of $\mathbf{P}_{(s)}$ requires that $\mathbf{\Phi}_k = \mathbf{H}\mathbf{A}_k$ is symmetric for $k \in \{0, 1\}$. □

A.2. Previously proposed regional models

The models of [8, 21] and the simplified one of [22] have $k_{\max} = 2$ in Eqs. (2)–(3), and use $\mathbf{W}_1 = \mathbf{W}_\text{U}$ and $\mathbf{W}_2 = \mathbf{W}_\text{U}^\top$ with $\mathbf{A}_1$ asymmetric, and $\mathbf{A}_2 = \mathbf{H}^{-1}\mathbf{A}_1^\top\mathbf{H}$ for Eq. (2) or $\mathbf{A}_2 = \mathbf{A}_1^\top$ for Eq. (3).

That of [8] has the form of Eq. (2) with $\mathbf{A}_1 = \mathbf{H}^{-1}(\mathbf{H} \circ \mathbf{A})$, for some $\mathbf{A} = (\alpha_{ii'})$, so that

$$\begin{aligned}\mathbf{P}_{(s)} &= \mathbf{D} \otimes \mathbf{H} - \mathbf{W}_\text{U} \otimes (\mathbf{H} \circ \mathbf{A}) - \mathbf{W}_\text{U}^\top \otimes (\mathbf{H} \circ \mathbf{A}^\top) \\ &= (\mathbf{I}_n \otimes \mathbf{H})\{\mathbf{D} \otimes \mathbf{I}_p - \mathbf{W}_\text{U} \otimes \mathbf{H}^{-1}(\mathbf{H} \circ \mathbf{A}) - \mathbf{W}_\text{U}^\top \otimes \mathbf{H}^{-1}(\mathbf{H} \circ \mathbf{A}^\top)\}.\end{aligned}$$

This generalizes the model of [11] for which $\mathbf{A}$ is symmetric, so that the model of [11] has $k_{\max} = 1$, with $\mathbf{P}_{(s)} = \mathbf{D} \otimes \mathbf{H} - \mathbf{W} \otimes (\mathbf{H} \circ \mathbf{A})$; which itself generalizes the separable (or factorized) model with $\mathbf{P}_{(s)} = (\mathbf{D} - \alpha\mathbf{W}) \otimes \mathbf{H}$, i.e., $\mathbf{A} = \alpha\mathbf{J}_p$.

The models of [21] use $\mathbf{M}_{(s)}$, with $\mathbf{W}_0 = \mathbf{I}_n$, and have the form of Eq. (3). Their (mvCAR form) CAMCAR model has

$$\begin{aligned}\mathbf{R}_{\text{S},(s)} &= (\mathbf{I}_n \otimes \mathbf{H}^{1/2})(\mathbf{I}_{np} - \mathbf{W}_\text{U} \otimes \mathbf{A}_1 - \mathbf{W}_\text{U}^\top \otimes \mathbf{A}_1^\top)\{\mathbf{I}_n \otimes (\mathbf{H}^{1/2})^\top\} = \\ &= \mathbf{I}_n \otimes \mathbf{H} - \mathbf{W}_\text{U} \otimes \{\mathbf{H}^{1/2}\mathbf{A}_1(\mathbf{H}^{1/2})^\top\} - \mathbf{W}_\text{U}^\top \otimes \{\mathbf{H}^{1/2}\mathbf{A}_1^\top(\mathbf{H}^{1/2})^\top\}.\end{aligned}$$

Their (uvCAR form) extended CAMCAR model is very similar, but uses $\mathbf{H}_\text{D}$ and a symmetric $\mathbf{A}_0$ instead of $\mathbf{H}$. It has

$$\begin{aligned}-\mathbf{R}_{\text{S},(s)} &= (\mathbf{I}_n \otimes \mathbf{H}_\text{D}^{1/2})(\mathbf{I}_n \otimes \mathbf{A}_0 + \mathbf{W}_\text{U} \otimes \mathbf{A}_1 + \mathbf{W}_\text{U}^\top \otimes \mathbf{A}_1^\top)t(\mathbf{I}_n \otimes \mathbf{H}_\text{D}^{1/2}) \\ &= \mathbf{I}_n \otimes (\mathbf{H}_\text{D}^{1/2}\mathbf{A}_0\mathbf{H}_\text{D}^{1/2}) + \mathbf{W}_\text{U} \otimes (\mathbf{H}_\text{D}^{1/2}\mathbf{A}_1\mathbf{H}_\text{D}^{1/2}) + \mathbf{W}_\text{U}^\top \otimes (\mathbf{H}_\text{D}^{1/2}\mathbf{A}_1^\top\mathbf{H}_\text{D}^{1/2}).\end{aligned}$$

The simplified model of [22] is the extended CAMCAR model with $\mathbf{M}_{(s)} = \mathbf{I}_{np}$, i.e., $\mathbf{P}_{(s)} = \mathbf{R}_{\text{S},(s)}$.

The model of [10] has $\mathbf{M}_{(s)} = \mathbf{M}_\text{s} \otimes \mathbf{I}_p$, and $\mathbf{R}_{(s)} = -(\mathbf{I}_n \otimes \mathbf{H}_\text{D})(\mathbf{I}_n \otimes \mathbf{A}_0 + \mathbf{W} \otimes \mathbf{A}_1)$, with $\mathbf{H}_\text{D}\mathbf{A}_k$ symmetric for $k \in \{0, 1\}$. This is essentially the same as the extended CAMCAR model with $\mathbf{A}_1$ symmetric and $\mathbf{M}_{(s)} = \mathbf{M}_\text{s} \otimes \mathbf{I}_p$, or the simplified model of [22] with $\mathbf{A}_1$ symmetric and $\mathbf{M}_\text{s} = \mathbf{I}_n$.

Two other models can be written using Eqs. (2)–(3), with $\mathbf{M}_{(s)} = \mathbf{I}_{np}$. The Twofold CAR model of [13] uses $\mathbf{I}_n$ as well as $\mathbf{D}$ and $\mathbf{W}$, and has

$$\begin{aligned}\mathbf{P}_{(s)} &= 2\mathbf{D} \otimes \mathbf{H}_\text{D} - \mathbf{I}_n \otimes (\mathbf{H}_\text{D}^{1/2}\mathbf{A}_0\mathbf{H}_\text{D}^{1/2}) - \mathbf{W} \otimes (\mathbf{H}_\text{D}^{1/2}\mathbf{A}_1\mathbf{H}_\text{D}^{1/2}) = \\ &= (\mathbf{I}_n \otimes \mathbf{H}_\text{D}^{1/2})(2\mathbf{D} \otimes \mathbf{I}_p - \mathbf{I}_n \otimes \mathbf{A}_0 - \mathbf{W} \otimes \mathbf{A}_1)(\mathbf{I}_n \otimes \mathbf{H}_\text{D}^{1/2}).\end{aligned}$$

This $\mathbf{P}_{(s)}$ adds $2\mathbf{D} \otimes \mathbf{H}_\text{D}$ to the symmetric [22] model, giving extra terms on the diagonal.

The model of [12] for $p = 2$ specifies $\mathbf{y}_1$ and $\mathbf{y}_2|\mathbf{y}_1$ as univariate CARs, and so the likelihood, etc., can be obtained from two univariate CARs. However, $\mathbf{y}_1$ and $\mathbf{y}_2$ are treated asymmetrically, and the joint distribution is not simple. For example, in the simplest case $\mathbf{y}_1$ and $\mathbf{y}_2|\mathbf{y}_1$ are first-order CARs (i.e., just involving adjacent neighbors), with $\mathbf{y}_2|\mathbf{y}_1$ having first-order dependence in $\mathbf{y}_1$. Then from the inverse of their Eq. (13), in the bivariate model $\mathbf{y}_1$ has first-order dependence, the cross-dependence is second-order involving neighbors of neighbors, and $\mathbf{y}_2$ has third-order dependence. Hence Eq. (2) has $k_{\max} = 3$, with $\mathbf{W}_2 = \mathbf{W}^2$ having non-zero elements on the diagonal, and $\mathbf{W}_3 = \mathbf{W}^3$ having non-zero elements overlapping with those of $\mathbf{W}$ (i.e., $\mathbf{W} \circ \mathbf{W}^3 \neq \mathbf{0}$).

A.3. The method of Martinez-Beneito [18]

For $\mathbf{E}_1$, $\mathbf{E}_2$ as in Section 4.1, let $\mathbf{P}$ be $(\mathbf{H}^{1/2} \otimes \mathbf{E}_1)(\mathbf{I}_{np} - \mathbf{A}^* \otimes \mathbf{E}_2)\{(\mathbf{H}^{1/2})^\top \otimes \mathbf{E}_1\}$, where $\mathbf{A}^*$ is diagonal, and the diagonal elements are allowable parameters for one-dimensional CARs, i.e., are between $1/\omega_{\min}$ and $1/\omega_{\max}$ for all i. Then using $\mathbf{H} \circ \mathbf{A} = \mathbf{H}^{1/2}\mathbf{A}^*(\mathbf{H}^{1/2})^\top$, gives $\mathbf{A} = \{\mathbf{H}^{1/2}\mathbf{A}^*(\mathbf{H}^{1/2})^\top\}\backslash \circ \mathbf{H}$, where $\backslash\circ$ denotes element-wise division. Any $\mathbf{H}^{1/2}$ can be used, and the resulting $\mathbf{P}$ will be pd.

However, since $\mathbf{A}^* = \mathbf{H}_{\mathrm{A,S}}$ is diagonal, it follows that $\mathbf{A}^*$ contains the eigenvalues of $\mathbf{H}_{\mathrm{A,S}}$, and hence those of $\mathbf{H}_{\mathrm{A}}$, so that $\mathbf{A}^* = \mathrm{diag}(\mu_1, \ldots, \mu_p)$. Thus, the results for $|\mathbf{P}|$ and checking $\mathbf{P}$ is pd are exactly those given initially in Section 4.1.2. If required, the appropriate $\mathbf{H}^{1/2}$ for transforming back from $\mathbf{H} \circ \mathbf{A}$ to $\mathbf{A}^*$ can be found by obtaining the standardized matrix $\mathbf{X}$ of eigenvectors of $\mathbf{H}_{\mathrm{A,S}}$ for any $\mathbf{H}^{1/2}$, and then $\mathbf{H}^{1/2}\mathbf{X}$ satisfies $(\mathbf{H}^{1/2}\mathbf{X})^{-1}(\mathbf{H} \circ \mathbf{A})\{(\mathbf{H}^{1/2}\mathbf{X})^\top\}^{-1} = \mathbf{A}^*$. For a given ordering of the μ_is in $\mathbf{A}^*$, $\mathbf{H}^{1/2}$ is usually unique, assuming its (1,1)-element is positive.

In particular, if $p = 2$, then μ_1 and μ_2 are the roots of $x^2 - \mathrm{tr}(\mathbf{H}_{\mathrm{A}})x + |\mathbf{H}_{\mathrm{A}}| = 0$, and $\prod_{i,j}(1 - \mu_i\omega_j) = \prod_j [1 - \mathrm{tr}(\mathbf{H}_{\mathrm{A}})\omega_j + \{|\mathbf{H}_{\mathrm{A}}|\}\omega_j^2]$. Note that if $|\mathbf{H} \circ \mathbf{A}| > 0$, as commonly occurs, then the μ_i have the same sign.

If it wished to obtain it directly, the squares of the first row elements of the $\mathbf{H}^{1/2}$ for the method of [18] are $\eta_{11}[1 \pm \{2\alpha_{11} - \mathrm{tr}(\mathbf{H}_{\mathrm{A}})\}/\sqrt{\{\mathrm{tr}(\mathbf{H}_{\mathrm{A}})^2 - 4|\mathbf{H}_{\mathrm{A}}|\}}]/2$; given one element of $\mathbf{H}^{1/2}$, the remaining three are determined by $\mathbf{H}^{1/2}(\mathbf{H}^{1/2})^\top = \mathbf{H}$.

A.4. Examples for Section 4.1

Example C. If $p = 2$,

$$\mathbf{H} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix} 0.2 & 0.3 \\ 0.3 & 0.1 \end{bmatrix},$$

then

$$\mathbf{H}_{\mathrm{A}} = \begin{bmatrix} 0.18 & 0.12 \\ 0.04 & 0.06 \end{bmatrix}.$$

The eigenvalues of $\mathbf{H}_{\mathrm{A}}$ satisfy $\mu^2 - 0.24\mu + 0.006 = 0$, so that $\mu_2 = 0.12 + (0.0084)^{1/2} \approx 0.2117$, with $\mu_1 \approx 0.0283$, i.e., $\min(\mu_i) > 0$. Then

$$\prod_{i,j}(1 - \mu_i\omega_j) = \prod_j (1 - 0.24\omega_j + 0.006\omega_j^2).$$

If $\omega_{\max} = 1$, $\mu_2 < 1$ shows that the parameters are acceptable. Using an upper triangular $\mathbf{T}_{\mathrm{s}}$, the method of [11] has

$$\mathbf{H}_{\mathrm{A}} = \begin{bmatrix} 0.2 & \sqrt{0.002} \\ \sqrt{0.002} & 0.04 \end{bmatrix},$$

with the same eigenvalues. The method of [18] has $\mathbf{A}^* = \mathrm{diag}(\mu_1, \mu_2)$ with

$$\mathbf{H}^{1/2} \approx \begin{bmatrix} 0.357 & 1.368 \\ -1.352 & 1.082 \end{bmatrix},$$

where the elements of the first row are the square roots of $1 \pm 4/\sqrt{21} \approx 0.127$ and 1.873. If however $\alpha_{11} = 0.9$ and $\alpha_{22} = 0.8$, the eigenvalues of $\mathbf{H}_{\mathrm{A}}$ satisfy $\mu^2 - 1.92\mu + 0.846 = 0$, so that $\mu_2 = 0.96 + (0.0756)^{1/2} \approx 1.235$. Thus, if $\omega_{\max} = 1$, $\mu_2\omega_{\max} > 1$, and the parameters are not acceptable. □

Example D. With $\mathbf{H}$ and $\mathbf{A}$ as in Example C, the pd conditions on α_{12} when $\omega_{\max} = 1$ and $\omega_{\min} = -1$ become $3.32 + 2\alpha_{12} - \alpha_{12}^2 > 0$, and $6.92 - 2\alpha_{12} - \alpha_{12}^2 > 0$, leading to (respectively) $\alpha_{12} > -1.078$ and $\alpha_{12} < 1.814$. If $\omega_{\min} = -1/2$, then the second condition becomes $23.72 - 4\alpha_{12} - \alpha_{12}^2 > 0$, and both lead to (from the first) $\alpha_{12} < 3.078$. If $\alpha_{11} = 0.9$ and $\alpha_{22} = 0.8$, the condition on α_{12} when $\omega_{\max} = 1$ becomes $-0.88 + 2\alpha_{12} - \alpha_{12}^2 > 0$, leading to $0.654 < \alpha_{12} < 1.346$. This condition is the only one even if $\omega_{\min} = -1$. □

Example E. If $\mathbf{H}$ and $\mathbf{A}$ are as in Example C, then

$$\mathbf{C}_j = \begin{bmatrix} 2(1 - 0.2\omega_j) & 1 - 0.3\omega_j \\ 1 - 0.3\omega_j & 3(1 - 0.1\omega_j) \end{bmatrix},$$

and the eigenvalues of the $\mathbf{C}_j$ are the roots of $x^2-(5-0.7\omega_j)x+(5-1.2\omega_j+0.03\omega_j^2)=0$, i.e., $\{(5-0.7\omega_j)\pm(5-2.2\omega_j+0.37\omega_j^2)^{1/2}\}/2$. These roots are positive provided $\omega_j<5$ (from $c_{j,11}$ and $c_{j,22}$), and $\omega_j<r_1=10\{2-(7/3)^{1/2}\}\approx 4.72$ (from $|\mathbf{C}_j|$). Thus the only pd condition is $\omega_j<r_1$, i.e., $\omega_{\max}<r_1$. □

Example F. Eq. (10) of [10] has $\mathbf{P}_{ii}=\eta_{ii}(\mathbf{I}_n-\alpha_{1,ii}\mathbf{W})$ for $i\in\{1,2\}$, and $\mathbf{P}_{12}=-\eta_{11}(\alpha_{0,12}\mathbf{I}_n+\alpha_{1,12}\mathbf{W})$. Then, apart from the $\eta_{ii}>0$, and the standard conditions on the $\alpha_{1,ii}$ for $\mathbf{P}$ to be pd, it is also necessary that $\eta_{22}(1-\alpha_{1,11}\omega_j)(1-\alpha_{1,22}\omega_j)>\eta_{11}(\alpha_{0,12}+\alpha_{1,12}\omega_j)^2$ for all ω_j. The same condition results from $\mathbf{C}_j=-(\mathbf{A}_0+\omega_j\mathbf{A}_1)$ after using the symmetry of $\mathbf{HA}$. For the model of [22] with $\mathbf{A}_1$ symmetric, using $\mathbf{C}_j=-(\mathbf{A}_0+\omega_j\mathbf{A}_1)$ the extra necessary condition for $\mathbf{P}$ to be pd is $(1-\alpha_{1,11}\omega_j)(1-\alpha_{1,22}\omega_j)>(\alpha_{0,12}+\alpha_{1,12}\omega_j)^2$ for all ω_j. It does not involve the parameters of $\mathbf{H}$. □

A.5. Boundary conditions for CARs on a regular rectangular lattice

The torus TBA is often used. It gives a stationary process, but the resulting correlations depend on n_1 and n_2. It implies, for example, that site 1 is the right neighbor of site n_2, and immediately below site $n_2(n_1-1)+1$. Thus the correlations are completely unrealistic for sites that are very distant, e.g., at lags (g_1,g_2) with n_m-g_m small for $m=1$ or 2.

The fixed FBA seems reasonable, and has frequently been used. But at border sites it results in low variances under positive correlations.

The reflection RBA reflects horizontal lags outside the lattice about 1/2 or $n_2+1/2$, and similarly for vertical lags. Thus, on a line site $j-g$, $g>j-1$, is replaced by site $g-j+1$, and site $j+g$, $g>n-j$, is replaced by site $2n+1-(j+g)$. This implies that border sites (1 or n) are first neighbors of themselves, sites 2 and $n-1$ are third neighbors of themselves, etc. Even for a CAR(1), the diagonal elements of $\mathbf{P}$ alter the conditional variance, and scale the coefficients for the conditional mean.

The symmetric Neumann SBA reflects horizontal lags outside the lattice about 1 or n_2+2, so that site 2 on a line is a left first-neighbor of site 1. Symmetry of $\mathbf{P}$ then strangely requires that on a line site 1 is twice a left first-neighbor of site 2. It leads to a large variation in the variances and covariances, and would rarely be useful.

Example G. Consider a univariate RS-CAR(2). The infinite-lattice stationary model has

$$\mathrm{E}(y_{j_1,j_2}|y_{-(j_1,j_2)})=\phi_{1,0}(y_{j_1-1,j_2}+y_{j_1+1,j_2})+\phi_{0,1}(y_{j_1,j_2-1}+y_{j_1,j_2+1})+\phi_{1,1}(y_{j_1-1,j_2-1}+y_{j_1-1,j_2+1}+y_{j_1+1,j_2-1}+y_{j_1+1,j_2+1}),$$

and $\mathrm{var}(y_{j_1,j_2}|y_{-(j_1,j_2)})=\tau$. For a finite lattice, using Eq. (1), this specifies all elements of $\mathbf{P}$ except for those corresponding to both sites on the border, i.e., each site has $j_m=1$ or n_m, $m=1$ or 2. The TBA has $\mathrm{E}(y_{1,1}|y_{-(1,1)})=\phi_{1,0}(y_{2,1}+y_{n_1,1})+\phi_{0,1}(y_{1,2}+y_{1,n_2})+\phi_{1,1}(y_{2,2}+y_{2,n_2}+y_{n_1,2}+y_{n_1,n_2})$, etc. The FBA has $\mathrm{E}(y_{1,1}|y_{-(1,1)})=\phi_{1,0}y_{2,1}+\phi_{0,1}y_{1,2}+\phi_{1,1}y_{2,2}$, $\mathrm{E}(y_{1,2}|y_{-(1,2)})=\phi_{1,0}y_{2,2}+\phi_{0,1}(y_{1,1}+y_{1,3})+\phi_{1,1}(y_{2,1}+y_{2,3})$, etc. The RBA alters the diagonal elements of $\mathbf{P}$ for $j_1=1$ or n_1, or $j_2=1$ or n_2, and the neighbor dependence, so that $\mathrm{var}(y_{1,1}|y_{-(1,1)})=\tau/(1-\phi_{1,0}-\phi_{0,1}-\phi_{1,1})$, and $\mathrm{var}(y_{1,2}|y_{-(1,2)})=\tau/(1-\phi_{1,0})$, with $(1-\phi_{1,0}-\phi_{0,1}-\phi_{1,1})\,\mathrm{E}(y_{1,1}|y_{-(1,1)})=(\phi_{1,0}+\phi_{1,1})y_{2,1}+(\phi_{0,1}+\phi_{1,1})y_{1,2}+\phi_{1,1}y_{2,2}$, $(1-\phi_{1,0})\,\mathrm{E}(y_{1,2}|y_{-(1,2)})=\phi_{1,0}y_{2,2}+(\phi_{0,1}+\phi_{1,1})(y_{1,1}+y_{1,3})+\phi_{1,1}(y_{2,1}+y_{2,3})$, etc. □

The NBA reflects horizontal lags outside the lattice about 0 or n_2+1, plus changing the sign, and similarly for vertical lags. Thus, on a line site $j-g$, $g>j-1$, is replaced by site $g-j$, and site $j+g$, $g>n-j$, is replaced by site $2(n+1)-(j+g)$, with the parameter multiplied by the sign of $j-g$ or $n+1-(j+g)$, respectively (0 if $g=j$ or $n+1-j$). This implies that border sites are negative second-neighbors of themselves (e.g., site (1,1) is a horizontal and vertical second-neighbor of itself), altering the conditional means and variances for border sites for CARs with a higher order than‘2, i.e., including lag $(2,0)$ neighbors, etc. For positive dependence parameters, this reduces the conditional variance at border sites. For CARs up to order 2, the NBA is the same as the FBA.

Example H. Consider a univariate RS-CAR(3), and the NBA. Then $\mathrm{var}(y_{1,1}|y_{-(1,1)})=\tau/(1+\phi_{2,0}+\phi_{0,2})$, and $(1+\phi_{2,0}+\phi_{0,2})\mathrm{E}(y_{1,1}|y_{-(1,1)})=\phi_{1,0}y_{2,1}+\phi_{0,1}y_{1,2}+\phi_{1,1}y_{2,2}+\phi_{2,0}y_{3,1}+\phi_{0,2}y_{1,3}$, $\mathrm{var}(y_{1,2}|y_{-(1,2)})=\tau/(1+\phi_{2,0})$, and $(1+\phi_{2,0})\mathrm{E}(y_{1,2}|y_{-(1,2)})=\phi_{1,0}y_{2,2}+\phi_{0,1}(y_{1,1}+y_{1,3})+\phi_{1,1}(y_{2,1}+y_{2,3})+\phi_{2,0}y_{3,2}+\phi_{0,2}y_{1,4}$, etc. □

A theoretical possibility for getting $\mathbf{P}$ close to the planar stationarity version is to embed the $n_1\times n_2$ planar lattice in a larger torus lattice. If the torus lattice is sufficiently large, a subset of the variance matrix corresponding to an

$n_1 \times n_2$ lattice will give a good approximation to $\mathbf{P}^{-1}$. This can then be inverted to get $\mathbf{P}$. Fast routines are available for inverting a specified correlation matrix, e.g., [7, 25]. However, the torus lattice may need to be very large to obtain a good approximation to $\mathbf{P}^{-1}$ for a stationary process unless the parameters are well within the acceptable region.

One of the suggestions in [3] uses Dempster's algorithm to improve $\mathbf{P}$, which requires knowledge of the neighbor correlation. It is not clear how to extend either of their two assumptions to other CARs.

Example I. Consider Example 2.2 of [3]. This concerns a stationary univariate 1-parameter CAR(1) on a 10×10 lattice, with variance 1 and neighbor correlation 0.75, and some approximate $\mathbf{P}$ matrices. They do not compare the RBA, or the SBA. They choose $\lambda = 4\phi_{1,0}$ so that the central neighbor correlation (between sites (5,5) and (5,6), etc.) is 0.75 (for the TBA, this is also the correlation between sites (1,1) and (1,10), etc.).

For comparison with their results, the RBA ($\lambda \approx 0.995763$) has variances varying from 1 to 1.786 (corner sites), neighbor covariances varying from 0.75 to 1.300 (sites (1,1) and (1,2), etc.), and neighboring correlations varying from 0.75 to 0.8033. These are slightly better than their average "rescaling" assumption, and much better than the FBA (0.31 to 1, 0.12 to 0.75, 0.35 to 0.75, respectively). Their own "new strategy", before using Dempster's algorithm, has 0.9949 to 1.1039, 0.75 to 0.8743, and 0.75 to 0.8160, respectively. The SBA compares very badly, giving ($\lambda \approx 0.86205347$) 1 to 99.04, 0.75 to 85.27, and 0.75 to 0.9974, respectively.

If the 10×10 planar lattice is embedded in a torus lattice, it needs the torus lattice to be around 700×700 to get reasonably close (absolute differences $< 10^{-4}$) to the stationary planar correlations for lags (g_1, g_2), $g_m \in \{0, \ldots, 9\}$; $m \in \{1, 2\}$. □

Acknowledgments

LI was partially funded by the grant MIUR Ministero dell'Istruzione dell'Università e della Ricerca, PRIN research project 2015 "Environmental processes and human activities: capturing their interactions via statistical methods" - EphaStat. The authors thank ARPA Emilia Romagna for providing the data of Section 5.2, and Lucia Paci for supplying further details. They also are grateful to the Editor-in-Chief, the Associate Editor, and a referee whom they thank for helpful suggestions. RJM is grateful for support from the University G. d'Annunzio, Chieti-Pescara, Italy.

References

[1] N. Balram, J.M.F. Moura, Noncausal Gauss Markov random fields: parameter structure and estimation, IEEE Trans. Inform. Theory 39 (1993) 1333–1355.
[2] J. Besag, Errors-in-variables estimation for Gaussian lattice schemes, J. R. Stat. Soc. Ser. B Stat. Methodol. 39 (1977) 73–78.
[3] J. Besag, C. Kooperberg, On conditional and intrinsic autoregressions, Biometrika 82 (1995) 733–746.
[4] N.A. Butler, A frequency domain comparison of estimation methods for Gaussian fields, Commun. Statist. Theor. Meth. 27 (1998) 2325–2342.
[5] N.A. Butler, A frequency domain algorithm for maximum likelihood estimation method of Gaussian fields, J. Comput. Statist. Simul. 64 (1999) 151–165.
[6] N. Cressie, Statistics for spatial data, Revised edition, Wiley, New York, 1993.
[7] C. Dietrich, G. Newsam, Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix, SIAM J. Sci. Comput. 18 (1997) 1088–1107.
[8] F.P. Greco, C. Trivisano, A multivariate CAR model for improving the estimation of relative risks, Statist. Medicine 28 (2009) 1707–1724.
[9] L. Ippoliti, R.J. Martin, R.J. Bhansali, Rational spectral density models for lattice data, Spatial Statist. 6 (2013) 91–108.
[10] L. Ippoliti, L. Romagnoli, G. Arbia, A Gaussian Markov random field approach to convergence analysis, Spatial Statist. 6 (2013) 78–90.
[11] X. Jin, S. Banerjee, B.P. Carlin, Order-free co-regionalized areal data models with application to multiple-disease mapping, J. R. Stat. Soc. Ser. B Stat. Methodol. 65 (2007) 817–838.
[12] X. Jin, B.P. Carlin, S. Banerjee, Generalized hierarchical multivariate CAR models for areal data, Biometrics 61 (2005) 950–961.
[13] H. Kim, D. Sun, R.K. Tsutakawa, A bivariate Bayes method for improving the estimates of mortality rates with a twofold conditional autoregressive model, J. Amer. Statist. Assoc. 96 (2001) 1506–1521.
[14] S. Lakshmanan, H. Derin, Valid parameter space for 2-D Gaussian Markov random fields, IEEE Trans. Inform. Theory 39 (1993) 703–709.
[15] K.V. Mardia, Multi-dimensional multivariate Gaussian Markov random fields with applications to image processing, J. Multivariate Anal. 24 (1988) 265–284.
[16] R.J. Martin, Exact maximum likelihood for incomplete data from a correlated Gaussian process, Commun. Statist. Theor. Meth. 13 (1984) 1275–1288.
[17] R.J. Martin, J. Dwyer, Approximations to the covariance properties of processes averaged over irregular spatial regions. Commun. Statist. Theor. Meth. 23 (1994) 913–945.
[18] M.A. Martinez-Beneito, A general modelling framework for multivariate disease mapping, Biometrika 100 (2013) 539–553.

[19] J.M.F. Moura, M.G.S. Bruno, DCT/DST and Gauss-Markov fields: conditions for equivalence, IEEE Trans. Signal Process 46 (1998) 2571–2574.
[20] H. Rue, L. Held, Gaussian Markov Random Fields:Theory and Applications, Chapman and Hall/CRC, Boca Raton, FL, 2005.
[21] S.R. Sain, N. Cressie, A spatial model for multivariate lattice data, J. Econometrics 140 (2007) 226–259.
[22] S.R. Sain, R. Furrer, N. Cressie, A spatial analysis of multivariate output from regional climate models, Ann. Appl. Statist. 5 (2011) 150–175.
[23] M. Stortini, Il sistema modellistico NINFA2015 per la qualità dell'aria. URL: `http://www.arpae.it/cms3/documenti/_cerca_doc/meteo/ambiente/ninfa2015.pdf`, 2015.
[24] The Mathworks, `Matlab, version 7.5.0.342`, Natick, MA, 2007.
[25] A.T.A. Wood, G. Chan, Simulation of stationary Gaussian processes in $[0; 1]^d$, J. Comput. Graph. Statist. 3 (1994) 409–432.