



Comments on: Some recent work on multivariate Gaussian Markov random fields

Fedele Greco¹ · Carlo Trivisano¹

Published online: 12 September 2018
© Sociedad de Estadística e Investigación Operativa 2018

1 Introduction

It is very stimulating to comment on this paper, which offers a very comprehensive review on multivariate GMRFs. Most of the latest developments on this field are discussed, and connections between them are efficiently highlighted. The coregionalisation framework plays a prominent role over the whole paper as well as in theory development, because it allows natural extension from a univariate setting to a multivariate setting, providing both theoretical advances and hints for efficient computational strategies. The entanglement of spatial and non-spatial components, which generates challenges from both theoretical and computational point of view, is well discussed, and we definitely agree with the author in claiming that it is a crucial topic in multivariate modelling of lattice data and in spatial data modelling in general.

In this brief commentary, we focus on a couple of neglected features concerning details that might be considered in prior specification and that may have impact on the results, particularly when dealing with disease mapping, where the amount of information is often very limited and posterior shrinkage is extremely sensitive to prior assumptions: in this framework, observed differences between models can be simply driven by different prior assumptions, and small changes to priors can cause changes in posterior inference, favouring different models. The discussion is developed with a notation coherent with the one used in MacNab's paper. Here, we briefly summarise the models discussed in this commentary. We denote as $\text{vec}(\boldsymbol{\eta})$ the np -dimensional vector of spatially interacting latent variables and with $\text{vec}(\boldsymbol{\zeta}) = (\mathbf{A} \otimes \mathbf{I}_n)\text{vec}(\boldsymbol{\eta})$ the vector of coregionalised variables, where $\mathbf{A}\mathbf{A}^T = \boldsymbol{\Sigma}$ is the $p \times p$ area-wise (non-spatial)

This comment refers to the invited paper available at <https://doi.org/10.1007/s11749-018-0605-3>.

✉ Carlo Trivisano
carlo.trivisano@unibo.it

Fedele Greco
fedele.greco@unibo.it

¹ Department of Statistical Sciences, University of Bologna, Bologna, Italy

covariance matrix. The joint precision matrix of the latent variables is:

$$\boldsymbol{\Omega}_{\text{vec}(\eta)}(\mathbf{C}) = \left(\mathbf{I}_p \otimes \mathbf{D}_m - \mathbf{C} \otimes \mathbf{W}_U - \mathbf{C}^\top \otimes \mathbf{W}_U^\top \right). \quad (1)$$

This delivers the following precision matrix of the coregionalised variables $\text{vec}(\boldsymbol{\zeta})$:

$$\boldsymbol{\Omega}_{\text{vec}(\boldsymbol{\zeta})}(\mathbf{C}, \mathbf{A}) = (\mathbf{A} \otimes \mathbf{I}_n) \boldsymbol{\Omega}_{\text{vec}(\eta)}(\mathbf{C}) (\mathbf{A}^\top \otimes \mathbf{I}_n) \quad (2)$$

Three different structures are considered for the spatial cross-correlation matrix \mathbf{C} , namely a diagonal structure denoted as $\mathbf{C}_d = c\mathbf{I}_p$, a non-diagonal but symmetric structure denoted as \mathbf{C}_s and a non-diagonal and asymmetric structure denoted as \mathbf{C}_a . As regards the covariance matrix $\boldsymbol{\Sigma}$, we consider both the Inverse Wishart prior and the prior based on Givens angles (GA), as proposed in Daniels and Kass (2012). Note that the same GA prior, with suitable modification to deal with positive definiteness conditions for (1), is used for both \mathbf{C}_s and \mathbf{C}_a .

Sensitivity to prior assumption is a crucial topic in Bayesian analysis, and we believe that this should deserve particular care in the context of Disease Mapping, where the noise might often be overwhelming with respect to empirical evidence, in particular when the diseases under study are rare. We provide a short discussion of:

- the effect of priors on model selection, by means of DIC;
- the effect of using priors based on Givens angles in terms of prior weights assigned to diagonal and off-diagonal elements of \mathbf{C}_s .

2 Prior choice and model comparison in terms of DIC

DIC is the most widely used model comparison tool in Bayesian applications, also because of its availability in BUGS. Nonetheless, DIC has been criticised as a model comparison tool, particularly when richly parameterised models, as MGMRFs, are involved: as an example, Chan and Grant (2016) argues that, in such cases, DIC may favour more complex models, i.e. it may favour overfitting. This should be considered when comparing multivariate models with univariate models in the context of Disease Mapping. A very interesting discussion with useful insights regarding model comparison is provided in Gelman et al. (2014), where WAIC is suggested as a promising model comparison measure, but with some pitfalls in spatial modelling, since it relies on a data partition.

To show sensitivity of model selection on prior assumptions, we present an application referred to multivariate modelling of $p = 3$ diseases: acute ischaemic heart diseases (D1), hypertension (D2) and chronic ischaemic heart disease (D3). Data refer to death counts observed from 1998 to 2001 in $n = 95$ municipalities of the Emilia-Romagna Region (Italy). The diseases show very different incidence, with total observed counts equal to 4315, 1670 and 171, respectively.

Table 1 Model comparison (DIC) with two different priors on Σ

	IW			Givens		
	C_d	C_a	C_s	C_d	C_a	C_s
D1	611.9	610.2	610.3	608.6	605.8	600.7
D2	494.7	491.7	492.6	496.5	495.4	491.8
D3	245.1	243.3	245.7	242.1	240.0	241.1
Total	1351.7	1345.2	1348.6	1347.2	1341.2	1333.6

2.1 Effect of prior on Σ

Table 1 reports model comparison in terms of DIC obtained with two different prior choices for Σ , namely $\Sigma \sim IW(I_p, p)$ and the GA prior specification. The latter starts by the spectral decomposition $\Sigma = P \Lambda P^T$ where P is the orthogonal matrix of the eigenvectors of Σ , and Λ is a diagonal matrix with ordered eigenvalues along the diagonal. Matrix P is parameterised in terms of $p(p - 1)/2$ Givens angles θ_{ij} , $i = 1, \dots, p - 1$, $j = i + 1, \dots, p$. Hence, $P = G_{12}G_{13} \dots G_{1p} \dots G_{(p-1)p}$ where G_{ij} is the $p \times p$ identity matrix with the i -th and j -th diagonal entries replaced by $\cos(\theta_{ij})$ and elements ij and ji replaced by $\pm \sin(\theta_{ij})$, respectively. A Uniform($-\pi/2, \pi/2$) prior is selected for the Givens angles, while a Uniform(0, 10) is specified for the eigenvalues.

Table 1 shows that, not surprisingly, DIC values change under the two specifications. Most importantly, a different model is selected in the two scenarios, with the model using the asymmetric C_a matrix preferred when using the IW prior and the model using the symmetric C_s matrix selected when using the GA prior. For a comprehensive discussion on the differences between IW and GA priors, see Daniels and Kass (2012).

2.2 Effect of Givens angles prior on C

In this section, we discuss some issues related to the GA prior when it is employed for the C_s matrix, but the same rationale applies to the specification of the prior on C_a , which is based on singular value decomposition $C_a = P S Q^T$. This requires a second set of priors on Givens angles to be specified for Q . Matrix S is a diagonal matrix of singular values: the prior on singular values needs to be specified such that matrix (1) is positive definite (see Greco and Trivisano 2009). The prior on C_s has been employed in the context of disease mapping by Jin et al. (2007) in the so-called order-free model. The prior is built following the same procedure outlined in Sect. 2.1, with the difference that diagonal elements of the eigenvalue matrix are specified as uniform distributions with their range fixed in order to be in compliance with positive definiteness conditions for the joint covariance matrix (1). In our case study, such range is $(-1.585, 1)$, where -1.585 is the reciprocal of the smallest eigenvalue of $D_m^{-\frac{1}{2}} W D_m^{-\frac{1}{2}}$.

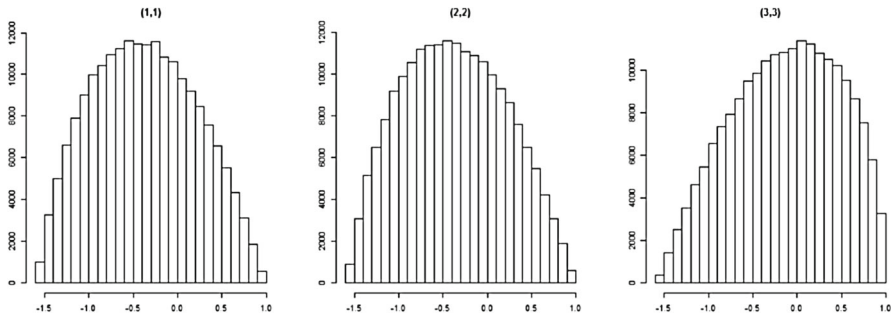


Fig. 1 Prior distribution on the diagonal entries of C_s when $p = 3$

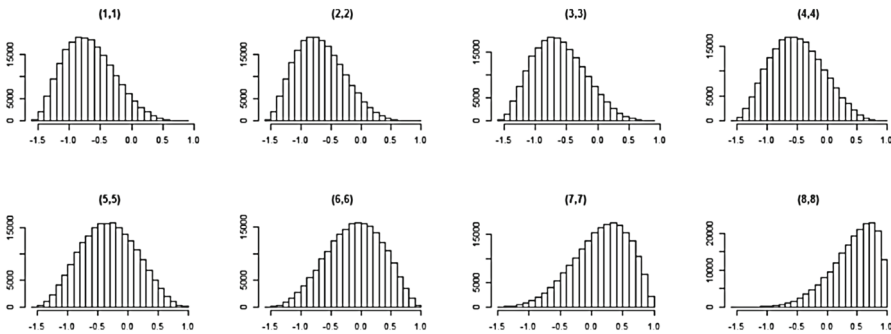


Fig. 2 Prior distribution on the diagonal entries of C_s when $p = 8$

As can be seen by simulation, the prior induces different prior distributions on the diagonal entries of C_s , as well as on out-of-diagonal elements; in this sense, the prior is not completely order-free. This is shown in Fig. 1, where simulations from the prior are reported for the diagonal entries of C_s : it can be seen that the prior is not exactly equivalent for such elements, with a left-skewed prior for $C_s(1, 1)$ and a right-skewed prior for $C_s(3, 3)$. Again, this might affect estimation results in a context where information provided by data is weak, as is the case in our application. As regards disease D3, results show a different amount of shrinkage, as measured by the standard deviation of the posterior means, when data on this disease are considered in the model as the first or the third data vector. We note that this pattern becomes more and more evident, for both diagonal (Fig. 2) and off-diagonal elements (not shown), when p increases.

3 Computational tools

As a concluding remark, we suggest that estimation of MGMRF models in the context of Disease Mapping would take great advantage from carefully programmed MCMC algorithms collected in an R package. While it is true that BUGS offers a very friendly interface for estimating Bayesian models accessible to practitioners, BUGS coding of

MGMRF models can become cumbersome and, more importantly, it does not allow to take full advantage of the sparseness of matrix (1). Such sparseness and the availability of very efficient packages for matrix algebra as `RcppArmadillo` could broaden the scope of MCMC algorithms coded in BUGS, allowing for both faster computation and the possibility to manage larger n and p .

References

- Chan JCC, Grant AL (2016) On the observed-data deviance information criterion for volatility modeling. *J Financ Econom* 14(4):772–802
- Daniels MJ, Kass RE (2012) Nonconjugate Bayesian estimation of covariance matrices and its use in hierarchical models. *J Am Stat Assoc* 94(448):1254–1263
- Gelman A, Hwang J, Vehtari A (2014) Understanding predictive information criteria for Bayesian models. *Stat Comput* 24(6):997–1016
- Greco F, Trivisano C (2009) A multivariate CAR model for improving the estimation of relative risks. *Stat Med* 28:1707–1724
- Jin X, Carlin BP, Banerjee S (2007) Order-free co-regionalized areal data models with application to multiple-disease mapping. *J R Stat Soc Ser B* 69(5):817–838