Statistica Sinica 20 (2010): Supplement

A GENERALIZED CONVOLUTION MODEL FOR MULTIVARIATE NONSTATIONARY SPATIAL PROCESSES

Anandamayee Majumdar, Debashis Paul and Dianne Bautista

Arizona State University, University of California, Davis, and Ohio State University, Columbus

Supplementary Material

S1. Technical details

Proof of Theorem 2

We need to verify that $C^\star(\cdot,\cdot)$ thus defined is indeed a covariance kernel. To show this, it is enough to show that for any finite set of spatial points $s_1,\ldots,s_m \in \mathbb{R}^d$, for any $m \geq 1$, $((C(s_l,s_{l'})))_{l,l'=1}^m$ is nonnegative definite. Since this is true from Theorem 1, and since given any valid covariance function on $\mathbb{R}^d \times \mathbb{R}^d$, there exists a mean-zero Gaussian spatial process $Y(s)$ which yields the same covariance function, the proof is completed by *Skorohod's Representation Theorem* and *Kolmogorov's Consistency Theorem* (Billingsley, 1999).

Proof of Theorem 3

Referring to Theorem 1, we only need to check the conditions on the functions $\{\rho_{jj'} : 1 \leq j, j' \leq N\}$ that ensure that the $N \times N$ matrix $\mathbf{R}(\omega)$ is positive definite (nonnegative definite). Note that since $\rho_{jj'} = \overline{\rho}_{j'j}$,

$$\mathbf{R}(\omega) = \begin{bmatrix} 1 & \rho_{12}(\omega) & \cdots & \rho_{1k}(\omega) \\ \overline{\rho_{12}(\omega)} & 1 & \cdots & \rho_{2k}(\omega) \\ \cdot & \cdot & \cdots & \cdot \\ \overline{\rho_{1k}(\omega)} & \overline{\rho_{2k}(\omega)} & \cdots & 1 \end{bmatrix}$$

The proof is based on the following well-known inversion formula for partitioned non-singular matrices:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A-BD^{-1}C)^{-1} & -A^{-1}B(D-CA^{-1}B)^{-1} \\ -(D-CA^{-1}B)^{-1}CA^{-1} & (D-CA^{-1}B)^{-1} \end{bmatrix},$$

where the condition for invertibility is that both $A-BD^{-1}C$ and $D-CA^{-1}B$ are invertible. If the matrix is Hermitian, i.e., $A^* = A$, $B^* = C$, $D^* = D$, then the necessary and sufficient condition for positive definiteness (semidefiniteness) of the matrix on the LHS is that one of the following holds:

- A and $D - CA^{-1}B$ are positive definite (semidefinite);
- D and $A - BD^{-1}C$ are positive definite (semidefinite).

In order to prove *Theorem 3*, we apply the last condition for 2×2, 3×3 and 4×4 principal submatrices of $\mathbf{R}(\omega)$.

Without loss of generality we take j, k, l, m to be 1, 2, 3 and 4, respectively. Then condition (i), viz. $1 - |\rho_{12}|^2 > 0$ is immediate.

For (ii), observe that we need

$$1 - \begin{bmatrix} \rho_{12} \\ \rho_{13} \end{bmatrix}^T \begin{bmatrix} 1 & \rho_{23} \\ \rho_{32} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_{21} \\ \rho_{31} \end{bmatrix} = 1 - \frac{1}{1 - |\rho_{23}|^2} \begin{bmatrix} \rho_{12} \\ \rho_{13} \end{bmatrix}^T \begin{bmatrix} 1 & -\rho_{23} \\ -\rho_{32} & 1 \end{bmatrix} \begin{bmatrix} \rho_{21} \\ \rho_{31} \end{bmatrix} > 0,$$

which translates into (ii) after a simplification. Note also that the quantity appearing in (ii) is really the determinant of the 3×3 principal submatrix (corresponding to rows 1,2, and 3) of $\mathbf{R}(\omega)$.

For (iii), we first consider the following scaler

$$\begin{aligned} & \begin{bmatrix} \rho_{12} \\ \rho_{13} \\ \rho_{14} \end{bmatrix}^T \begin{bmatrix} 1 & \rho_{23} & \rho_{24} \\ \rho_{32} & 1 & \rho_{34} \\ \rho_{42} & \rho_{43} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_{21} \\ \rho_{31} \\ \rho_{41} \end{bmatrix} \\ = & \frac{1}{\Delta_{234}} \begin{bmatrix} \rho_{12} \\ \rho_{13} \\ \rho_{14} \end{bmatrix}^T \begin{bmatrix} 1 - |\rho_{34}|^2 & \rho_{23} - \rho_{24}\rho_{43} & \rho_{24} - \rho_{23}\rho_{34} \\ \rho_{32} - \rho_{34}\rho_{42} & 1 - |\rho_{24}|^2 & \rho_{34} - \rho_{32}\rho_{24} \\ \rho_{42} - \rho_{43}\rho_{32} & \rho_{43} - \rho_{42}\rho_{23} & 1 - |\rho_{23}|^2 \end{bmatrix} \begin{bmatrix} \rho_{21} \\ \rho_{31} \\ \rho_{41} \end{bmatrix} \end{aligned}$$

where $\Delta_{234} = 1 - |\rho_{23}|^2 - |\rho_{24}|^2 - |\rho_{34}|^2 + 2\mathrm{Re}(\rho_{23}\rho_{34}\rho_{42})$ is the determinant of the $(2,3,4)$ submatrix of $\mathbf{R}(\omega)$. In order that the 4×4 principal submatrix corresponding to $(1,2,3,4)$ is positive definite (semidefinite) we need that the quantity in the display above is less than $(\leq)$ 1. Some straightforward algebra yields condition (iii).

Expression for $\Gamma_{jj'}(s; \widetilde{\theta}_l, \widetilde{\theta}_{l'}, \kappa)$

$$\begin{aligned} & \Gamma_{12}(s; \widetilde{\theta}_l, \widetilde{\theta}_{l'}, \kappa) = \Gamma_{21}(s; \widetilde{\theta}_{l'}, \widetilde{\theta}_l, \kappa) \\ = & \frac{1}{2\sqrt{\pi}} \left[g_1(\alpha_1, \widetilde{\theta}_l, \widetilde{\theta}_{l'}) e^{-g_2(\alpha_1, \widetilde{\theta}_l, \widetilde{\theta}_{l'}) \|s\|^2} - \beta g_1(\alpha_2, \widetilde{\theta}_l, \widetilde{\theta}_{l'}) e^{-g_2(\alpha_2, \widetilde{\theta}_l, \widetilde{\theta}_{l'}) \|s\|^2} \right], \end{aligned} \tag{S1.1}$$

$$\Gamma_{jj}(s; \widetilde{\theta}_l, \widetilde{\theta}_{l'}; \kappa) = \frac{1}{2\sqrt{\pi}} \frac{1}{\sqrt{\widetilde{\theta}_l + \widetilde{\theta}_{l'}}} \exp\left(-\frac{\widetilde{\theta}_l \widetilde{\theta}_{l'}}{\widetilde{\theta}_l + \widetilde{\theta}_{l'}} \parallel s \parallel^2 \right), \quad j = 1, 2. \tag{S1.2}$$

where, for $k = 1, 2$, $1 \leq l, l' \leq L$,

$$g_1(\alpha_k, \widetilde{\theta}_l, \widetilde{\theta}_{l'}) = \sqrt{\frac{\alpha_k}{\alpha_k(\widetilde{\theta}_l + \widetilde{\theta}_{l'}) + \widetilde{\theta}_l \widetilde{\theta}_{l'}}}, \quad g_2(\alpha_k, \widetilde{\theta}_l, \widetilde{\theta}_{l'}) = \frac{\alpha_k \widetilde{\theta}_l \widetilde{\theta}_{l'}}{\alpha_k(\widetilde{\theta}_l + \widetilde{\theta}_{l'}) + \widetilde{\theta}_l \widetilde{\theta}_{l'}}.$$

S2. Gibbs sampling

Here we provide a detailed derivation of the posterior distributions that were used for the developments in Section 3.1. We focus on the parameters in model (12) - (15).

Let $\mathbf{y}^T = (\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_N^T)^T$ denote the vector of measurements arranged in sub-vectors corresponding to each process and ordered by the location-vector $\mathbf{s} = (s^{(1)}, \ldots, s^{(d)})$, with

$$s^{(i)} = (s_1^{(i)}, \ldots, s_n^{(i)})^T, \quad i = 1, \ldots, d,$$

and

$$\mathbf{y}_j = (y_{j1}, \ldots, y_{jn})^T, \quad j = 1, \ldots, N.$$

The vector of parameters for data from model (12) - (15)

$$\boldsymbol{\theta}^T = (c_{11}, \ldots, c_{NL}, \widetilde{\theta}_{11}, , \ldots, \widetilde{\theta}_{NL}, \Sigma_l, \ldots, \Sigma_l, \{\nu_{jj'} \ 1 \leq j < j' \leq N\}, \alpha_1, \alpha_2, \beta, \tau) \tag{S2.3}$$

has the associated likelihood function

$$L(\mathbf{y}|\boldsymbol{\theta}) \propto |C^\star|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}^T C^{\star-1}\mathbf{y}\right), \tag{S2.4}$$

where $C^\star$ denotes the $Nn \times Nn$ matrix of the spatial variance-covariance matrix whose elements are defined by (12) - (15). The developments in Section 3.1 now entail that the joint distribution of $\mathbf{y}$ and $\boldsymbol{\theta}$ is the product of $L(\mathbf{y}|\boldsymbol{\theta})$ with the prior density

$$\pi(\boldsymbol{\theta}) = \pi(\nu_{jj'}, 1 \leq j < j' \leq N) \prod_{l=1}^{L} \left(\pi(\Sigma_l) \prod_{\substack{j<j' \\ j,j'=1}}^{N} \pi(\widetilde{\theta}_{jl}) \pi(c_{jl})\right) \pi(\alpha_1)\pi(\alpha_2)\pi(\beta).$$

The next step is to derive the posterior densities for each of the components of (S2.3) that will allow us to carry out Bayesian inference via the Gibbs sampler (e.g., Robert and Casella, 2004, Ch. 10). For this development we shall employ the symbol "\" to indicate removal of a particular parameter or a group of parameters from $\boldsymbol{\theta}$.

We begin the marginalization process by noting that the posterior density for parameter τ is

$$\pi(\tau|\mathbf{y}, \boldsymbol{\theta}\backslash\tau) \quad \propto \quad |C^\star|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}^T C^{\star-1}\mathbf{y}\right) \times \exp\left(-b_\tau \tau\right) \tau^{a_\tau - 1}.$$

The posterior distributions of $\alpha_k, \ k = 1, 2$ would follow the posterior density:

$$\pi(\alpha_k|\mathbf{y}, \boldsymbol{\theta}\backslash\alpha_k) \quad \propto \quad |C^\star|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}^T C^{\star-1}\mathbf{y}\right) \exp\left(-b_{\alpha_k}\alpha_k\right) {\alpha_k}^{a_{\alpha_k} - 1}.$$

For the posterior distribution of the local, process-dependent decay parameter $\widetilde{\theta}_{jl}, \ j = 1, \ldots, N, \ l = 1, \ldots, L$, we obtain

$$\pi(\widetilde{\theta}_{jl}|\mathbf{y}, \boldsymbol{\theta}\backslash\widetilde{\theta}_{jl}) \quad \propto \quad |C^\star|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}^T C^{\star-1}\mathbf{y}\right) \exp\left(-b_{\widetilde{\theta}_{jl}}\widetilde{\theta}_{jl}\right) (\widetilde{\theta}_{jl})^{a_{\widetilde{\theta}_{jl}} - 1}.$$

Similarly, the local process-dependent scale parameters c_{jl} have a posterior distribution

$$\pi(c_{jl}|\mathbf{y}, \boldsymbol{\theta}\backslash c_{jl}) \quad \propto \quad |C^\star|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}^T C^{\star-1}\mathbf{y}\right) \exp\left(-b_{c_{jl}} c_{jl}\right) (c_{jl})^{a_{c_{jl}} - 1}.$$

The posterior distribution of β is proportional to the likelihood in (S2.4) times the appropriate indicator function corresponding to the specific range, since the prior specification of this parameter is a Uniform distribution.

Since the mapping $\nu^\star \to (\mathbf{N}, diag(\nu^\star))$ is one-to-one, we view the prior (posterior) of $\mathbf{N}$ (equivalently $(\nu_{jj'})_{j<j'}$) as a marginal (corresponding to $\mathbf{N}$) of the prior (posterior) of $(\mathbf{N}, (\nu^\star_{jj})_{j=1}^N)$. Note also that the posterior distribution of $\nu^\star$ is

$$\begin{aligned}\pi(\nu^\star|\mathbf{y}, \boldsymbol{\theta}\backslash\mathbf{N}) \quad &\propto \quad |C^\star|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}^T {C^\star}^{-1}\mathbf{y}\right) \\ &\times \exp\left(-\frac{1}{2}\text{trace}(\tilde{\nu}^{-1}{\nu^\star}^{-1})\right) \det(\nu^\star)^{-(d+1)/2}\end{aligned}$$

In order to sample from the posterior distribution of $\mathbf{N}$ we just need to sample from the posterior distribution of $\nu^\star$ and use the normalization $\mathbf{N} = diag(\nu^\star)^{-\frac{1}{2}}\nu^\star diag(\nu^\star)^{-\frac{1}{2}}$.

Finally, the posterior for $\Sigma_l, l = 1, \ldots, L$ can be deduced from

$$\begin{aligned}\pi(\Sigma_l|\mathbf{y}, \boldsymbol{\theta}\backslash\Sigma_l) \quad &\propto \quad |C^\star|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{y}^T {C^\star}^{-1}\mathbf{y}\right) \\ &\times \exp\left(-\frac{1}{2}\text{trace}(\Psi^{-1}{\Sigma_l}^{-1})\right) \det(\Sigma_l)^{-(2+1)/2}\end{aligned}$$

S3. Sample realization from models for simulation

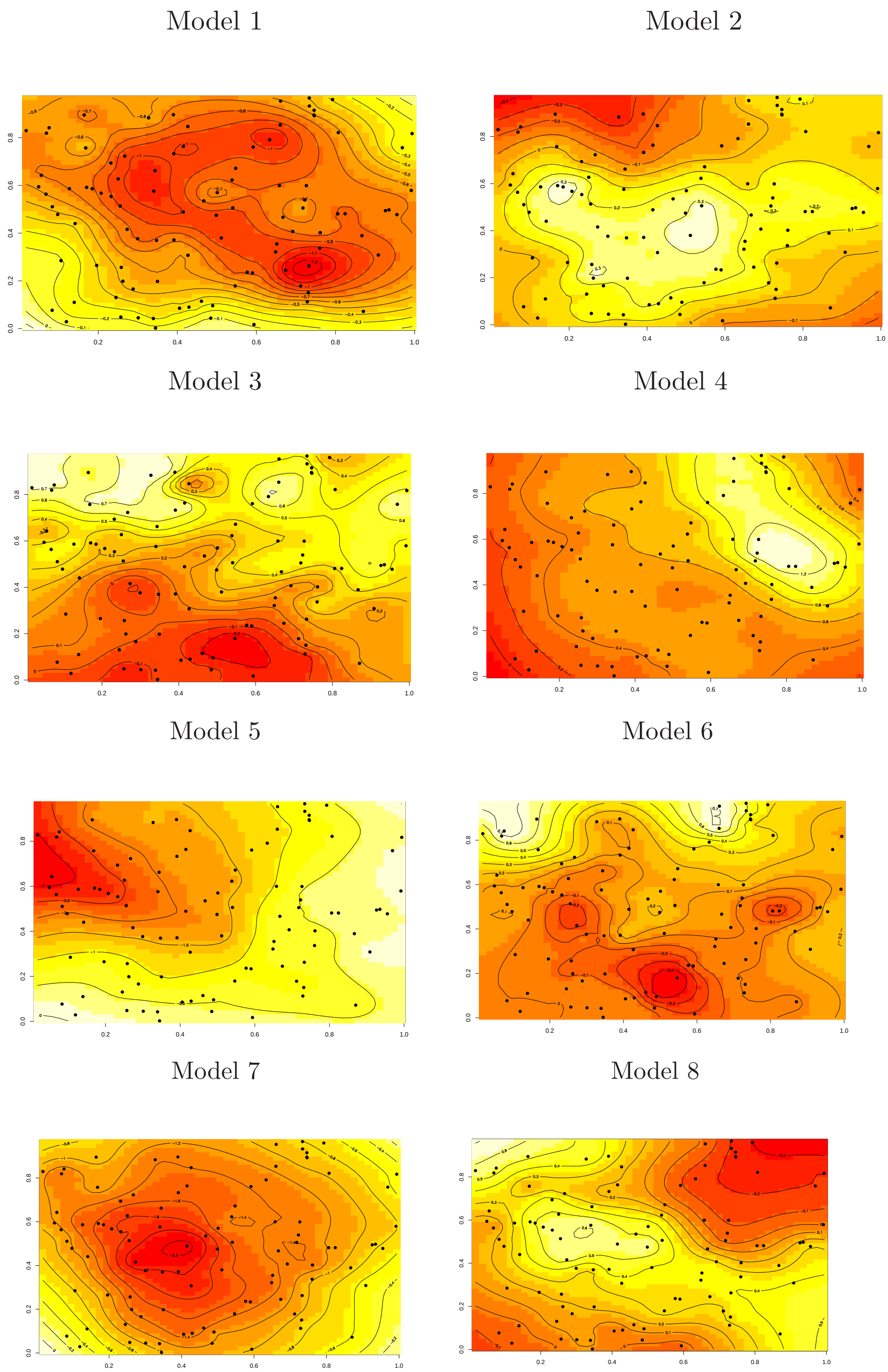

Figure S3.1: Realization of the first coordinate process (Y_1) under the 8 different models