WEB SUPPLEMENTARY MATERIAL FOR "A BAYESIAN HIERARCHICAL SPATIAL POINT PROCESS MODEL FOR MULTI-TYPE NEUROIMAGING META-ANALYSIS"

JIAN KANG, TIMOTHY D. JOHNSON, THOMAS E. NICHOLS, AND TOR D. WAGER

1. HPGRF Model

In this Appendix, we provide proofs for all the theorems that are presented in the paper for the HPGRF model, except for Theorem 2. A proof of Theorem 2 is provided in Wolpert and Ickstadt (1998). We first introduce the following lemmas.

Lemma 1. *If* $\mathbf{Y} \sim \mathcal{PP}(\mathcal{B}, \Lambda)$*, then for any* $A, B \subset \mathcal{B}$*, we have*

$$\mathrm{Cov}\{N_{\mathbf{Y}}(A), N_{\mathbf{Y}}(B)\} = \mathrm{Var}\{N_{\mathbf{Y}}(A \cap B)\} = \Lambda(A \cap B).$$

Proof. Note that $N_{\mathbf{Y}}(A) = N_{\mathbf{Y}}(A \cap B) + N_{\mathbf{Y}}(A \setminus B)$ and $N_{\mathbf{Y}}(B) = N_{\mathbf{Y}}(A \cap B) + N_{\mathbf{Y}}(B \setminus A)$. By the independence property of the Poisson process, we have that

$$\begin{aligned} \mathrm{Cov}\{N_{\mathbf{Y}}(A), N_{\mathbf{Y}}(B)\} &= \mathrm{Cov}\{N_{\mathbf{Y}}(A \cap B) + N_{\mathbf{Y}}(A \setminus B), N_{\mathbf{Y}}(A \cap B) + N_{\mathbf{Y}}(B \setminus A)\} \\ &= \mathrm{Cov}\{N_{\mathbf{Y}}(A \cap B), N_{\mathbf{Y}}(A \cap B)\} = \mathrm{Var}\{N_{\mathbf{Y}}(A \cap B)\}. \qquad \square \end{aligned}$$

Lemma 2. *Let* $\Gamma(d\mathrm{x}) \sim \mathcal{GRF}\{\alpha(d\mathrm{x}), \beta\}$*, and let* $f_1(\mathrm{x})$ *and* $f_2(\mathrm{x})$ *be measurable functions on* $\mathcal{B}$*, Then*

$$\mathrm{Cov}\left\{\int_{\mathcal{B}} f_1(\mathrm{x})\Gamma(d\mathrm{x}), \int_{\mathcal{B}} f_2(\mathrm{x})\Gamma(d\mathrm{x})\right\} = \frac{1}{\beta^2}\int_{\mathcal{B}} f_1(\mathrm{x}) f_2(\mathrm{x})\alpha(d\mathrm{x}).$$

Proof. We start with the case that $f_1(\mathrm{x})$ and $f_2(\mathrm{x})$ are simple functions on $\mathcal{B}$, i.e. for $m = 1, \ldots, M$ and $n = 1, \ldots, M$, there exist numbers $a_m, b_n \in \mathbb{R}$ and disjoint sets A_m and B_n with $\mathcal{B} = \bigcup_{m=1}^M A_m = \bigcup_{n=1}^N B_n$ such that $f_1(\mathrm{x}) = \sum_{m=1}^M a_m \delta_{A_m}(\mathrm{x})$ and $f_2(\mathrm{x}) = \sum_{n=1}^N b_n \delta_{B_n}(\mathrm{x})$. Then $\int_{\mathcal{B}} f_1(\mathrm{x})\Gamma(d\mathrm{x}) = \sum_{m=1}^M a_m \Gamma(A_m)$ and $\int_{\mathcal{B}} f_2(\mathrm{x})\Gamma(d\mathrm{x}) =$

$\sum_{n=1}^{N} b_n\Gamma(B_n)$. Thus,

$$\mathrm{E}\left\{\int_{\mathcal{B}} f_1(\mathrm{x})\Gamma(d\mathrm{x}) \times \int_{\mathcal{B}} f_2(\mathrm{x})\Gamma(d\mathrm{x})\right\} = \mathrm{E}\left\{\sum_{m=1}^{M} a_m\Gamma(A_m) \times \sum_{n=1}^{N} b_n\Gamma(B_n)\right\}$$

$$\begin{aligned}
&= \sum_{m=1}^{M}\sum_{n=1}^{N} a_m b_n \mathrm{E}\{\Gamma^2(A_m \cap B_n)\} \\
&\quad + \sum_{(m,n)\neq(m',n')} a_m b_{n'} \mathrm{E}\{\Gamma(A_m \cap B_n)\}\mathrm{E}\{\Gamma(A_{m'} \cap B_{n'})\} \\
&= \frac{1}{\beta^2}\sum_{m=1}^{M}\sum_{n=1}^{N} a_m b_n \alpha(A_m \cap B_n) \\
&\quad + \frac{1}{\beta^2}\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{m'=1}^{M}\sum_{n'=1}^{N} a_m b_{n'} \alpha(A_m \cap B_n)\alpha(A_{m'} \cap B_{n'}) \\
&= \frac{1}{\beta^2}\int_{\mathcal{B}} f_1(\mathrm{x})f_2(\mathrm{x})\alpha(d\mathrm{x}) + \frac{1}{\beta^2}\int_{\mathcal{B}} f_1(\mathrm{x})\alpha(d\mathrm{x})\int_{\mathcal{B}} f_2(\mathrm{x})\alpha(d\mathrm{x}).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\mathrm{Cov}\left\{\int_{\mathcal{B}} f_1(\mathrm{x})\Gamma(d\mathrm{x}), \int_{\mathcal{B}} f_2(\mathrm{x})\Gamma(d\mathrm{x})\right\} \\
&\quad = \mathrm{E}\left\{\int_{\mathcal{B}} f_1(\mathrm{x})\Gamma(d\mathrm{x}) \times \int_{\mathcal{B}} f_2(\mathrm{x})\Gamma(d\mathrm{x})\right\} - \frac{1}{\beta^2}\int_{\mathcal{B}} f_1(\mathrm{x})\alpha(d\mathrm{x})\int_{\mathcal{B}} f_2(\mathrm{x})\alpha(d\mathrm{x}) \\
&\quad = \frac{1}{\beta^2}\int_{\mathcal{B}} f_1(\mathrm{x})f_2(\mathrm{x})\alpha(d\mathrm{x}).
\end{aligned}$$

For general measurable functions, $f_1(\mathrm{x})$ and $f_2(\mathrm{x})$, a routine passage to the limit completes the proof. □

Theorem 1. *Within emotion type j and for all $A, B \subseteq \mathcal{B}$,*

$$E\{N_{\mathbf{Y}_j}(A) \mid \sigma_j^2, \tau, \alpha, \beta\} = \frac{1}{\tau\beta}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})\alpha(d\mathrm{x}).$$

$$\begin{aligned}
&\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_j}(B) \mid \sigma_j^2, \tau, \alpha, \beta\} \\
&\quad = \frac{1}{\tau\beta}\int_{\mathcal{B}} K_{\sigma_j^2}(A \cap B, \mathrm{x})\alpha(d\mathrm{x}) + \frac{1+\beta}{\tau^2\beta^2}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})K_{\sigma_j^2}(B, \mathrm{x})\alpha(d\mathrm{x}).
\end{aligned} \tag{1}$$

Between emotion types j and k ($j \neq k$),

$$\begin{aligned}
&\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_k}(B) \mid \sigma_j^2, \sigma_k^2, \tau, \alpha, \beta\} \\
&\quad = \frac{1}{\tau^2\beta^2}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})K_{\sigma_k^2}(B, \mathrm{x})\alpha(d\mathrm{x}).
\end{aligned} \tag{2}$$

Proof. First, we note that given G_0, τ and σ_j^2, the conditional expectation of $N_{\mathbf{Y}_j}(A)$ is

$$\begin{aligned}\mathrm{E}\{N_{\mathbf{Y}_j}(A) \mid G_0, \tau, \sigma_j^2\} &= \mathrm{E}_{G_j}\{\mathrm{E}\{N_{\mathbf{Y}_j}(A) \mid G_j\} \mid G_0, \tau, \sigma_j^2\} \\ &= \int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})\mathrm{E}\left\{G_j(d\mathrm{x}) \mid G_0, \tau\right\} = \frac{1}{\tau}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})G_0(d\mathrm{x}).\end{aligned}$$

Also, the conditional covariance between $N_{\mathbf{Y}_j}(A)$ and $N_{\mathbf{Y}_j}(B)$ is

$$\begin{aligned}&\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_j}(B) \mid G_0, \tau, \sigma_j^2\} \\ &\quad= \mathrm{E}_{G_j}\{\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_j}(B) \mid G_j\} \mid G_0, \tau, \sigma_j^2\} \\ &\qquad\qquad + \mathrm{Cov}_{G_j}\{\mathrm{E}\{N_{\mathbf{Y}_j}(A) \mid G_j\}, \mathrm{E}\{N_{\mathbf{Y}_j}(B) \mid G_j\} \mid G_0, \tau, \sigma_j^2\} \\ &\quad= \frac{1}{\tau}\int_{\mathcal{B}} K_{\sigma_j^2}(A\cap B, \mathrm{x})G_0(d\mathrm{x}) + \frac{1}{\tau^2}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})K_{\sigma_j^2}(B, \mathrm{x})G_0(d\mathrm{x}).\end{aligned}$$

Now, given σ_j^2, α, β, within type j, for any $A \subseteq \mathcal{B}$,

$$\begin{aligned}\mathrm{E}\{N_{\mathbf{Y}_j}(A) \mid \sigma_j^2, \tau, \alpha, \beta\} &= \mathrm{E}_{G_0}\{\mathrm{E}\{N_{\mathbf{Y}_j}(A) \mid G_0, \sigma_j^2, \tau\} \mid \sigma_j^2, \tau, \alpha, \beta\} \\ &= \frac{1}{\tau}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})\mathrm{E}\{G_0(d\mathrm{x}) \mid \alpha, \beta\} = \frac{1}{\tau\beta}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})\alpha(d\mathrm{x}).\end{aligned}$$

Within type j, the conditional covariance between $N_{\mathbf{Y}_j}(A)$ and $N_{\mathbf{Y}_j}(B)$ is

$$\begin{aligned}&\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_j}(B) \mid \sigma_j^2, \tau, \alpha, \beta\} \\ &\quad= \mathrm{E}_{G_0}\{\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_j}(B) \mid G_0, \tau, \sigma_j^2\} \mid \sigma_j^2, \tau, \alpha, \beta\} \\ &\qquad\qquad + \mathrm{Cov}_{G_0}\{\mathrm{E}\{N_{\mathbf{Y}_j}(A) \mid G_0, \tau, \sigma_j^2\}, \mathrm{E}\{N_{\mathbf{Y}_j}(B) \mid G_0, \tau, \sigma_j^2\} \mid \tau, \sigma_j^2, \alpha, \beta\} \\ &\quad= \frac{1}{\tau\beta}\int_{\mathcal{B}} K_{\sigma_j^2}(A\cap B, \mathrm{x})\alpha(d\mathrm{x}) + \frac{1+\beta}{\tau^2\beta^2}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})K_{\sigma_j^2}(B, \mathrm{x})\alpha(d\mathrm{x}).\end{aligned}$$

For $j \neq k$, the conditional covariance between $N_{\mathbf{Y}_j}(A)$ and $N_{\mathbf{Y}_k}(B)$ is

$$\begin{aligned}&\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_k}(B) \mid \sigma_j^2, \sigma_k^2, \tau, \alpha, \beta\} \\ &\quad= \mathrm{E}_{G_0}\{\mathrm{Cov}\{N_{\mathbf{Y}_j}(A), N_{\mathbf{Y}_k}(B) \mid G_0, \sigma_j^2, \sigma_k^2, \tau\} \mid \sigma_j^2, \sigma_k^2, \tau, \alpha, \beta\} \\ &\qquad + \mathrm{Cov}_{G_0}\{\mathrm{E}\{N_{\mathbf{Y}_j}(A) \mid G_0, \sigma_j^2, \tau\}, \mathrm{E}\{N_{\mathbf{Y}_k}(B) \mid G_0, \sigma_k^2, \tau\} \mid \sigma_k^2, \sigma_j^2, \tau, \alpha, \beta\} \\ &\quad= \frac{1}{\tau^2\beta^2}\int_{\mathcal{B}} K_{\sigma_j^2}(A, \mathrm{x})K_{\sigma_k^2}(B, \mathrm{x})\alpha(d\mathrm{x}).\end{aligned}$$

□

The following theorem states that our model can be approximated by the truncated model to any desired level of accuracy.

Theorem 3. *For $j = 1, \ldots, J$, for any $\epsilon > 0$ and for any measurable $A \subseteq \mathcal{B}$, there exists a natural number M_ϵ, such that*

$$E\{\Lambda_j(A) - \Lambda_j^{M_\epsilon}(A) \mid \beta, \tau\} < \epsilon,$$

where $\Lambda_j^M(A) = \sum_{m=1}^M \mu_{jm} K_{\sigma_j^2}(A, \theta_m)$ is the conditional expectation of $N_{\mathbf{Y}_j}(A)$ in the truncated model (3).

Proof. For any $M > 0$, the conditional expected truncation error given σ_j^2, τ, β, and $\{\nu_m, \theta_m\}_{m=1}^M$ is

$$
\begin{aligned}
\mathrm{E}\left[\Lambda_j(A) - \Lambda_j^M(A) \mid \sigma_j^2, \tau, \{\nu_m, \theta_m\}_{m=1}^M\right] &= \frac{1}{\tau}\sum_{m=M+1}^{\infty} \mathrm{E}\{\nu_m \mid \{\nu_m\}_{m=1}^M\} K_{\sigma_j^2}(A, \theta_m) \\
&\le \frac{1}{\tau}\sum_{m=M+1}^{\infty} \mathrm{E}\{\nu_m \mid \{\nu_m\}_{m=1}^M, \beta\} = \frac{1}{\tau}\sum_{m=1}^{\infty} \mathrm{E}\{\nu_{m+M} \mid \{\nu_m\}_{m=1}^M, \beta\} \\
&= \frac{1}{\tau\beta}\sum_{m=1}^{\infty} \mathrm{E}\left\{E_1^{-1}(\zeta_{m+M}/\alpha(\mathcal{B})) \mid \{\nu_m\}_{m=1}^M\right\} \quad (\text{let } \zeta_m' = \zeta_{m+M} - \zeta_M) \\
&= \frac{1}{\tau\beta}\sum_{m=1}^{\infty} \mathrm{E}\left\{E_1^{-1}((\zeta_m' + \zeta_M)/\alpha(\mathcal{B})) \mid \{\nu_m\}_{m=1}^M\right\} (\text{note that } \zeta_m' \sim \mathrm{Gamma}(m,1)) \\
&= \frac{1}{\tau\beta}\sum_{m=1}^{\infty} \int_0^\infty E_1^{-1}((s+\zeta_M)/\alpha(\mathcal{B}))\frac{s^{m-1}}{\Gamma(m)}\exp(-s)ds \\
&= \frac{1}{\tau\beta}\left[1 - \exp\{-E_1^{-1}(\zeta_M/\alpha(\mathcal{B}))\}\right] \le \frac{E_1^{-1}(\zeta_M/\alpha(\mathcal{B}))}{\tau\beta} \le \frac{1}{(\exp(\zeta_M/\alpha(\mathcal{B}))-1)\tau\beta}.
\end{aligned}
$$

Taking the expectation with respect to ζ_M on both sides of the above inequality results in

$$
\begin{aligned}
\mathrm{E}\left[\Lambda_j(A) - \Lambda_j^M(A) \mid \tau, \beta\right] &\le \frac{1}{\tau\beta\Gamma(M)}\int_0^\infty \frac{s^{M-1}\exp(-s)}{(\exp(s/\alpha(\mathcal{B}))-1)}ds \\
&\le \frac{\alpha(\mathcal{B})}{\tau\beta\Gamma(M)}\int_0^\infty s^{M-2}\exp(-s)ds = \frac{\alpha(\mathcal{B})}{\tau\beta(M-1)}.
\end{aligned}
$$

When $\alpha(\mathcal{B}) = 1$, we have a more accurate upper bound $\mathrm{E}\left[\Lambda_j(A) - \Lambda_j^M(A) \mid \tau, \beta\right] \le (\zeta(M) - 1)/(\tau\beta)$, where $\zeta(M) = \sum_{k=1}^\infty k^{-M}$. Therefore, for any $\epsilon > 0$, take $M_\epsilon = \lfloor a(\mathcal{B})/\tau\beta\epsilon \rfloor + 1$, where $\lfloor x \rfloor$ is the largest integer less than or equal to x.

2. Posterior Computation

In this section, we provide the details on the posterior compuation algorithm for the truncated HPGRF model

$$
\begin{aligned}
\left[(\mathbf{Y}_j, \mathbf{X}_j) \mid \{(\mu_{j,m}, \theta_m)\}_{m=1}^M, \sigma_j^2\right] &\sim \mathcal{PP}\left\{\mathcal{B}\times\mathcal{B}, K_{\sigma_j^2}(\mathrm{dy}, \mathrm{x})\sum_{m=1}^M \mu_{j,m}\delta_{\theta_m}(\mathrm{dx})\right\}, \\
[\mu_{j,m} \mid \nu_m, \tau] &\overset{\text{iid}}{\sim} \mathrm{Gamma}(\nu_m, \tau), \qquad (3) \\
\{(\theta_m, \nu_m)\}_{m=1}^M &\sim \mathrm{InvLévy}\{\alpha(d\mathrm{x}), \beta, M\},
\end{aligned}
$$

2.1. Posterior Computation.
The truncated model (3) only involves a fixed number of parameters allowing computation of the posterior.

The target distribution
Let $(\mathbf{y}_{i,j}, \mathbf{x}_{i,j})$, for $i = 1, 2, \ldots, n_j$, be multiple independent realizations (e.g. emotion studies) of the Cox process $(\mathbf{Y}_j, \mathbf{X}_j)$ in model (3), where n_j is the number of realizations of $(\mathbf{Y}_j, \mathbf{X}_j)$. For each i and j, write $(\mathbf{y}_{i,j}, \mathbf{x}_{i,j}) = \{(\mathrm{x}_{i,j,l}, \mathrm{y}_{i,j,l})\}_{l=1}^{m_{i,j}}$, where $(\mathrm{y}_{i,j,l}, \mathrm{x}_{i,j,l})$ is an observed point in $(\mathbf{y}_{i,j}, \mathbf{x}_{i,j})$ indexed by l, for $l = 1, \ldots, m_{i,j}$, and $m_{i,j}$ is the observed number of points in $(\mathbf{y}_{i,j}, \mathbf{x}_{i,j})$. The joint density of $\{\{\mathbf{x}_{i,j}\}_{i=1}^{n_j}\}_{j=1}^{J}$, $\{\boldsymbol{\mu}_j\}_{j=1}^{J}$, $\{\sigma_j^2\}_{j=1}^{J}$, $\boldsymbol{\nu}$, $\boldsymbol{\theta}$, τ and β given $\{\{\mathbf{y}_{i,j}\}_{i=1}^{n_j}\}_{j=1}^{J}$ is

$$
\begin{aligned}
&\prod_{j=1}^{J}\left[\prod_{i=1}^{n_j}\left[\pi(\mathbf{y}_{i,j}, \mathbf{x}_{i,j} \mid \boldsymbol{\mu}_j, \boldsymbol{\theta}, \sigma_j^2)\right]\pi(\sigma_j^2)\pi(\boldsymbol{\mu}_j \mid \boldsymbol{\nu}, \tau)\right]\pi(\tau)\pi(\boldsymbol{\nu}, \boldsymbol{\theta} \mid \beta)\pi(\beta) \\
&\propto \prod_{j=1}^{J}\left[\exp\left\{-\sum_{m=1}^{M} n_j K_{\sigma_j^2}(\mathcal{B}, \theta_m)\mu_{j,m}\right\}\prod_{i=1}^{n_j}\prod_{l=1}^{m_{i,j}} k_{\sigma_j^2}(\mathrm{y}_{i,j,l}, \mathrm{x}_{i,j,l})\sum_{m=1}^{M}\mu_{j,m} I_{\theta_m}(\mathrm{x}_{i,j,l})\right] \\
&\qquad \times \prod_{j=1}^{J}\left[\pi(\sigma_j^2)\prod_{m=1}^{M}\left\{\frac{\tau^{\nu_m}\mu_{j,m}^{\nu_m-1}}{\Gamma(\nu_m)}\exp\{-\tau\mu_{j,m}\}\right\}\right] \\
&\qquad \times \pi(\tau)\exp\{-E_1(\beta\nu_M)/\ell(\mathcal{B})\}\pi(\beta)\prod_{m=1}^{M}[\nu_m^{-1}\exp\{-\nu_m\beta\}],
\end{aligned}
\tag{4}
$$

where $\pi(\mathbf{y}_{i,j}, \mathbf{x}_{i,j} \mid \boldsymbol{\mu}_j, \boldsymbol{\theta}, \sigma_j^2)$ is the density of $(\mathbf{Y}_j, \mathbf{X}_j)$ with respect to a unit rate Poisson process (Møller and Waagepetersen 2004). The densities of other parameters are all with respect to a product of Lebesgue measures. In this article, we assume $\alpha(d\mathrm{x}) = \ell(d\mathrm{x})$ (i.e. Lebesgue measure) which is non-atomic; thus, the θ_m, for $m = 1, 2, \ldots, M$, are distinct with probability one. Also, $I_x(y)$ is the indicator function with $I_x(y) = 1$ if $x = y$, $I_x(y) = 0$, otherwise. Now we summarize the key steps in the posterior simulation.
Sampling $\mathbf{x}_j$: From (4), it is straightforward to obtain the conditional distribution of $\mathrm{x}_{i,j,l}$ given all other parameters:

$$
\Pr(\mathrm{x}_{i,j,l} = \theta_m \mid \cdot) \propto \mu_{j,m} k_{\sigma_j^2}(\mathrm{y}_{i,j,l}, \theta_m).
$$

Sampling $\boldsymbol{\theta}$: The full conditional density of $\boldsymbol{\theta}$ is given by

$$
\pi(\boldsymbol{\theta} \mid \cdot) \propto \exp\left\{-\sum_{m=1}^{M}\sum_{j=1}^{J} K_{\sigma_j^2}(\mathcal{B}, \theta_m)\mu_{jm} n_j\right\}\prod_{j=1}^{J}\prod_{i=1}^{n_j}\prod_{l=1}^{m_{i,j}}\left[\sum_{m=1}^{M}\mu_{j,m} I_{\mathrm{x}_{i,j,l}}(\theta_m)\right].
\tag{5}
$$

This implies that $\sum_{m=1}^{M} I_{\mathrm{x}_{i,j,l}}(\theta_m) > 0$ for all i, j and l. Let $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_{\widetilde{M}}$ be $\widetilde{M}$ distinct points in $\{\{\{\mathrm{x}_{i,j,l}\}_{l=1}^{m_{ij}}\}_{i=1}^{n_j}\}_{j=1}^{J}$. Due to the symmetry of $\{\mu_{1,m}, \ldots, \mu_{J,m}, \theta_m\}_{m=1}^{M}$ in (5) and noting that $\theta_1, \ldots, \theta_M$ are distinct points, there exists one and only one θ_m that is equal to one of $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_{\widetilde{M}}$. Thus, to sample $\boldsymbol{\theta}$, we first draw a random permutation of $\{1, \ldots, M\}$ denoted by $\{p_1, \ldots, p_M\}$. Then for $1 \le m \le \widetilde{M}$, let $\theta_{p_m} = \widetilde{\theta}_m$. For $m > \widetilde{M}$, draw θ_{p_m} according to the following density,

$$\pi(\theta_{p_m} \mid \cdot) \propto \exp\left\{-\sum_{j=1}^{J} K_{\sigma_j^2}(\mathcal{B}, \theta_{p_m})\mu_{j,p_m} n_j\right\}.$$

Note that this is a discrete distribution with normalizing constant equal to the sum of the right hand side over $m = 1, \dots, M$.

Sampling $\boldsymbol{\mu}_j$:

Proposition 1. *The full conditional distribution of $\mu_{j,m}$, for $j = 1, \dots, J$ and $m = 1, \dots, M$, is given by*

$$[\mu_{j,m} \mid \cdot] \sim \mathit{Gamma}\left[\nu_m + \sum_{i=1}^{n_j}\sum_{l=1}^{m_{i,j}} I_{\theta_m}(\mathrm{x}_{i,j,l}), n_j K_{\sigma_j^2}(\mathcal{B}, \theta_m) + \tau\right].$$

Proof. Write $a_{i,j,l} = \sum_{m=1}^{M} m I_{\theta_m}(\mathrm{x}_{i,j,l})$, then we have $I_m(a_{i,j,l}) = I_{\theta_m}(\mathrm{x}_{i,j,l})$. Note that the θ_m are distinct, $\sum_{m=1}^{M} I_{\theta_m}(\mathrm{x}_{i,j,l}) = 1 = \sum_{m=1}^{M} I_m(a_{i,j,l})$. Thus,

$$\sum_{m=1}^{M} \mu_{j,m} I_{\theta_m}(\mathrm{x}_{i,j,l}) = \sum_{m=1}^{M} \mu_{j,m} I_m(a_{i,j,l}) = \mu_{j,a_{i,j,l}}. \tag{6}$$

Furthermore,

$$\prod_{i=1}^{n_j}\prod_{l=1}^{m_{i,j}} \mu_{j,a_{i,j,l}} = \prod_{i=1}^{n_j}\prod_{l=1}^{m_{i,j}} \mu_{j,a_{i,j,l}}^{\sum_{m=1}^{M} I_m(a_{i,j,l})} = \prod_{i=1}^{n_j}\prod_{l=1}^{m_{i,j}}\prod_{m=1}^{M} \mu_{j,a_{i,j,l}}^{I_m(a_{i,j,l})} = \prod_{m=1}^{M} \mu_{j,m}^{b_{j,m}}, \tag{7}$$

where $b_{j,m} = \sum_{i=1}^{n_j}\sum_{l=1}^{m_{i,j}} I_{\theta_m}(\mathrm{x}_{i,j,l})$. From the joint posterior distribution of all the parameters, (6) and (7), we have $\pi(\mu_{j,m} \mid \cdot) \propto \mu_{j,m}^{\nu_m + b_{j,m} - 1} \exp\{-(n_j K_{\sigma_j}(\mathcal{B}, \mathrm{x}) + \theta_m)\mu_{j,m}\}$. □

Sampling $\boldsymbol{\nu}$: The full conditional distribution of ν_m is

$$\pi(\nu_m \mid \cdot) \propto \begin{cases} \frac{c_m^{\nu_m}}{\Gamma(\nu_m)} \nu_m^{-1}, & m = 1, \dots, M-1 \\ \exp\{-E_1(\beta\nu_M)/\ell(\mathcal{B})\} \frac{c_M^{\nu_M}}{\Gamma(\nu_M)} \nu_M^{-1}, & m = M \end{cases}$$

where $c_m = \tau^J \prod_{j=1}^{J} \mu_{j,m} e^{-\beta}$. We use a symmetric random walk to update v_m: sample $v^* \sim N(v_m, \sigma_v^2)$ and accept v^* with probability $\min\{1, \pi(v^* \mid \cdot)/\pi(v_m \mid \cdot)\}$.

2.2. **Sampling Hyperparameters.** We update σ_j^2, for $j = 1, \dots, J$, τ using Metropolis within Gibbs sampling (Smith and Roberts 1993). In this article, we choose $k_{\sigma_j^2}(\mathrm{y}, \mathrm{x}) = \left(2\pi\sigma_j^2\right)^{-d/2} \exp\{-\|\mathrm{y} - \mathrm{x}\|^2/(2\sigma_j^2)\}$ (an isotropic Gaussian density) and assume, a priori, $\sigma_j^{-2} \sim \text{Uniform}[a_\sigma, b_\sigma]$, $\tau \sim \text{Gamma}(a_\tau, b_\tau)$ and $\beta \sim \text{Gamma}(a_\beta, b_\beta)$ (hyperprior parameters will be specified in the next section).

The full conditional of σ_j^2 is

$$\pi(\sigma_j^2 \mid \cdot) \\ \propto \exp\left[-\sum_{m=1}^{M}\left\{K_{\sigma_j^2}(\mathcal{B},\theta_m)\mu_{j,m}n_j\right\} - \frac{S_{\cdot,j}^2}{\sigma_j^2} - \frac{m_{\cdot,j}d}{2}\log(\sigma_j^2)\right] I_{[a_\sigma,b_\sigma]}(\sigma_j^{-2}),$$

where $S_{\cdot,j}^2 = \frac{1}{2}\sum_{i=1}^{n_j}\sum_{l=1}^{m_{i,j}}\|\mathrm{y}_{i,j,l} - \mathrm{x}_{i,j,l}\|^2$ and $m_{\cdot,j} = \sum_{j=1}^{n_j} m_{i,j}$. Thus, to update σ_j^2, we use a random walk and first draw $\sigma_j^{2*} \sim N(\sigma_j^2, \theta_\sigma^2)$, if $\sigma_j^{2*} \in (a_\sigma, b_\sigma)$, then set $\sigma_j^2 = \sigma_j^{2*}$ with probability $\min\{r(\sigma), 1\}$, where $r(\sigma)$ is

$$\exp\left[\sum_{m=1}^{M}\left\{\left[K_{\sigma_j^2}(\mathcal{B},\theta_m) - K_{\sigma_j^{2*}}(\mathcal{B},\theta_m)\right]\mu_{j,m}n_j\right\} + S_{\cdot,j}^2\left(\frac{1}{\sigma_j^2} - \frac{1}{\sigma_j^{2*}}\right)\right]\left(\frac{\sigma_j^2}{\sigma_j^{2*}}\right)^{\frac{1}{2}m_{\cdot,j}d}.$$

The full conditional of τ is

$$\pi(\tau \mid \cdot) \propto \tau^{J\sum_{m=1}^{M}\nu_m + a_\tau - 1}\exp\left\{-\left(b_\tau + \sum_{j=1}^{J}\sum_{m=1}^{M}\mu_{j,m}\right)\tau\right\},$$

which implies that we can update τ by drawing

$$[\tau \mid \cdot] \sim \text{Gamma}\left(J\sum_{m=1}^{M}\nu_m + a_\tau - 1, \sum_{j=1}^{J}\sum_{m=1}^{M}\mu_{j,m} + b_\tau\right).$$

The full conditional of β is

$$\pi(\beta \mid \cdot) \propto \beta^{b_\beta - 1}\exp\left\{-\left[a_\beta + \sum_{i=1}^{M}\nu_m\right]\beta - E_1(\beta\nu_M)/\ell(\mathcal{B})\right\}.$$

We update β using a symmetric random walk by sampling $\beta^* \sim N(\beta, \sigma_\beta^2)$ and accepting it with probability $\min\{1, \pi(\beta^* \mid \cdot)/\pi(\beta \mid \cdot)\}$.

3. Sensitivity Analysis

In this section, we conduct simulation studies to study how the posterior inference on the intensity function varies with different prior specifications of σ_j^2, β and τ and truncation approximations M. Specifically, we consider nine scenarios in Table 1. We simulate the posterior distribution with 20,000 iterations after a burn-in of 2,000 iterations. Table 2 shows that summary statistics of posterior mean intensity estimates over the whole brain regions for nine different scenarios. Theses summary statistics are quite similar and show that posterior results are not very sensitive to prior specification.

TABLE 1. Prior specifications and truncation approximations of nine scenarios for sensitivity analysis

Scenarios	σ_j^{-2}	β	τ	M
1	$U(0,10)$	$G(0.1,0.1)$	$G(0.1,0.1)$	$10,000$
2	$U(0,10)$	$G(0.2,0.2)$	$G(0.2,0.2)$	$10,000$
3	$U(0,10)$	$G(0.1,0.1)$	$G(0.1,0.1)$	$5,000$
4	$U(0,10)$	$G(0.2,0.2)$	$G(0.2,0.2)$	$5,000$
5	$U(0,10)$	$G(2.0,2.0)$	$G(2.0,2.0)$	$10,000$
6	$U(0,10)$	$G(2.0,2.0)$	$G(2.0,2.0)$	$12,500$
7	$G(2.0,2.0)$	$G(2.0,2.0)$	$G(2.0,2.0)$	$10,000$
8	$G(1.0,1.0)$	$G(2.0,2.0)$	$G(2.0,2.0)$	$10,000$
9	$G(3.0,3.0)$	$G(2.0,2.0)$	$G(2.0,2.0)$	$10,000$

TABLE 2. Summary statistics of posterior mean intensity estimates over the whole brain regions for nine different scenarios

		Scenarios								
Emotions	Stats	1	2	3	4	5	6	7	8	9
	Min.	9.1e-18	2.8e-17	3.6e-20	1.2e-17	3.8e-20	4.0e-16	1.5e-16	7.4e-17	3.0e-17
Sad	Med.	1.5e-07	1.4e-07	1.2e-07	1.6e-07	1.3e-07	1.4e-07	1.5e-07	1.7e-07	1.8e-07
	Max.	1.7e-03	1.6e-03	1.6e-03	1.6e-03	1.7e-03	1.7e-03	1.6e-03	1.6e-03	1.6e-03
	Min.	1.7e-17	1.6e-17	3.9e-23	5.8e-19	6.5e-24	6.7e-17	3.8e-18	1.4e-17	4.7e-18
Happy	Med.	3.1e-08	3.2e-08	1.2e-08	1.7e-08	8.0e-09	3.4e-08	2.2e-08	2.0e-08	2.0e-08
	Max.	1.5e-03	1.4e-03	1.6e-03	1.6e-03	1.5e-03	1.6e-03	1.6e-03	1.5e-03	1.5e-03
	Min.	6.2e-19	1.7e-17	3.9e-22	1.6e-18	6.5e-24	3.9e-17	1.3e-18	7.3e-19	6.4e-19
Anger	Med.	2.8e-08	3.2e-08	2.5e-08	2.7e-08	2.4e-08	3.3e-08	2.0e-08	3.4e-08	2.1e-08
	Max.	1.8e-03	1.9e-03	1.6e-03	1.8e-03	1.7e-03	1.9e-03	1.8e-03	1.7e-03	1.9e-03
	Min.	7.8e-21	1.6e-17	2.2e-21	3.1e-17	1.7e-22	2.3e-17	5.2e-18	1.2e-17	3.5e-18
Fear	Med.	6.4e-08	8.0e-08	6.1e-08	6.2e-08	6.3e-08	7.8e-08	6.9e-08	9.0e-08	6.2e-08
	Max.	1.5e-03	1.5e-03	1.5e-03	1.5e-03	1.5e-03	1.4e-03	1.4e-03	1.5e-03	1.4e-03
	Min.	3.7e-19	2.9e-19	4.1e-22	1.2e-18	2.1e-22	3.0e-17	2.6e-18	2.9e-18	3.4e-19
Disgust	Med.	5.6e-08	5.3e-08	5.8e-08	4.8e-08	4.2e-08	4.2e-08	5.2e-08	5.6e-08	5.9e-08
	Max.	2.0e-03	2.0e-03	2.0e-03	2.0e-03	2.1e-03	2.1e-03	2.1e-03	2.0e-03	2.1e-03
	Min.	8.3e-07	6.8e-07	7.7e-08	5.9e-07	4.8e-08	9.0e-07	6.8e-07	6.0e-07	5.8e-07
Pop. Mean	Med.	1.6e-05	1.5e-05	1.1e-05	1.4e-05	1.1e-05	1.4e-05	1.4e-05	1.2e-05	1.3e-05
	Max.	3.2e-04	2.8e-04	4.6e-04	3.0e-04	6.7e-04	2.2e-04	2.8e-04	2.6e-04	2.5e-04

4. Bayesian Spatial Point Process Classifier

In this section, we present the important sampling approach to sampling T_{n+1}. We have the follow proposition:

Proposition 2. *The posterior predictive distribution for T_{n+1} is given by*

$$\Pr[T_{n+1} = j \mid \mathbf{x}_{n+1}, \mathcal{D}_n] = \frac{p_j \int \pi(\mathbf{x}_{n+1} \mid T_{n+1} = j, \Theta)\pi(\Theta \mid \mathcal{D}_n)d\Theta}{\sum_{j'=1}^{J} p_{j'} \int \pi(\mathbf{x}_{n+1} \mid T_{n+1} = j', \Theta)\pi(\Theta \mid \mathcal{D}_n)d\Theta} \tag{8}$$

where

$$\pi(\mathbf{x}_{n+1} \mid T_{n+1} = j, \Theta) \quad = \quad \exp\left\{|\mathcal{B}| - \int_{\mathcal{B}} \lambda_j(x \mid \Theta)dx\right\} \prod_{x \in \mathbf{x}_{n+1}} \lambda_j(x \mid \Theta).$$

Proof. First, we notice that

$$\begin{aligned}
\pi(\Theta \mid \mathbf{x}_{n+1}, \mathcal{D}_n) &= \frac{\pi(\Theta \mid \mathbf{x}_{n+1}, \mathcal{D}_n)}{\pi(\Theta \mid \mathcal{D}_n)} \pi(\Theta \mid \mathcal{D}_n) \\
&= \frac{\pi(\mathbf{x}_{n+1}, \mathcal{D}_n \mid \Theta)\pi(\Theta)}{\pi(\mathbf{x}_{n+1}, \mathcal{D}_n)} \cdot \frac{\pi(\mathcal{D}_n)}{\pi(\mathcal{D}_n \mid \Theta)\pi(\Theta)} \cdot \pi(\Theta \mid \mathcal{D}_n) \\
&= \frac{\pi(\mathbf{x}_{n+1} \mid \Theta)\pi(\mathcal{D}_n \mid \Theta)\pi(\Theta)}{\pi(\mathbf{x}_{n+1}, \mathcal{D}_n)} \cdot \frac{\pi(\mathcal{D}_n)}{\pi(\mathcal{D}_n \mid \Theta)\pi(\Theta)} \cdot \pi(\Theta \mid \mathcal{D}_n) \\
&= \frac{\pi(\mathbf{x}_{n+1} \mid \Theta)}{\pi(\mathbf{x}_{n+1} \mid \mathcal{D}_n)} \pi(\Theta \mid \mathcal{D}_n) = \frac{\pi(\mathbf{x}_{n+1} \mid \Theta)\pi(\Theta \mid \mathcal{D}_n)}{\int \pi(\mathbf{x}_{n+1} \mid \Theta)\pi(\Theta \mid \mathcal{D}_n)d\Theta}. \qquad (9)
\end{aligned}$$

Thus, the probability becomes

$$\Pr[T_{n+1} = j \mid \mathbf{x}_{n+1}, \mathcal{D}_n] = \frac{\int p_j \pi(\mathbf{x}_{n+1} \mid T_{n+1} = j, \Theta)\pi(\Theta \mid \mathcal{D}_n)d\Theta}{\int \sum_{j'=1}^{J} p_{j'} \pi(\mathbf{x}_{n+1} \mid T_{n+1} = j', \Theta)\pi(\Theta \mid \mathcal{D}_n)d\Theta}. \tag{10}$$

□

This proposition leads to the following algorithm used to estimate the posterior predictive probability used for reverse inference.
Reverse inference algorithm.

- **Input**: The observed data $\mathcal{D}_n$, a foci pattern, $\mathbf{x}_{n+1}$, reported from a new study and the total number of simulations K.
- **Step 1**: Run a Bayesian spatial point process model to obtain the posterior draws $\Theta^{(k)} \sim \pi(\Theta \mid \mathcal{D}_n)$, for $k = 1, \ldots, K$.
- **Step 2**: Compute

$$\pi_j^{(k)} = \exp\left\{-\int_{\mathcal{B}} \lambda_j(x \mid \Theta^{(k)})dx\right\} \left\{\prod_{x \in \mathbf{x}_{n+1}} \lambda_j(x \mid \Theta^{(k)})\right\}.$$

- **Output**: The posterior predictive probability is given by

$$\widehat{\Pr}[T_{n+1} = j \mid \mathbf{x}_{n+1}, \mathcal{D}_n] = \frac{p_j \sum_{k=1}^{K} \pi_j^{(k)}}{\sum_{j'=1}^{J} p_{j'} \sum_{k=1}^{K} \pi_{j'}^{(k)}}. \tag{11}$$

The prediction of T_{n+1} is given by

$$\hat{T}_{n+1} = \arg\max_j \left(p_j \sum_{k=1}^{K} \pi_j^{(k)} \right). \tag{12}$$

To evaluate the performance of our method, we are interested in computing the leave-one-out cross-validation (LOOCV) classification rates. More specifically, we use data $\mathcal{D}_{-i} = \{(\mathbf{x}_l, t_l)\}_{l \neq i}$ to make a predication of T_i denoted by $\hat{T}_i$, and focus on the $J \times J$ LOOCV confusion matrix $\boldsymbol{C} = \{c_{jj'}\}$, defined by

$$c_{jj'} = \frac{\sum_{i=1}^{n} I_j(t_i) I_{j'}(\hat{T}_i)}{\sum_{i=1}^{n} I_j(t_i)}, \tag{13}$$

where $I_a(b)$ is an indicator function. $I_a(b) = 1$ if $a = b$, $I_a(b) = 0$, otherwise. Then the overall and the average correct classification rates are respectively given by

$$c_o \;=\; \frac{1}{n}\sum_{i=1}^{n} I_{t_i}(\hat{T}_i), \quad \text{and} \quad c_a = \frac{1}{n}\sum_{i=1}^{n} c_{jj}. \tag{14}$$

In order to obtain $\hat{T}_i$, we compute the LOOCV predictive probabilities. For $i = 1, \ldots, n$,

$$\Pr[T_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}] = \int \Pr[T_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}, \Theta] \pi(\Theta \mid \mathbf{x}_i, \mathcal{D}_{-i}) d\Theta, \tag{15}$$

which can be estimated via Monte Carlo simulation. However, it is not straightforward and very inefficient to draw Θ from $\pi(\Theta \mid \mathbf{x}_i, \mathcal{D}_{-i})$ for each i. Thus, to avoid having to run the multiple posterior simulations, we consider another representation of (15) in the following proposition.

Proposition 3. *The LOOCV predictive probabilities of T_i, for $i = 1, \ldots, n$, is given by*

$$\Pr[T_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}] = \frac{p_j Q_{jt_i}}{\sum_{j'=1}^{J} p_{j'} Q_{j't_i}}, \tag{16}$$

where

$$Q_{jj'} = \int \frac{\pi(\mathbf{x}_i \mid T_i = j, \Theta)}{\pi(\mathbf{x}_i \mid T_i = j', \Theta)} \pi(\Theta \mid \mathcal{D}_n) d\Theta.$$

Proof. The LOOCV posterior predictive probability is

$$\Pr[T_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}] = \int \Pr[T_i = j \mid \mathbf{x}_i, \Theta] \pi(\Theta \mid \mathbf{x}_i, \mathcal{D}_{-i}) d\Theta$$

$$= \int \Pr[T_i = j \mid \mathbf{x}_i, \Theta] \frac{\pi(\Theta \mid \mathbf{x}_i, \mathcal{D}_{-i})}{\pi(\Theta \mid \mathbf{x}_i, T_i = t_i, \mathcal{D}_{-i})} \pi(\Theta \mid \mathbf{x}_i, T_i = t_i, \mathcal{D}_{-i}) d\Theta.$$

Note that

$$\frac{\pi(\Theta \mid \mathbf{x}_i, \mathcal{D}_{-i})}{\pi(\Theta \mid \mathbf{x}_i, T_i = t_i, \mathcal{D}_{-i})} = \frac{\pi(\Theta, \mathbf{x}_i, \mathcal{D}_{-i}) \pi(\mathbf{x}_i, T_i = t_i, \mathbf{x}_{-i}, t_{-i})}{\pi(\mathbf{x}_i, \mathcal{D}_{-i}) \cdot \pi(\Theta, \mathbf{x}_i, T_i = t_i, \mathcal{D}_{-i})} = \frac{\Pr[T_i = t_i \mid \mathbf{x}_i, \mathcal{D}_{-i}]}{\Pr[T_i = t_i \mid \mathbf{x}_i, \Theta]}.$$

This implies that

$$\frac{\Pr[t_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}]}{\Pr[T_i = t_i \mid \mathbf{x}_i, \mathcal{D}_{-i}]} = \int \frac{\Pr[T_i = j \mid \mathbf{x}_i, \Theta]}{\Pr[T_i = t_i \mid \mathbf{x}_i, \Theta]} \pi(\Theta \mid \mathbf{x}_i, T_i = t_i, \mathcal{D}_{-i}) d\Theta$$

$$= \int \frac{\pi[\mathbf{x}_i \mid T_i = j, \Theta] p_j}{\pi[\mathbf{x}_i \mid T_i = t_i, \Theta] p_{t_i}} \pi(\Theta \mid \mathbf{x}_i, T_i = t_i, \mathcal{D}_{-i}) d\Theta := Q_{j,t_i}.$$

By $\sum_{j=1}^{J} \Pr[T_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}] = 1$, we have that

$$\frac{1 - \Pr[T_i = t_i \mid \mathbf{x}_i, \mathcal{D}_{-i}]}{\Pr[T_i = t_i \mid \mathbf{x}_i, \mathcal{D}_{-i}]} = \sum_{j \neq t_i} Q_{j,t_i}.$$

This implies

$$\Pr[T_i = t_i \mid \mathbf{x}_i, \mathcal{D}_{-i}] = \frac{1}{1 + \sum_{j \neq t_i} Q_{j,t_i}}, \quad \Pr[T_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}] = \frac{Q_{j,t_i}}{1 + \sum_{j \neq t_i} Q_{j,t_i}}.$$

□

This proposition leads to the following algorithm.
LOOCV algorithm.

- **Input**: The observed data $\mathcal{D}_n$ and the total number of simulations K.
- **Step 1**: Run a Bayesian spatial point process model to obtain the posterior draws $\Theta^{(k)} \sim \pi(\Theta \mid \mathcal{D}_n)$, for $k = 1, \ldots, K$.
- **Step 2**: For $i = 1, \ldots, n$ and $j = 1, \ldots, J$, compute

$$\widehat{Q}_{jt_i} = \frac{1}{K} \sum_{k=1}^{K} \frac{\pi(\mathbf{x}_i \mid T_i = j, \Theta_i^{(k)})}{\pi(\mathbf{x}_i \mid T_i = t_i, \Theta_i^{(k)})}. \tag{17}$$

- **Output**: The posterior of the predictive probabilities of T_i, for $i = 1, \ldots, n$, are given by

$$\widehat{\Pr}[T_i = j \mid \mathbf{x}_i, \mathcal{D}_{-i}] = \frac{p_j \hat{Q}_{jt_i}}{\sum_{j'=1}^{J} p_{j'} \hat{Q}_{j't_i}}. \tag{18}$$

And the estimate of T_i is

$$\hat{T}_i = \arg\max_j \left(p_j Q_{jt_i}\right). \tag{19}$$

References

Møller, J. and Waagepetersen, R. P. (2004), *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC.

Smith, A. and Roberts, G. (1993), "Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society. Series B (Methodological)*, 3–23.

Wolpert, R. L. and Ickstadt, K. (1998), "Poisson/gamma random field models for spatial statistics," *Biometrika*, 85, 251–267.