SpGPCW: Spatially Varying Gaussian Process Model for Critical Window Estimation

Statistical Model

$$Y_{ij}|\boldsymbol{\beta}, \boldsymbol{\theta}_i \stackrel{\text{ind}}{\sim} \text{Binomial}\left\{c_{ij}, p_{ij}\left(\boldsymbol{\beta}, \boldsymbol{\theta}_i\right)\right\},\ i = 1, ..., s;\ j = 1, ..., n_i;$$

$$\log\left\{\frac{p_{ij}\left(\boldsymbol{\beta}, \boldsymbol{\theta}_i\right)}{1 - p_{ij}\left(\boldsymbol{\beta}, \boldsymbol{\theta}_i\right)}\right\} = \mathbf{x}_{ij}^{\text{T}}\boldsymbol{\beta} + \sum_{k=1}^{m_{ij}} \text{z}_{ijk}\theta_i\left(k\right);$$

$$\boldsymbol{\theta} = \left(\boldsymbol{\theta}_1^{\text{T}}, ..., \boldsymbol{\theta}_s^{\text{T}}\right)^{\text{T}};\ \boldsymbol{\theta}_i = \left\{\theta_i\left(1\right), ..., \theta_i\left(m\right)\right\}^{\text{T}};$$

$$\boldsymbol{\theta}_i|\boldsymbol{\theta}_{-i}, \rho, \phi, \sigma_{\theta}^2, \boldsymbol{\eta}, \stackrel{\text{ind}}{\sim} \text{MVN}\left\{\frac{\rho\sum_{j=1}^s w_{ij}\boldsymbol{\theta}_j + \left(1-\rho\right)\boldsymbol{\eta}}{\rho\sum_{j=1}^s w_{ij} + \left(1-\rho\right)}, \frac{\sigma_{\theta}^2\Sigma\left(\phi\right)}{\rho\sum_{j=1}^s w_{ij} + \left(1-\rho\right)}\right\},\ i = 1, ..., s;$$

$$\boldsymbol{\eta} = \left\{\eta\left(1\right), ..., \eta\left(m\right)\right\}^{\text{T}}|\sigma_{\eta}^2, \phi \sim \text{MVN}\left\{\mathbf{0}_m, \sigma_{\eta}^2\Sigma\left(\phi\right)\right\};$$

$$\Sigma\left(\phi\right)_{ij} = \exp\left\{-\phi|i-j|\right\},\ \phi > 0;$$

- $m = \max\left\{m_{ij} : i = 1, ..., s;\ j = 1, ..., n_i\right\};$
- s: Number of unique spatial locations;
- n_i: Number of observations with spatial location i;
- $\mathbf{0}_m$: Length m vector with each entry equal to zero.

Prior Information

$\beta_j \stackrel{\text{iid}}{\sim} \text{N}\left(0, \sigma_{\beta}^2\right),\ j = 1, ..., p;$

- p: Length of $\mathbf{x}_{ij}$ vector (same for all i, j);
- Default setting: $\sigma_{\beta}^2 = 10,000$.

$\sigma_{\theta}^2 \sim \text{Inverse Gamma}\left(a_{\sigma_{\theta}^2}, b_{\sigma_{\theta}^2}\right);$

- Default setting: $a_{\sigma_{\theta}^2} = 3$, $b_{\sigma_{\theta}^2} = 2$.

$\sigma_{\eta}^2 \sim \text{Inverse Gamma}\left(a_{\sigma_{\eta}^2}, b_{\sigma_{\eta}^2}\right);$

- Default setting: $a_{\sigma_{\eta}^2} = 3$, $b_{\sigma_{\eta}^2} = 2$.

$\rho \sim \text{Uniform}\left(a_{\phi}, b_{\phi}\right);$

- Default setting: $a_{\phi} = 0$, $b_{\phi} = 1$.

$\phi \sim \text{Uniform}\left(a_{\phi}, b_{\phi}\right);$

- Default setting: $a_{\phi} = \log(0.9999)/(-(m-1))$, $b_{\phi} = \log(0.0001)/(-1)$.

Default Initial Values

- $\beta_j = 0$ for all j;
- $\theta_j = 0$ for all j;
- $\eta_j = 0$ for all j;
- $\sigma_\theta^2 = 1.00$;
- $\sigma_\eta^2 = 1.00$;
- $\rho = 0.50 * (b_\phi - a_\phi)$;
- $\phi = 0.01\,(b_\phi - a_\phi)$.

Alternate Likelihood: Gaussian

$Y_{ij}|\boldsymbol{\beta}, \boldsymbol{\theta}_i, \sigma_\epsilon^2 \overset{\text{ind}}{\sim} \text{Normal}\left(\mathbf{x}_{ij}^{\text{T}}\boldsymbol{\beta} + \sum_{k=1}^{m_{ij}} \text{z}_{ijk}\theta_i\left(k\right), \sigma_\epsilon^2\right),\ i = 1, ..., s;\ j = 1, ..., n_i;$

- $\sigma_\epsilon^2 \sim \text{Inverse Gamma}\left(a_{\sigma_\epsilon^2}, b_{\sigma_\epsilon^2}\right);$
- Default setting: $a_{\sigma_\epsilon^2} = 0.01$, $b_{\sigma_\epsilon^2} = 0.01$;
- Default initial value: $\sigma_\epsilon^2 = 1.00$.

Alternate Likelihood: Negative Binomial

$Y_{ij}|\boldsymbol{\beta}, \boldsymbol{\theta}, r \overset{\text{ind}}{\sim} \text{Negative Binomial}\left\{r, \lambda_{ij}\left(\boldsymbol{\beta}, \boldsymbol{\theta}_i\right)\right\},\ i = 1, ..., s;\ j = 1, ..., n_i;$

$\ln\left\{\frac{\lambda_{ij}(\boldsymbol{\beta}, \boldsymbol{\theta}_i)}{1-\lambda_{ij}(\boldsymbol{\beta}, \boldsymbol{\theta}_i)}\right\} = \text{O}_{ij} + \mathbf{x}_{ij}^{\text{T}}\boldsymbol{\beta} + \sum_{k=1}^{m_{ij}} \text{z}_{ijk}\theta_i\left(k\right).$

- $r \sim \text{Discrete Uniform}\left[a_r, b_r\right];$
- Default setting: $a_r = 1$, $b_r = 100$;
- Default initial value: $r = b_r$.

Likelihood Indicator

- likelihood_indicator = 0: Binomial;
- likelihood_indicator = 1: Gaussian;
- likelihood_indicator = 2: Negative binomial.

Special Notes

- Do not center the exposures. Instead, we recommend that they are only scaled at each outcome time period (e.g., by interquartile range).