Statistica Sinica: Supplement

Bayesian Spatial-Temporal Modeling of Ecological Zero-Inflated Count Data

Xia Wang[1], Ming-Hui Chen[2], Rita C. Kuo[3], and Dipak K. Dey[2]

[1] *University of Cincinnati* [2] *University of Connecticut*
[3] *Lawrence Berkeley National Laboratory*

Supplementary Material

S1 The MCMC Sampling Algorithm for the Proposed Spatial-Temporal Model

Gibbs sampling on the latent spatial random fields Z

The conditional distribution of $Z_{t,i}$ in year t depends on the observed counts $y_{t,i}$ and all the other $Z_{(t,i)} = \{Z_{t,\ell}, \ell \neq i\}$. Assume independence between Z_t and Z_s, $t, s = 1, \ldots, T$, $t \neq s$. If $y_{t,i} > 0$, $E_{t,i} = 1$ and $Z_{t,i} > 0$ with probability 1. The full conditional distribution of $Z_{t,i}$ is a truncated Gaussian given by

$$p(Z_{t,i}|Z_{(t,i)}, y_{t,i} > 0, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\xi}_t(t = 1, \ldots, T), \rho, \phi, \phi_1)$$
$$\propto (2\pi)^{-1/2} \exp\{-(Z_{t,i} - \mu^*_{t,i})^2/2\} I(Z_{t,i} > 0),$$

where $\mu^*_{t,i} = \mathbf{x}_i'\boldsymbol{\beta} + \phi \sum_{l \neq i} \mathbf{w}_{il}(Z_{t,l} - \mathbf{x}_l'\boldsymbol{\beta})$. When $y_{t,i} = 0$, the full conditional distribution of $Z_{t,i}$ is

$$p(Z_{t,i}|\cdot) \propto (2\pi)^{-1/2} \exp\{-(Z_{t,i} - \mu^*_{t,i})^2/2\}$$
$$\times \{I(Z_{t,i} \leq 0) + I(Z_{t,i} > 0) \exp[-\exp\{\eta_{t,i}\}]\}^{I_{t,i}}, \quad \text{(S1.1)}$$

where the indicator function $I_{t,i} = 1$ if grid i is surveyed in year t and 0 otherwise.

When $I_{t,i} = 1$, the technique to sample from the full conditional distribution in (S1.1) is described in Fei and Rathbun (2006). Let $q_{t,i} = \Phi(\mu^*_{t,i})$, where $\Phi(\cdot)$ is the cumulative standard normal distribution function. With the probability

$$\frac{q_{t,i}\{\exp[-\exp(\eta_{t,i})]\}}{q_{t,i}\{\exp[-\exp(\eta_{t,i})]\} + (1 - q_{t,i})},$$

we sample from $N(\mu^*_{t,i}, 1)$ truncated to the left at zero, and otherwise we sample from $N(\mu^*_{t,i}, 1)$ truncated to the right at zero. When $I_{t,i} = 0$, $Z_{t,i}$ is sampled from $N(\mu^*_{t,i}, 1)$ without truncation.

For the grid with $\sum_{t=1}^{T} I_{t,i} = 0$, there is no contribution to the likelihood by this grid, while the statistical inference can still be drawn for these grids based on their neighborhood.

Gibbs sampling on $\boldsymbol{\beta}$

The full conditional distribution of $\boldsymbol{\beta}$ is of the form

$$p(\boldsymbol{\beta}|\mathbf{X}, \mathbf{Z}, \phi) \propto \exp[-\{\sum_{t=1}^{T}(\mathbf{Z}_t - \mathbf{X}'\boldsymbol{\beta})'\Sigma(\phi)^{-1}(\mathbf{Z}_t - \mathbf{X}'\boldsymbol{\beta}) + \boldsymbol{\beta}' g_{\boldsymbol{\beta}}^{-1}(\mathbf{X}'\mathbf{X})\boldsymbol{\beta}\}/2],$$

which is a multivariate normal distribution $N(\tilde{\mu}_{\boldsymbol{\beta}}, \tilde{\mathbf{V}}_{\boldsymbol{\beta}})$ with $\tilde{\mathbf{V}}_{\boldsymbol{\beta}} = \{T\mathbf{X}'\Sigma(\phi)^{-1}\mathbf{X} + g_{\boldsymbol{\beta}}^{-1}(\mathbf{X}'\mathbf{X})\}^{-1}$ and $\tilde{\mu}_{\boldsymbol{\beta}} = \tilde{\mathbf{V}}_{\boldsymbol{\beta}}(\sum_{t=1}^{T}\mathbf{X}'\Sigma(\phi)^{-1}\mathbf{Z}_t)$.

Metropolis-Hastings sampling on ϕ

The full conditional distribution of ϕ is given by

$$p(\phi|\cdot) \propto |\Sigma(\phi)|^{-T/2}\exp\Big\{-\frac{1}{2}\sum_{t=1}^{T}(\mathbf{Z}_t - \mathbf{X}'\boldsymbol{\beta})'\Sigma^{-1}(\phi)(\mathbf{Z}_t - \mathbf{X}'\boldsymbol{\beta})\Big\}\cdot\pi(\phi).$$

To facilitate sampling ϕ from the given boundary of the uniform prior, let $\eta = \text{logit}\{(\phi - \phi_{\min})/(\phi_{\max} - \phi_{\min})\}$ and let $\eta_{m+1} \sim N(\eta_m, \phi_\eta^2)$, where η_m is the value of η at the m^{th} iteration and ϕ_η^2 is the tuning constant to achieve the desired acceptance rate. Then $\phi_{m+1} = \{\exp(\eta_{m+1})\phi_{\max} + \phi_{min}\}/\{1 + \exp(\eta_{m+1})\}$. The probability of updating ϕ is decided by the minimum of 0 and $\log\{p(\phi_{m+1}|\cdot)(\phi_{m+1} - \phi_{\min})(\phi_{\max} - \phi_{m+1})\} - \log\{p(\phi_m|\cdot)(\phi_m - \phi_{\min})(\phi_{\max} - \phi_m)\}$, where $\phi_{\min} = -0.2517119$ and $\phi_{\max} = 0.1257267$ as described in Section 3.

Metropolis-Hastings or adaptive rejection sampling on $\boldsymbol{\alpha}$

The full conditional distribution of $\boldsymbol{\alpha}$ takes the form

$$\begin{aligned} p(\boldsymbol{\alpha}|\cdot) \propto \exp\big[&\sum_{t=1}^{T}\sum_{i=1}^{N}\{(\tilde{\mathbf{x}}_{t,i}'\boldsymbol{\alpha} + \boldsymbol{\Psi}_{t,i}\boldsymbol{\xi}_t)y_{t,i} - \exp(\tilde{\mathbf{x}}_{t,i}'\boldsymbol{\alpha} + \boldsymbol{\Psi}_{t,i}\boldsymbol{\xi}_t)\}I(Z_{t,i} > 0) \\ &- 0.5 \times \boldsymbol{\alpha}'\Sigma_{\boldsymbol{\alpha}}^{-1}\boldsymbol{\alpha}\big]. \end{aligned} \tag{S1.2}$$

To accelerate convergence, each element of $\boldsymbol{\alpha}$ is sampled separately. A new value for the p^{th} element in $\boldsymbol{\alpha}$, $\alpha_{p,m+1}^*$, is sampled from $N(\alpha_{p,m}, \phi_p^2)$, where $\alpha_{p,m}$ is the value of α_p at the m^{th} iteration and ϕ_p^2 is the tuning parameter which is selected to yield an empirical acceptance rate around 0.25, as used by Fei and Rathbun (2006). Each element is updated based on the fraction $\min[\{p(\alpha_{p,m+1}|\cdot)\}/\{p(\alpha_{p,m}|\cdot)\}, 1]$.

Once spatial correlation in the count part is introduced into the model, the convergence is very slow using the above Metropolis-Hastings sample. Since the full conditional density in

(S1.2) is log-concave in each component of $\boldsymbol{\alpha}$, we use the adaptive rejection sampling algorithm of Gilks and Wild (1992).

Adaptive rejection sampling on $\boldsymbol{\xi}_t$ $(t = 1, \ldots, T)$

The posterior densities of $\xi_t, t = 1, \ldots, T$, are log-concave in each component of $\boldsymbol{\xi}$. Therefore, we can use the adaptive rejection sampling algorithm of Gilks and Wild (1992) to sample from the full conditional density of $\boldsymbol{\xi}$. The logarithm of the full conditional density $h(\xi_{t,j})$ and its first derivative $h'(\xi_{t,j}), t = 1, \ldots, T, j = 1, \ldots, M$ are given as follows:

$$h(\xi_{t,j}) \propto$$
$$\begin{cases} \sum_{i=1}^{N} \left[I(Z_{1,i} > 0) \{ \eta_{1,i}\, y_{1,i} - \exp(\eta_{1,i}) \} \right] - \frac{(\xi_{2,j} - \rho\xi_{1,j})^2}{2\tau^2\lambda_{j,j}} - \frac{\xi_{1,j}^2}{2 \cdot \sigma_\xi^2}, & t = 1 \\ \sum_{i=1}^{N} \left[I(Z_{t,i} > 0) \{ \eta_{t,i}\, y_{t,i} - \exp(\eta_{t,i}) \} \right] - \frac{(\xi_{t,j} - \rho\xi_{t-1,j})^2 + (\xi_{t+1,j} - \rho\xi_{t,j})^2}{2\tau^2\lambda_{j,j}}, & 1 < t < T \\ \sum_{i=1}^{N} \left[I(Z_{T,i} > 0) \{ \eta_{T,i}\, y_{T,i} - \exp(\eta_{T,i}) \} \right] - \frac{(\xi_{T,j} - \rho\xi_{T-1,j})^2}{2\tau^2\lambda_{j,j}}, & t = T \end{cases}$$

$$h'(\xi_{t,j}) \propto$$
$$\begin{cases} \sum_{i=1}^{N} \left[I(Z_{1,i} > 0) \psi_{i,j} \{ y_{1,i} - \exp(\eta_{1,i}) \} \right] + \frac{\rho(\xi_{2,j} - \rho\xi_{1,j})}{\tau^2\lambda_{j,j}} - \frac{\xi_{1,j}}{\sigma_\xi^2}, & t = 1 \\ \sum_{i=1}^{N} \left[I(Z_{t,i} > 0) \psi_{i,j} \{ y_{t,i} - \exp(\eta_{t,i}) \} \right] - \frac{(\xi_{t,j} - \rho\xi_{t-1,j}) - \rho(\xi_{t+1,j} - \rho\xi_{t,j})}{\tau^2\lambda_{j,j}}, & 1 < t < T \\ \sum_{i=1}^{N} \left[I(Z_{T,i} > 0) \psi_{i,j} \{ y_{T,i} - \exp(\eta_{T,i}) \} \right] - \frac{\xi_{T,j} - \rho\xi_{T-1,j}}{\tau^2\lambda_{j,j}}, & t = T, \end{cases}$$

where $\eta_{t,i} = \tilde{\mathbf{x}}_{t,i}\boldsymbol{\alpha} + \boldsymbol{\Psi}_i'\boldsymbol{\xi}_t$ and $\sigma_\xi^2 = 10^6$.

Metropolis-Hastings sampling on ϕ_1

We draw ϕ_1 using the Metropolis-Hastings algorithm. For the m^{th}-current MCMC iteration, $\log(\phi_1^*)$ is drawn from $N(\log(\phi_1^{\text{old}}), \delta_{\phi_1})$. The full conditional density of ϕ_1 is given by

$$\pi(\phi_1|\cdot) \propto \exp\left[\sum_{t=1}^{T} \sum_{i=1}^{N} \{ (\tilde{\mathbf{x}}_{t,i}'\boldsymbol{\alpha} + \boldsymbol{\Psi}_{t,i}(\phi_1)\boldsymbol{\xi}_t) y_{t,i} - \exp(\tilde{\mathbf{x}}_{t,i}'\boldsymbol{\alpha} + \boldsymbol{\Psi}_{t,i}(\phi_1)\boldsymbol{\xi}_t) \} \right] \cdot \phi_1^{-\alpha_{\phi_1} - 1} \exp(-h/\phi_1),$$

where $\alpha_{\phi_1} = 2$ and $h = \max \text{distance}/[-2\log(0.05)]$. The new value is accepted with the probability given by

$$\min\left\{ 1, \frac{\pi(\phi_1^*|\cdot) \cdot \phi_1^*}{\pi(\phi_1^{\text{old}}|\cdot) \cdot \phi_1^{\text{old}}} \right\}.$$

Gibbs sampling on ρ

The full conditional density of ρ is a truncated normal $N(\mu_\rho, \sigma_\rho^2)I(-1,1)$ with

$$\sigma_\rho^2 = \left(\sum_{t=2}^{T} \boldsymbol{\xi}_{t-1}' \tau^{-2} \boldsymbol{\Lambda}^{-1} \boldsymbol{\xi}_{t-1}\right)^{-1},$$

$$\mu_\rho = \sigma_\rho^2 \sum_{t=2}^{T} \boldsymbol{\xi}_{t-1}' \tau^{-2} \boldsymbol{\Lambda}^{-1} \boldsymbol{\xi}_{t}.$$

Gibbs sampling on τ^2

The full conditional density of τ^2 is an inverse Gamma distribution $IG(c^*, d^*)$ with $c^* = c + \frac{M(T-1)}{2}$, $d^* = d + \sum_{t=2}^{T}(\xi_t - \rho\xi_{t-1})'\Lambda^{-1}(\xi_t - \rho\xi_{t-1})/2$, $c = 2$ and $d = 1$.

S2 The Algorithm for Particle Learning and the Log Predictive Score

The proposed spatial-temporal model contains a Poisson dynamic linear model (DLM) in (1) and (2) with the observational and state equations:

$$\begin{aligned} \log(\lambda_{t,i}) &= \tilde{\mathbf{x}}_{t,i}' \boldsymbol{\alpha} + \boldsymbol{\Psi}(\mathbf{s}_i)\boldsymbol{\xi}_t, \\ \boldsymbol{\xi}_t &= \rho \boldsymbol{\xi}_{t-1} + v_i, \;\; v_i \sim N_M(\mathbf{0}, \tau^2 \boldsymbol{\Lambda}). \end{aligned}$$

For this DLM, it can be shown that, for $t = 2, \ldots, T$, the prior (propagation) density is $\boldsymbol{\xi}_t | \boldsymbol{\lambda}^{t-1} \sim N(\mathbf{a}_t, \mathbf{R}_t)$, the predictive density is $\lambda_{t,i}|\lambda^{t-1} \sim N(\mathbf{f}_{t,i}, \mathbf{Q}_{t,i})$, and the posterior (filtering) density is $\boldsymbol{\xi}_t|\boldsymbol{\lambda}^t \sim N(m_t, C_t)$, where $\boldsymbol{\lambda}^t = (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_t)'$, $\boldsymbol{\lambda}_t = (\lambda_{t,1}', \ldots, \lambda_{t,N}')'$ and $N(a, b)$ denotes the normal distribution with mean a and variance b. The means and variances for these densities can be obtained using standard Kalman filtering recursions (West and Harrison (1997)) as follows:

$$\underset{M\times 1}{\mathbf{a}_t} = \rho \underset{M\times 1}{\mathbf{m}_{t-1}}, \qquad \underset{M\times M}{\mathbf{R}_t} = \rho^2 \underset{M\times M}{\mathbf{C}_{t-1}} + \tau^2 \underset{M\times M}{\boldsymbol{\Lambda}},$$

$$\underset{M\times 1}{\mathbf{m}_t} = \underset{M\times 1}{\mathbf{a}_t} + \underset{M\times N}{A_t} \underset{N\times 1}{\mathbf{e}_t}, \qquad \underset{M\times M}{\mathbf{C}_t} = \underset{M\times M}{\mathbf{R}_t} - \underset{M\times N}{A_t} \underset{N\times N}{Q_t} \underset{N\times M}{A_t'},$$

$$\underset{M\times 1}{A_{t,i}} = \underset{M\times M}{R_t} \underset{M\times 1}{\boldsymbol{\Psi}(\mathbf{s}_i)'} \underset{1\times 1}{Q_{t,i}^{-1}}, \qquad \underset{1\times 1}{\mathbf{e}_{t,i}} = \underset{1\times 1}{\lambda_{t,i}} - \underset{1\times 1}{\mathbf{f}_{t,i}},$$

$$\mathbf{f}_{t,i} = \tilde{\mathbf{x}}_{t,i}' \boldsymbol{\alpha} + \underset{1\times M}{\boldsymbol{\Psi}_i} \underset{M\times 1}{\mathbf{a}_t}, \qquad \underset{1\times 1}{Q_{t,i}} = \underset{1\times M}{\boldsymbol{\Psi}(\mathbf{s}_i)} \underset{M\times M}{R_t} \underset{M\times 1}{\boldsymbol{\Psi}(\mathbf{s}_i)'}.$$

The basic particle learning algorithm can be found in Carvalho, Johannes, Lopes, and Polson (2010) and Lopes and Tsay (2011). Assume that the static parameters $(\boldsymbol{\alpha}, \phi_1, \rho, \tau^2)$ and the binary probabilities of presence $(p_{t,i}, t = 1, \dots, T, i = 1, \dots, N)$ are specified as their corresponding posterior means from the MCMC simulation. The algorithm for particle learning and the log predictive score is stated as follows.

- *Resample:* Sample $\boldsymbol{\lambda}_{t+1,i}^{(j)} | \boldsymbol{\lambda}_t \sim N(\log(\boldsymbol{\lambda}_{t+1,i}); \mathbf{f}_{t+1,i}, Q_{t+1,i})$. The indices $\{\zeta(l)\}_{l=1}^{J}$ are resampled with the weights

$$p(\zeta(l) = j) = \prod_{i=1}^{N} \{P(Y_{i,t+1} = y_{i,t+1} | \lambda_{t+1,i}^{(j)})\}^{I_{t+1,i}},$$

  where

$$P(Y_{t+1,i} = y_{t+1,i} | \lambda_{t+1,i}^{(j)}) = \begin{cases} (1 - p_{t+1,i}) + p_{t+1,i} \cdot \text{Poisson}(y_{t+1,i}; \lambda_{t+1,i}^{(j)}) & \text{if } y_{t+1,i} = 0, \\ p_{t+1,i} \cdot \text{Poisson}(y_{t+1,i}; \lambda_{t+1,i}^{(j)}) & \text{if } y_{t+1,i} > 0, \end{cases}$$

  $p_{t+1,i}$ is the binary probability of presence at grid i in year $t+1$, $I_{t+1,i}$ is the survey indicator variable, $I_{t+1,i} = 1$ if grid i is surveyed in year $t+1$ and 0 otherwise, and $\text{Poisson}(z; \gamma)$ is the density at integer z of a Poisson distribution with mean γ.

- *Propagate:* The state sufficient statistics are the updates using the deterministic relations obtained in the standard DLM recursions. A new state parameter is then propagated from the posterior density $p(\xi_{t+1} | (m_{t+1}, C_{t+1})^{\zeta(l)})$.

- *The log predictive likelihood:* This can be obtained by

$$\begin{aligned} & \log p(Y_{t+1} | \mathbf{Y}_{1:t}, \text{static parameter}, \text{sufficient state statistics at } t) \\ \approx \quad & \frac{1}{J} \sum_{l=1}^{J} \sum_{i=1}^{N} I_{t+1,i} \log P^{\zeta(l)}(Y_{t+1,i} = y_{t+1,i}), \end{aligned}$$

  where $\zeta(l)$ is the resampling indicator and $P(Y_{t+1,i} = y_{t+1,i})$ is the probability of observing the data $y_{t+1,i}$. The survey indicator variable $I_{t+1,i}$ equals 1 if the grid i was surveyed in year $t+1$ and 0 otherwise. $P^{\zeta(l)}(Y_{t+1,i} = y_{t+1,i})$ is defined as

$$\begin{cases} (1 - p_{t+1,i}) + p_{t+1,i} \cdot \text{Poisson}(y_{t+1,i}; \lambda_{t+1,i}^{\zeta(l)}) & \text{if } y_{t+1,i} = 0, \\ p_{t+1,i} \cdot \text{Poisson}(y_{t+1,i}; \lambda_{t+1,i}^{\zeta(l)}) & \text{if } y_{t+1,i} > 0. \end{cases}$$

S3 Additional Discussion

The negative binomial model offers an alternative two-parameter family that is comparable to the ZIP model. The Poisson model is in fact a special case of the negative binomial model with the heterogeneity parameter going to zero (Hilbe (2007)). However, parameter estimation for the negative binomial model with covariates is not as straightforward as the ZIP model (Agarwal, Gelfand, and Citron-Pousty (2002)). Most noticeably, the negative binomial model does not provide a parameter for the probability of the presence of the species and its interpretation is not as easy as that of the ZIP model. The computation in the spatial dynamic model is also not stable (Fernandes, Schmidt, and Migon (2009)). Due to these reasons, we do not consider a negative binomial model to fit the ecological count data.

Another interesting approach is to relate the knots to biologically meaningful locations, such as species "hot spots" or "cold spots." We have tried to select the knots based on the spawning locations. O'Brien, Gregory, Mayo, and Hunt (2005) describe a general seasonal movement caused by spawning: the cod move inshore in the autumn to the spawning grounds, and offshore in the summer into the central part of the area for the Gulf of Maine stock, while on Georges Bank, the cod form spawning aggregations in the late autumn and early spring and disperse after spawning by moving to the northeast and southwest. Dynamics of the cod stocks are reflected in the NEFSC data. The distributions of the cods are more scattered in the spring survey than those in the fall. O'Brien, Gregory, Mayo, and Hunt (2005) summarized the general spawning areas for the cods in the Gulf of Maine and Georges Bank areas. For details, see the references cited in O'Brien, Gregory, Mayo, and Hunt (2005). However, our results showed that simply putting the knots around the spawning locations did not significantly improve the fit of the model. Better tools are needed for biologically-related knot locations.

Multiscale spatial-temporal modeling is another approach to modeling the dynamic evolution of massive ecological data. It achieves scalable and efficient computation by decomposing large data analysis problems into many smaller components (multiscale) (Ferreira, Holan, and Bertolde (2011)). The multiscale Poisson spatial dynamic model has been developed by Ferreira and Fonseca (2013) for tornado data, where the administrative county boundary is used for the finest levels to construct the multiscale structure. To implement this method for ecological data, a challenging problem is how to appropriately identify and cluster distinct areas of the ecological population when the observed data are sparse, such as the Atlantic cod data we studied here.

S4 Figures

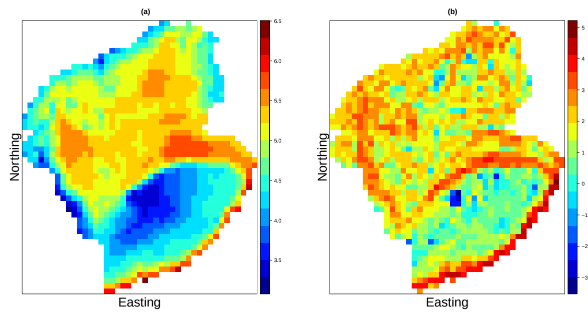

Figure S1: The ocean depth in the Gulf of Maine - Georges Bank region (in the unit of natural logarithm): (a) left panel: the mean ocean depth; (b) right panel: the standard deviation of the ocean depth at each grid. The whole study area is divided into 1325 grids.

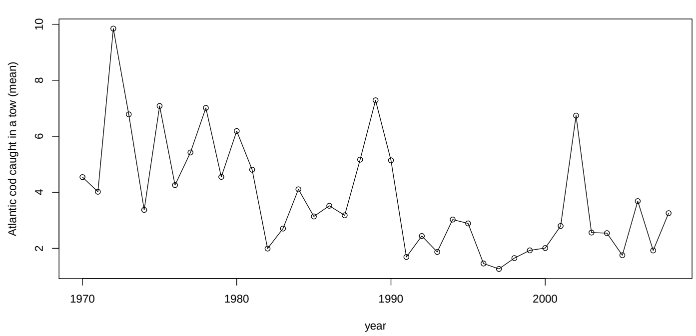

Figure S2: The average number of Atlantic cod caught in a tow in the Gulf of Maine - Georges Bank region during 1970-2008.

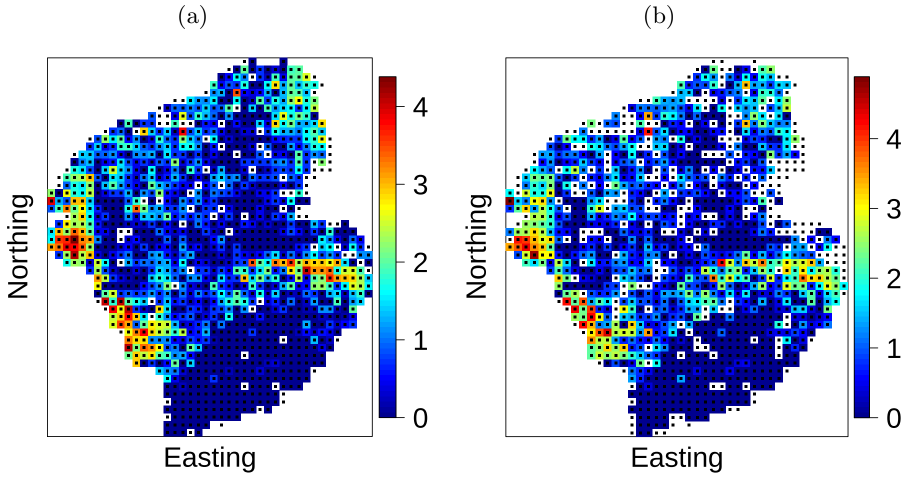

Figure S3: The whole study area is divided into 1325 grids with a black dot indicating the center of the grid. (a) The average number of fish caught at each grid during 1970-2008. To ease visualization, it is calculated as logarithm of (1+average number of fish caught at each grid during 1970-2008). The white grids are those that have never been sampled. (b) The standard deviation of fish caught at each grid during the year 1970-2008. To ease visualization, it is calculated as logarithm of (1+ standard deviation of fish caught at each grid during 1970-2008). The white grids are those that have never been sampled or had only one tow during 1970-2008.

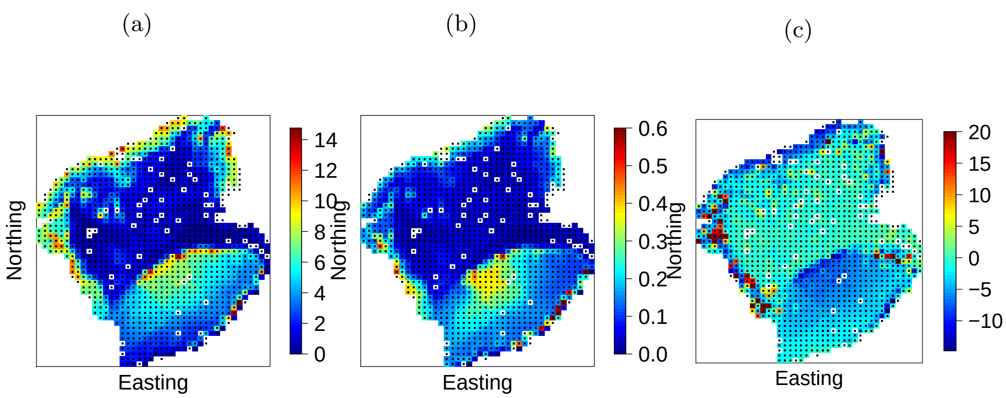

Figure S4: Posterior estimates under Model 1 (in the unit of counts): (a) posterior mean count at each grid; (b) standard deviation of the posterior mean count at each grid; and (c) residuals.

Figure S5: Knots selected. Left panel: 16 evenly spaced knots were selected in the study area. Right panel: 32 evenly spaced knots were selected in the study area.

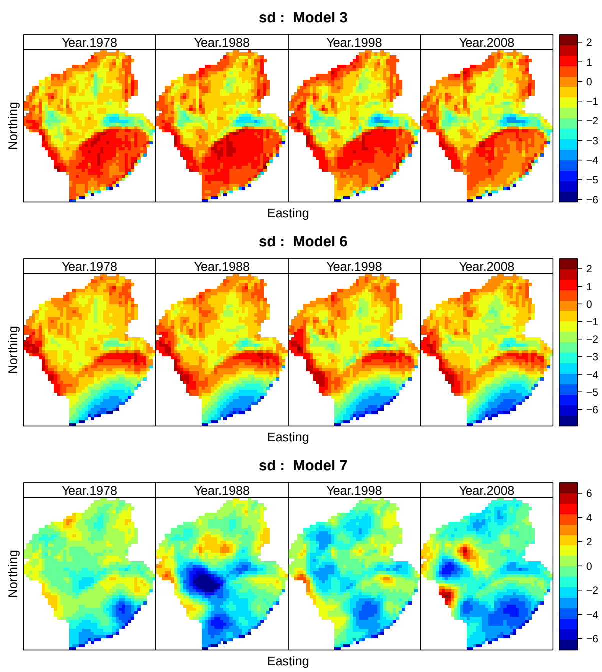

Figure S6: The logarithm of the standard deviation of posterior mean count from Model 3, Model 6 (32 knots spatial), and Model 7 (32 knots spatial-temporal) at each grid.

References

Agarwal, D. K., Gelfand, A. E. and Citron-Pousty S. (2002). Zero-inflated models with application to spatial count data. *Environ. Ecol. Stat.* 9, 341–355.

Carvalho, C. M., Johannes, M. S., Lopes, H. F. and Polson, N. G. (2010). Particle learning and smoothing. *Statist. Sci.* 25, 88–106.

Fei, S. and Rathbun, S. L. (2006). A spatial zero-inflated Poisson model for oak regeneration. *Environ. Ecol. Stat.* 13, 406–426.

Fernandes, M. V., Schmidt, A. M. and Migon, H. S. (2009). Modelling zero-inflated spatio-temporal processes. *Stat. Model.* 9, 3–25.

Ferreira, M. A. R. and Fonseca, T. C. O. (2013). Dynamic Multiscale Spatiotemporal Models for Poisson Data. Presented at SAMSI MD Transitional Workshop.

Ferreira, M. A. R., Holan, S. H. and Bertolde, A. I. (2011). Dynamic multiscale spatiotemporal models for Gaussian areal data. *J. Roy. Statist. Soc. B* 73, 661–688.

Gilks W. R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *J. Roy. Statist. Soc. C* 41, 337–348.

Hilbe, J. M. (2007). *Negative Binomial Regression.* Cambridge University Press, Cambridge.

Lopes, H. F. and Tsay, R. S. (2011). Particle filters and Bayesian inference in financial econometrics. *J. of Forecasting* 30, 168–209.

O'Brien, L., Gregory, L. R., Mayo, R. K., Hunt and J. J. (2005). Gulf of Maine and Georges Bank (NAFO Subareas 5 and 6). In: Spawning and Life History Information for North Atlantic Cod Stocks, Brander, K. (ed), 95–104, Prepared by the ICES/GLOBEC Working Group on Cod and Climate Change, ICES Cooperative Research Report No. 274.

West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*, 2nd edn. Springer, New York.