Self-Exciting Spatio-Temporal Models for Count Data

Nicholas Clark

Iowa State University

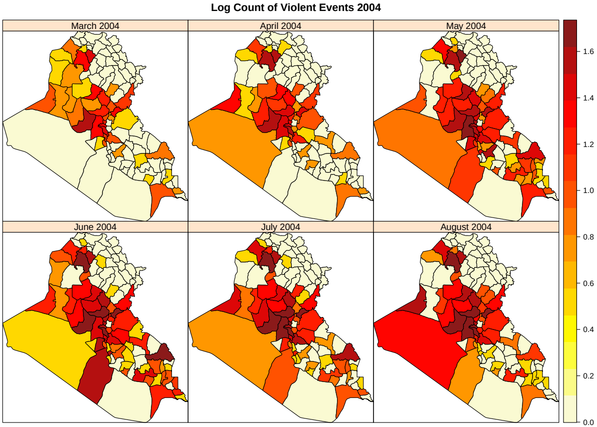
nclark1@iastate.edu

February 9, 2018

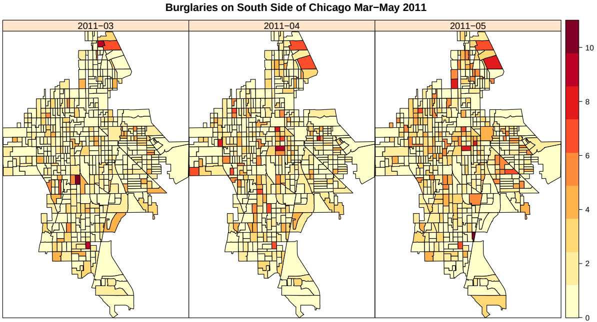
Motivation: The Evolution of Violence in Space and Time

At present, the most under-researched area of spatial criminology is that of spatio-temporal crime patterns... the temporal component of the underlying crime distributions has languished as a largely ignored area of study - **Crime mapping: Spatial and Temporal Challenges**, Ratcliffe (2010)

The Spread of Violence in Iraq 2004



Burglaries South Side of Chicago



Goals

- General statistical model for diffusion of violence in space-time
  - Accurately reflects beliefs on how violence/crime evolves
  - Extends traditional statistical models for count data
  - Stationary with extremely flexible second order properties
  - Inference via traditional MCMC techniques

Overview

A Model of Criminal Behavior (Short et al. 2008)

- $Z(s_i, t)$ - number of observed burglaries from $(t - \Delta t, t)$
- $s_i \in \{s_1, \cdots, s_{n_d}\}$ - fixed regions in $\mathbb{R}^2$
- $t \in \{1, \cdots, T\}$ - discrete time
- Define $A(s_i, t) \equiv A(s_i, 0) + B(s_i, t)$ as attractiveness

$$B(s_i, t) = (1 - \chi \Delta t)\, B(s_i, t - \Delta t) + \eta Z(s_i, t - \Delta t) \qquad (1)$$

- $Z(s_i, t) \sim \text{Pois}\,(A(s_i, t))$
- Three factors impact change in crime rate, base attractiveness , decay χ, and repeat victimization, η

Relationship to INGARCH Model

Integer Auto-Regressive Conditionally Heteroskedastic, INGARCH (1,1), or Poisson Auto-Regression Model

$$Z(s_i, t) \sim \text{Pois}\left(\lambda(s_i, t)\right) \tag{2}$$
$$\lambda(s_i, t) = d + a\lambda(s_i, t-1) + bZ(s_i, t-1) \tag{3}$$

- Unlike GARCH, not solely a variance property
- Short model is similar to INGARCH(1,1) with $A(s_i, 0) = \sum_{k=0}^{t} a^k d$, $a = (1 - \chi \Delta t)$, and $b = \eta$

Relationship to Self-Exciting Models

- Point process introduced by Alan Hawkes with intensity

$$\lambda(t) = \nu(t) + \int_0^t g(t-u)N(ds) \tag{4}$$

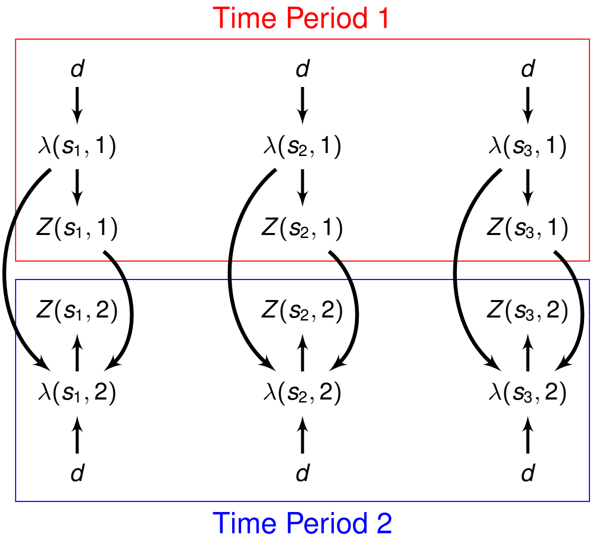
- Commonly discretized as

$$Z(s_i, t) \sim \text{Pois}\left(\lambda(s_i, t)\right) \tag{5}$$

$$\lambda(s_i, t) = \nu + \sum_{j<t} \eta^{t-j} Z(s_i, t-j) \tag{6}$$

- Equivalent to stationary INGARCH(1,1)

Structural Diagram - INGARCH(1,1)



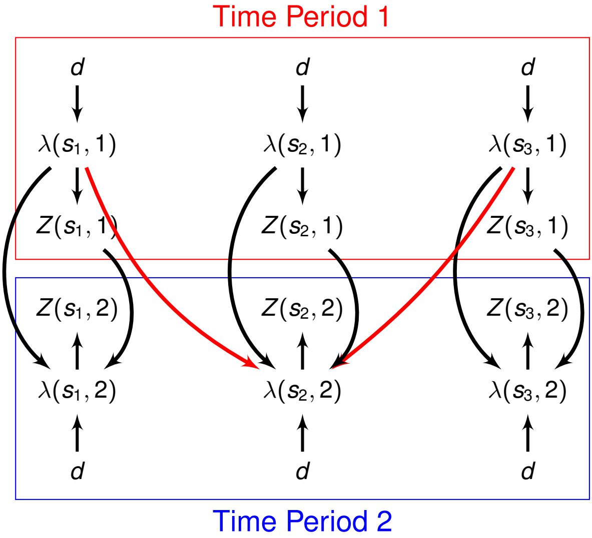
Short (2008) Extension - Spatial Spread

- Motivated by Reaction-Diffusion PDE
- Recall $A(s_i, t) \equiv A(s_i, 0) + B(s_i, t)$ as attractiveness
- Define: $N_i = \{s_j : s_j \text{ is a spatial neighbor of } s_i\}$, $|N_{s_i}|$ is number of neighbors

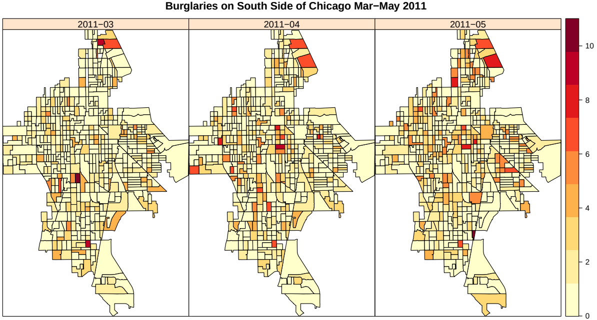
$$B(s_i, t) = \kappa B(s_i, t - \Delta t) + \frac{\psi}{|N_{s_i}|} \sum_{s_j \in N_i} \left[B(s_j, t - \Delta t) - B(s_i, t - \Delta t)\right] + \eta Z(s_i, t - \Delta t) \tag{7}$$

- Four factors impact change in crime rate, base attractiveness, decay κ, and repeat victimization, η, and spatial spread ψ
- Resulting model is MINGARCH (1,1)

Short (2008) Extension - Spatial Spread



Applied to Residential Burglaries in Chicago



552 Spatial Locations, 72 Months, residential burglaries

Applied to Residential Burglaries in Chicago

$$
\begin{aligned}
B(s_i, t) = &\ \kappa B(s_i, t - \Delta t) \\
&+ \frac{\psi}{|N_{s_i}|} \sum_{s_j \in N_{s_i}} \left[B(s_j, t - \Delta t) - B(s_i, t - \Delta t) \right] + \eta Z(s_i, t - \Delta t)
\end{aligned} \tag{8}
$$

- Further structure $A(s_i, 0)$ to account for socio-economic factors
- MLEs are $\hat{\psi} = .02$, $\hat{\eta} = .23$ $\hat{\kappa} = .461$
- Simulate data from asymptotic distribution - unable to replicate lag-one autocorrelation, spatial correlation, or variance to mean ratio of original data

Properties of INGARCH Model

$$Z(s_i, t) \sim \text{Pois}\left(\lambda(s_i, t)\right)$$
$$\lambda(s_i, t) = d + a\lambda(s_i, t-1) + bZ(s_i, t-1)$$

Stationarity yields:

$$E[Z(s_i, t)] = \frac{d}{1-(a+b)} \tag{9}$$

$$\text{Var}[Z(s_i, t)] = \frac{1-(a+b)^2+b^2}{1-(a+b)^2} E[Z(s_i, t)] \tag{10}$$

$$\text{Cov}[Z(s_i, t), Z(s_i, t-h)] = \frac{b(1-a(a+b))(a+b)^h}{1-(a+b)^2} E[Z(s_i, t)] \tag{11}$$

$$\text{Var-Mean Ratio}[Z(s_i, t)] = 1 + \frac{b^2}{1-(a+b)^2} \tag{12}$$

Issues

Allows for Overdispersion... But at a cost!

$$\text{Cor}[Z(s_i,t), Z(s_i,t-1)] = \frac{b(a+b)(a^2+ab-1)}{a^2+2ab-1} \quad (13)$$

$$\text{Var-Mean Ratio}[Z(s_i,t)] = 1 + \frac{b^2}{1-(a+b)^2} \quad (14)$$

$$(15)$$

For fixed Var-Mean Ratio at 2 $\implies b = 1/2(-a+\sqrt{2-a^2})$.
Implies Lag-one Cor in (.5,.707)

For fixed Var-Mean Ratio at 1.2, Implies Lag-one Cor in (.2,.29)

Actual crime data, Var-Mean is 2, Lag-one Cor is .3

Spatially Correlated Self-Exciting Model (Clark & Dixon, 2018)

$Y(s_i, t)$ - Spatially Correlated Latent Gaussian

Time Period 1

$$
\begin{tikzcd}
\exp(Y(s_1,1)) \arrow[r, leftrightarrow] \arrow[d] & \exp(Y(s_2,1)) \arrow[r, leftrightarrow] \arrow[d] & \exp(Y(s_3,1)) \arrow[d] \\
\lambda(s_1,1) \arrow[d] & \lambda(s_2,1) \arrow[d] & \lambda(s_3,1) \arrow[d] \\
Z(s_1,1) \arrow[dd, bend right] & Z(s_2,1) \arrow[dd, bend right] & Z(s_3,1) \arrow[dd, bend right] \\
Z(s_1,2) & Z(s_2,2) & Z(s_3,2) \\
\lambda(s_1,2) \arrow[u] & \lambda(s_2,2) \arrow[u] & \lambda(s_3,2) \arrow[u] \\
\exp(Y(s_1,2)) \arrow[r, leftrightarrow] \arrow[u] & \exp(Y(s_2,2)) \arrow[r, leftrightarrow] \arrow[u] & \exp(Y(s_3,2)) \arrow[u]
\end{tikzcd}
$$

Time Period 2

- **Theory:** Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation

- **Theory:** Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation
- Mixture of two processes that influence expectation : LGCP and Hawkes process
- Hawkes process letting $g(t-j)=\eta$ if $(t-j)=1$, 0 otherwise

$$\begin{aligned}
Z(\boldsymbol{s_i},t)|\lambda(\boldsymbol{s_i},t) &\sim \text{Pois}\left(\lambda(\boldsymbol{s_i},t)\right) \\
\lambda(\boldsymbol{s_i},t) &= \exp(Y(\boldsymbol{s_i},t)) + \eta Z(\boldsymbol{s_i},t-1) \\
Y(\boldsymbol{s_i},t) &= \theta_1 \sum_{\boldsymbol{s_j}\in N(\boldsymbol{s_i})} Y(\boldsymbol{s_j},t) + \epsilon(\boldsymbol{s_i},t) \\
\epsilon(\boldsymbol{s_i},t) &\sim Gau(0,\sigma^2)
\end{aligned} \tag{16}$$

$Y(s_i, t)$ - Spatially Correlated Latent Gaussian

Time Period 1

$$
\begin{tikzcd}
\exp(Y(s_1,1)) \arrow[r, leftrightarrow] \arrow[d] & \exp(Y(s_2,1)) \arrow[r, leftrightarrow] \arrow[d] & \exp(Y(s_3,1)) \arrow[d] \\
\lambda(s_1,1) \arrow[d] & \lambda(s_2,1) \arrow[d] & \lambda(s_3,1) \arrow[d] \\
Z(s_1,1) \arrow[dd, bend right] & Z(s_2,1) \arrow[dd, bend right] & Z(s_3,1) \arrow[dd, bend right] \\
Z(s_1,2) & Z(s_2,2) & Z(s_3,2) \\
\lambda(s_1,2) \arrow[u] & \lambda(s_2,2) \arrow[u] & \lambda(s_3,2) \arrow[u] \\
\exp(Y(s_1,2)) \arrow[r, leftrightarrow] \arrow[u] & \exp(Y(s_2,2)) \arrow[r, leftrightarrow] \arrow[u] & \exp(Y(s_3,2)) \arrow[u]
\end{tikzcd}
$$

Time Period 2

SPINGARCH(1,1) Model

Spatially Correlated INGARCH(1,1) Model

Time Period 1

$$
\begin{tikzcd}
\exp(Y(s_1,1)) \arrow[r, leftrightarrow] \arrow[d] & \exp(Y(s_2,1)) \arrow[r, leftrightarrow] \arrow[d] & \exp(Y(s_3,1)) \arrow[d] \\
\lambda(s_1,1) \arrow[d] \arrow[ddd, bend right] & \lambda(s_2,1) \arrow[d] \arrow[ddd, bend right] & \lambda(s_3,1) \arrow[d] \arrow[ddd, bend right] \\
Z(s_1,1) \arrow[dd, bend left] & Z(s_2,1) \arrow[dd, bend left] & Z(s_3,1) \arrow[dd, bend left] \\
Z(s_1,2) & Z(s_2,2) & Z(s_3,2) \\
\lambda(s_1,2) \arrow[u] & \lambda(s_2,2) \arrow[u] & \lambda(s_3,2) \arrow[u] \\
\exp(Y(s_1,2)) \arrow[r, leftrightarrow] \arrow[u] & \exp(Y(s_2,2)) \arrow[r, leftrightarrow] \arrow[u] & \exp(Y(s_3,2)) \arrow[u]
\end{tikzcd}
$$

Time Period 2

SPINGARCH Model

Theory: Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation, ***absence of violence or exogeneous effects reduces tension***

SPINGARCH Model

Theory: Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation, ***absence of violence or exogeneous effects reduces tension***

- $Z(s_i, t) | Y(s_i, t), \mathcal{H}_{Z(s_i)} \sim \text{Pois}\,(\lambda(s_i, t))$ where $\mathcal{H}_{Z(s_i)}$ is history of process at location s_i

$$\lambda(s_i, t) = \exp\left[Y(s_i, t)\right] + \eta Z(s_i, t-1) + \kappa\lambda(s_i, t-1) \tag{17}$$

SPINGARCH Model

Theory: Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation, ***absence of violence or exogeneous effects reduces tension***

- $Z(s_i, t) | Y(s_i, t), \mathcal{H}_{Z(s_i)} \sim$ Pois $(\lambda(s_i, t))$ where $\mathcal{H}_{Z(s_i)}$ is history of process at location s_i

$$\lambda(s_i, t) = \exp\left[Y(s_i, t)\right] + \eta Z(s_i, t-1) + \kappa\lambda(s_i, t-1) \tag{17}$$

$$Y(s_i, t) | \boldsymbol{Y}(N_i) \sim N(\mu(s_i, t), \sigma^2_{sp}) \tag{18}$$

$$\mu(s_i, t) = \alpha(s_i) + \zeta \sum_{s_j \in N_i} \{Y(s_j, t) - \alpha(s_j)\}.$$

SPINGARCH Model

Theory: Exist common cause between geographically similar locations, regions that experience uptick in violence likely to have short term self-excitation, ***absence of violence or exogeneous effects reduces tension***

- $Z(s_i,t) | Y(s_i,t), \mathcal{H}_{Z(s_i)} \sim$ Pois $(\lambda(s_i,t))$ where $\mathcal{H}_{Z(s_i)}$ is history of process at location s_i

$$\lambda(s_i,t) = \exp\left[Y(s_i,t)\right] + \eta Z(s_i,t-1) + \kappa\lambda(s_i,t-1) \tag{17}$$

$$Y(s_i,t) | \boldsymbol{Y}(N_i) \sim N(\mu(s_i,t), \sigma^2_{sp}) \tag{18}$$

$$\mu(s_i,t) = \alpha(s_i) + \zeta \sum_{s_j \in N_i} \{Y(s_j,t) - \alpha(s_j)\}.$$

- $\eta = 0, \kappa = 0$ Poisson - CAR, $\sigma^2_{sp} \to 0$, INGARCH(1,1)/Short model

Relationship to INGARCH(1,1)/Short Model

$$\frac{\lambda(s_i, t) - \lambda(s_i, t-1)}{\Delta t} = d - \chi\lambda(s_i, t-1) + \eta Z(s_i, t-1) \qquad (19)$$

- Change in violence due to exogeneous d, natural decay, χ, and excitement, η

Relationship to INGARCH(1,1)/Short Model

$$\frac{\lambda(s_i, t) - \lambda(s_i, t-1)}{\Delta t} = d - \chi\lambda(s_i, t-1) + \eta Z(s_i, t-1) \tag{19}$$

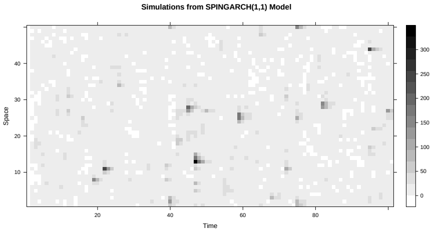
- Change in violence due to exogeneous d, natural decay, χ, and excitement, η
- Assume each time period, exogeneous impact is stochastic and spatially correlated yields SPINGARCH

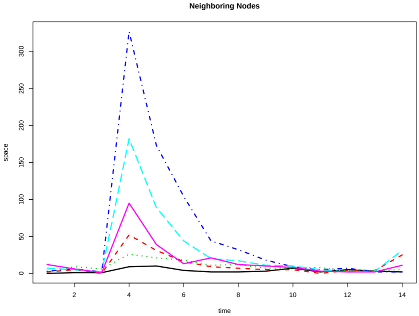
$$\frac{\lambda(s_i, t) - \lambda(s_i, t-1)}{\Delta t} = \exp(Y(s_i, t)) - \chi\lambda(s_i, t-1) + \eta Z(s_i, t-1) \tag{20}$$

- Change in intensity due to three factors, spatially correlated exogeneous, natural decay, and excitement

Data Realizations

50 Spatial Observations on $\mathbb{R}^1$, 100 Temporal Observations





$$
\begin{aligned}
\lambda(s_i, t) &= \exp\left[Y(s_i, t)\right] + 0.1\, Z(s_i, t-1) + 0.4\, \lambda(s_i, t-1) \\
Y(s_i, t) | \boldsymbol{Y}(N_i) &\sim \text{Gau}(\mu(s_i, t), 0.5) \\
\mu(s_i, t) &= 0 + 0.49 \sum_{s_j \in N_i} \{Y(s_j, t)\}.
\end{aligned} \tag{21}
$$

Model Properties - Parameter Space

- $Z(s_i, t) | Y(s_i, t), \mathcal{H}_{Z(s_i)} \sim$ Pois $(\lambda(s_i, t))$ where $\mathcal{H}_{Z(s_i)}$ is history of process at location s_i

$$\lambda(s_i, t) = \exp\left[Y(s_i, t)\right] + \eta Z(s_i, t-1) + \kappa \lambda(s_i, t-1)$$
$$Y(s_i, t) | \boldsymbol{Y}(N_i) \sim N(\mu(s_i, t), \sigma^2_{sp})$$
$$\mu(s_i, t) = \alpha(s_i) + \zeta \sum_{s_j \in N_i} \{Y(s_j, t) - \alpha(s_j)\}.$$

- $\zeta \in (\psi^{-1}_{(1)}, \psi^{-1}_{(n)})$ where $\psi_{(i)}$ is the ith largest eigenvalue of adjacency matrix
- For stationarity, $\eta > 0$, $\kappa > 0$, $\eta + \kappa < 1$

Model as Markov Chain

- Let $\boldsymbol{\lambda_t} = (\lambda(s_1, t), \lambda(s_2, t), \cdots, \lambda(s_{n_d}, t))^T$
- $\boldsymbol{C}$ is $n_d \times n_d$ with $C(i, j) = \zeta$ if $s_j \in N_i$
- $\boldsymbol{M} = \text{diag } \sigma^2_{sp}$

$$\begin{aligned}
Z(s_i, t) | \lambda(s_i, t) &\sim \text{Pois}(\lambda(s_i, t)) \\
E[Z(s_i, t)] &= \lambda(s_i, t) \\
\boldsymbol{\lambda_t} &= \exp(\boldsymbol{Y_t}) + \eta \boldsymbol{Z_{t-1}} + \kappa \boldsymbol{\lambda_{t-1}} \\
\boldsymbol{Y_t} &\sim \text{Gau}(\boldsymbol{\alpha_t}, (I_{n_d, n_d} - \boldsymbol{C})^{-1} \boldsymbol{M})
\end{aligned} \tag{22}$$

- Markov chain for $\boldsymbol{\lambda_t}$ on State space, $(\mathbb{R}^+)^{n_d}$

Impact of Initial Condtions and Recursion

By recursion

$$
\begin{aligned}
&[\lambda(s_i,t)|\lambda(s_i,0)=B] = \exp(Y(s_i,t)) + \kappa\lambda(s_i,t-1) + \eta Z(s_i,t-1)\\
=&\exp(Y(s_i,t)) + \kappa\left[\exp(Y(s_i,t-1)) + \kappa\lambda(s_i,t-2)\right.\\
&\left.+\eta Z(s_i,t-2)\right] + \eta Z(s_i,t-1)\\
&\dots\\
=&\sum_{k=0}^{t-1}\kappa^k\exp(Y(s_i,t-k)) + \sum_{k=0}^{t-1}\kappa^k\eta Z(s_i,t-k-1) + \kappa^t B.
\end{aligned}
$$

Geometric Ergodicity with Finite Moments

Under the parameter space restriction, $\eta, \kappa \geq 0$ and $\eta + \kappa < 1$, the SPINGARCH (1,1) is geometrically ergodic and admits a unique stationary distributions that has finite first two moments.

Sketch of Proof.

Meyn and Tweedie (15.0.1) need to show aperiodic, ϕ-irreducible and $\exists$ test function $V(.)$ such that $E[V(\boldsymbol{\lambda}_{t+1})|\boldsymbol{\lambda}_t = \boldsymbol{\lambda}_*] \leq \psi V(\boldsymbol{\lambda}_*) + L\,\mathrm{I}(\boldsymbol{\lambda}_* \in C)$ holds where $\psi \in (0,1)$, $L \in (0,\infty)$ and $I(.)$ is the indicator function and C is a petite set.

Basic Idea: With positive probability, $\exists$ a realization $Z(s_i,1) = Z(s_i,2) = \cdots = Z(s_i,t-1) = 0$. Along that chain, $P(\lambda(s_i,t)) \in A = P(\exp(Y(s_i,t)) \in A - \kappa^t B)$. If $\kappa^T B > \sup A$ run chain longer. $\square$

Geometric Ergodicity with Finite Moments

Under the parameter space restriction, $\eta, \kappa \geq 0$ and $\eta + \kappa < 1$, the SPINGARCH (1,1) is geometrically ergodic and admits a unique stationary distributions that has finite first two moments.

Sketch of Proof Cont.

Test function $V(\lambda) = 1 + \lambda^2$ works for
$E[V(\boldsymbol{\lambda}_{t+1})|\boldsymbol{\lambda}_t = \boldsymbol{\lambda}_*] \leq \psi V(\boldsymbol{\lambda}_*) + L \, \mathrm{I}(\boldsymbol{\lambda}_* \in C)$.

$\implies$ Unique stationary distribution, goes to geometrically fast.
Specific choice of $V(.)$ gives (at least) finite first two moments (can be extended likely as in Fokianos, 2009.)

$\square$

Increased Modeling Flexibility with SPINGARCH(1,1)

$$
\begin{aligned}
Z(s_i,t) &\sim \text{Pois}\left(\lambda(s_i,t)\right)\\
\lambda(s_i,t) &= \exp\left[Y(s_i,t)\right] + \eta Z(s_i,t-1) + \kappa E\left[Z(s_i,t-1)\right]\\
Y(s_i,t)|\boldsymbol{Y}(N_i) &\sim N(\mu(s_i,t),\sigma^2_{sp})\\
\mu(s_i,t) &= \alpha(s_i) + \zeta \sum_{s_j \in N_i}\left\{Y(s_j,t)-\alpha(s_j)\right\}
\end{aligned}
$$

Define $\Sigma_{i,j}$ as i,j entry of $(I_{n_d,n_d} - \boldsymbol{C})^{-1}\boldsymbol{M}$

$$
E\left[Z(s_i,t)\right] = \frac{1}{1-\eta-\kappa}\exp\left(\alpha + \frac{\Sigma_{1,1}}{2}\right) \tag{23}
$$

$$
\text{Var}\left(Z(s_i,t)\right) = \frac{1}{1-(\kappa+\eta)^2}\text{Var}\left(\exp(Y(s_i,t))\right) + \frac{1-\kappa^2-2\kappa\eta}{1-(\kappa+\eta)^2}E(Z(s_i,t)) \tag{24}
$$

Increased Modeling Flexibility with SPINGARCH(1,1)

Temporal Covariance:

$$\text{Cov}\left(Z(s_i,t), Z(s_i,t-1\right) = (\eta+\kappa)\text{Var}(Z(s_i,t)) - \kappa E\left[Z(s_i,t)\right] \quad (25)$$

Increased Modeling Flexibility with SPINGARCH(1,1)

Temporal Covariance:

$$\text{Cov}\left(Z(s_i,t),Z(s_i,t-1\right) = (\eta+\kappa)\text{Var}(Z(s_i,t)) - \kappa E\left[Z(s_i,t)\right] \quad (25)$$

Recall: Var-Mean Ratio at 2 $\implies$ Lag-1 correlation is between 0.5 and $\sqrt{1/2}$ for INGARCH(1,1)

Increased Modeling Flexibility with SPINGARCH(1,1)

Temporal Covariance:

$$\text{Cov}\left(Z(s_i,t), Z(s_i,t-1)\right) = (\eta+\kappa)\text{Var}(Z(s_i,t)) - \kappa E\left[Z(s_i,t)\right] \quad (25)$$

Recall: Var-Mean Ratio at 2 $\implies$ Lag-1 correlation is between 0.5 and $\sqrt{1/2}$ for INGARCH(1,1)

Let $\kappa = 0$ (SPINGARCH(0,1)), Var-Mean Ratio at 2

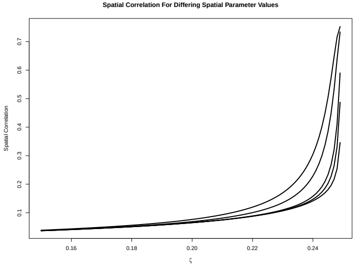
$$\implies 2 = \frac{\text{Var}\left(\exp(Y(s_i,t))\right)}{(1-\eta)^2 E\left[Z(s_i,t)\right]} + \frac{1}{1-\eta^2} \quad (26)$$

$$\text{Cor}\left(Z(s_i,t), Z(s_i,t-1)\right) = \eta \quad (27)$$

$\forall \eta \in (0, \sqrt{1/2}) \quad \exists \alpha, \sigma^2_{sp}$ such that equality holds

Spatial Correlation

$$\mathrm{Corr}(Z(s_i,t),Z(s_j,t)) = \frac{(\exp(\Sigma_{i,i}+\Sigma_{i,j})-\exp(\Sigma_{i,i}))}{\exp(2\Sigma_{i,i})-\exp(\Sigma_{i,i})+\exp(-\alpha+\frac{\Sigma_{i,i}}{2})\frac{1}{1-(\kappa+\eta)}}$$



$\eta = .3$, $\sigma^2_{sp} = .5$, 4×4 to 15×15 size lattice

Inference

- Likelihood roots for INGARCH(1,1) easily found, asymptotically Gaussian
- Inclusion of latent process in SPINGARCH(1,1) complicates
- $\boldsymbol{\theta} \equiv (\eta, \alpha, \zeta, \sigma^2_{sp})$

$$\pi(\boldsymbol{\theta}|\boldsymbol{Z}, \boldsymbol{Y}) \propto \prod_t \pi(\boldsymbol{Z}_t|\boldsymbol{\lambda}_t)\pi(\boldsymbol{\lambda}_t|\boldsymbol{\lambda}_{t-1}, \boldsymbol{Z}_{t-1}, \boldsymbol{\theta}, \boldsymbol{Y}_t)\pi(\boldsymbol{Y}_t|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \quad (28)$$

$$\pi(\boldsymbol{Y}|\boldsymbol{Z}, \boldsymbol{\theta}) \propto \prod_t \pi(\boldsymbol{Z}_t|\boldsymbol{\lambda}_t)\pi(\boldsymbol{\lambda}_t|\boldsymbol{\lambda}_{t-1}, \boldsymbol{Z}_{t-1}, \boldsymbol{\theta}, \boldsymbol{Y}_t)\pi(\boldsymbol{Y}_t|\boldsymbol{\theta}). \quad (29)$$

Efficient Bayesian Inference

$$\log(\boldsymbol{Y}|\boldsymbol{\alpha}, \sigma_{sp}, \zeta) \propto \frac{1}{2}\log|\Sigma_f^{-1}(\theta)| - \frac{1}{2}(Y-\alpha)^T\Sigma_f^{-1}(\theta)(Y-\alpha), \quad (30)$$

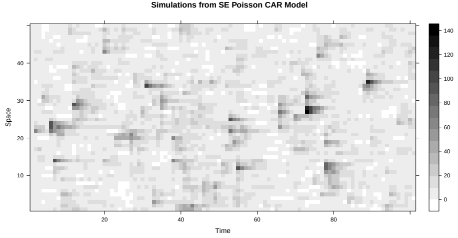
- $\Sigma_f^{-1} \equiv \left(I_{n_d \times T, n_d \times T} - I_{t,t} \otimes \boldsymbol{C}\right)^{-1} I_{t,t} \otimes \boldsymbol{M}$
- $\log|\Sigma^{-1}(\theta)| = \frac{n_d}{2\log\sigma_{sp}^2} + \log|I_{n_d,n_d} - \zeta N|$
- Letting $V\Lambda V^T$ be the spectral decomposition of N we have $|I_{n_d,n_d} - \zeta N| = |V||I_{n_d,n_d} - \zeta\Lambda||V^T| = \prod_{j=1}^{n_d}\left(1-\zeta\chi_j\right)$ where χ_j are the eigenvalues of the neighborhood matrix

$$\log|\Sigma_f^{-1}(\theta)| = T\log|\Sigma^{-1}(\theta)| \quad (31)$$

$$\propto \frac{n_d \times T}{\log\sigma_{sp}^2} + T\sum_{j=1}^{n_d}(1-\zeta\chi_j) \quad (32)$$

Impacts of Misspecificaiton

$$
\begin{aligned}
Z(s_i,t) &\sim \text{Pois}(\lambda(s_i,t)) \\
\lambda(s_i,t) &= \exp\left[Y(s_i,t)\right] + 0.66Z(s_i,t-1) \\
Y(s_i,t)|\boldsymbol{Y}(N_i) &\sim \text{Gau}(\mu(s_i,t),0.5) \\
\mu(s_i,t) &= 0 + 0.49\sum_{s_j\in N_i}\{Y(s_j,t)\}.
\end{aligned} \tag{33}
$$



Generating Mechanism is SPINGARCH(0,1)

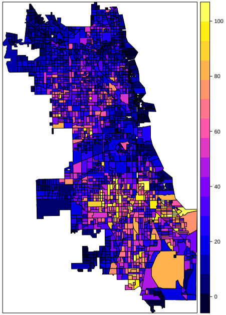
- Fit to SPINGARCH(0,1), SPINGARCH(1,0), and INGARCH(1,1)
- Vague, proper priors, e.g $\eta \sim$ Unif$(0, 1)$, $\zeta \sim$ Unif$(0, .5)$, $\sigma_{sp} \sim (Cau)^{+}(0, 1)$, $\alpha \sim$ Gau$(0, 100)$
- Model assessment using posterior predictive P values
  - Pick ancillary statistic, $T(.)$ and calculate $T(\boldsymbol{Z})$
  - for $m = 1...M$, draw a value of $\boldsymbol{\theta_m}$ according to $\pi(\boldsymbol{\theta}|\boldsymbol{Z})$
  - Simulate $Z^*(\boldsymbol{s_i}, \boldsymbol{t})_m$ of the same dimension as $\boldsymbol{Z}$ and compute $T(\boldsymbol{Z}^*_m)$
  - Compute $\frac{1}{M}\sum_{m=1}^{M} I[T(\boldsymbol{Z}^*_m) > T(\boldsymbol{Z})]$

Simulation and Estimation

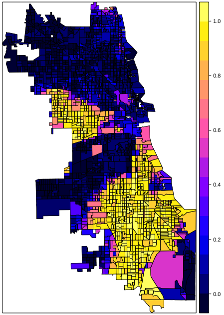
- SPINGARCH(0,1) 95% credible intervals: $\alpha \in (-0.24, 0.1)$, $\sigma^2 \in (0.46, 0.59)$, $\zeta \in (0.486, 0.492)$, and $\eta \in (0.64, 0.67)$
- SPINGARCH(1,0) 95% credible intervals: $\alpha \in (-0.54, -0.2)$, $\sigma^2 \in (0.96, 1.2)$, $\zeta \in (0.47, 0.48)$, and $\kappa \in (0.65, 0.67)$

	SPINGARCH(1,0)	SPINGARCH(0,1)
p_1 - Moran's I	.05	.46
p_2 - Var to Mean	.99	.65
p_3 - Lag 1 Corr	.45	.60

Burglaries in South Side of Chicago

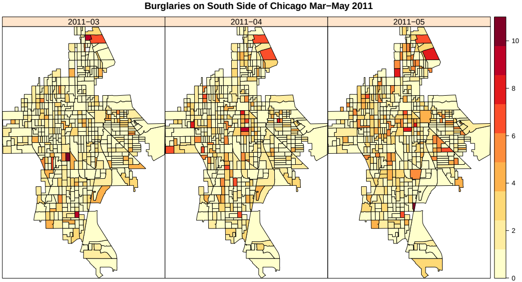


Aggregated Burglaries



Racial Segregation

Burglaries South Side of Chicago



- Crime data from city of Chicago
- 72 months (2010-2015), 552 locations (Census block groups)
- Demographic data from Census bureau

SPINGARCH(1,1) Model

$$
\begin{aligned}
Z(s_i, t) &\sim \text{Pois}(\lambda(s_i, t)) \\
\boldsymbol{\lambda_t} &= \exp(\boldsymbol{Y_t} + \boldsymbol{U}) + \eta \boldsymbol{Z_{t-1}} + \kappa \boldsymbol{\lambda_{t-1}} \\
\boldsymbol{Y_t} &\sim \text{Gau}(\boldsymbol{0}, \sigma^2_{ind} \boldsymbol{I}_{n_d, n_d}) \\
\boldsymbol{U} &\sim \text{Gau}(\alpha, \sigma^2_{sp} [(\boldsymbol{N} - \boldsymbol{C})]^{-1})
\end{aligned} \tag{34}
$$

- WCAR specification (Weighted conditional variance of CAR)
- $\boldsymbol{N}$ is diag$(|N_i|)$, $\implies \zeta \in (-1, 1)$, fix at $\zeta = 0.999$

$$
\begin{aligned}
\alpha_{s_i} = \exp \big(&\beta_0 + \beta_{pop} \log(\text{Pop}_{s_i}) \\
&+ \beta_{ym} \text{Young Men}_{s_i} + \beta_{wealth} \text{Wealth}_{s_i} + \beta_{unemp} \text{Unemp}_{s_i} \big)
\end{aligned} \tag{35}
$$

Impacts of Including Spatial Correlation

Parameter	SPINGARCH(1,1)	INGARCH(1,1)
β_0	(-3.3,-1.0)	(-4.2,-3.4)
β_{pop}	(0.11,0.34)	(0.33,0.46)
β_{ym}	$(-0.75, 0.17)$	$(0.06, 0.09)$
β_{wealth}	$(0.05, 0.16)$	$(-0.04, 0.01)$
β_{unemp}	(0.006,0.07)	(0.002,0.03)
η	$(0.04, 0.07)$	$(0.22, 0.24)$
κ	(0.31,0.39)	(0.44,0.48)
σ^2_{sp}	(0.40,0.54)	-
σ^2_{ind}	(0.40,0.47)	-

Model Assessment - Posterior Predictive Checks

	SPINGARCH(1,1)	INGARCH(1,1)
p_1 - Moran's I Statistic	0.43	0
p_2 - Variance to Mean Ratio	0.62	0
p_3 - Lag 1 Auto Correlation	0.67	0.83

- SPINGARCH(1,1) - observed maximum (p=.67), number of zeros (p=.49)

Summary

- INGARCH(1,1) process unable to replicate second order properties of burglaries in Chicago, SPINGARCH(1,1) much more so
- Exogeneous covariates offer some structure for crime, but rarely, if ever, adequately account for all
- Failure to account for small scale spatial structure leads to differing conclusions - possible policy implications

Future Work

- Impacts of aggregation
- Laplace approximations greatly speed up SPINGARCH(0,1) - Can extend to SPINGARCH(1,1)?
- Reaction Diffusion Self-Exciting Model from (Clark & Dixon, 2018) does not fit in framework (temporally correlated errors)
  - RDSEM captures reaction diffusion process of Short in Latent Process
- Dropping self-excitement leads to SPDE with exact solution - Sparse approximation?

Conclusion

- SPINGARCH model has potential to model phenomena where there is expected data correlation and limited spatial correlation
- Data model correlation manifests differently then latent correlation structure and should be accounted for accordingly
- Although derived from crime and violence, potential use for suicides, weather, etc

Thank you for your time!

Post Talk Slides

- Research Chapter 1 - SCSEM and RDSEM Models applied to Iraq Data, Laplace Approximation based exploration of posterior density of parameters
- Research Chapter 2 - Extend SCSEM and put in context of other statistical models, apply to Burglaries in Chicago
- Research Chapter 3 - Discovered bias in Laplace approximation for subset of parameter space. Explains why bias occurs and how to fix (in some instances)

Ch. 3 - An Extended Laplace Approximation Technique for Bayesian Inference

Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise **The Future of Data Analysis.** Annals of Mathematical Statistics 33: 1-67 John W. Tukey 1962

SPINGARCH (0,1) - Spatially Correlated Self-Exciting Model

Previously used in inference on violence in Iraq

$$Z(s_i, t) \sim \text{Pois}(\lambda(s_i, t)) \tag{36}$$

$$E[Z(s_i, t)] = \lambda(s_i, t) \tag{37}$$

$$\boldsymbol{\lambda_t} = \exp(\boldsymbol{Y_t}) + \eta \boldsymbol{Z_{t-1}} \tag{38}$$

$$\boldsymbol{Y_t} \sim \text{Gau}(\boldsymbol{\alpha_t}, (I_{n_d, n_d} - \boldsymbol{C})^{-1} \sigma_{sp}^2). \tag{39}$$

Integration Free Technique

Laplace Approximation to marginals

$$\tilde{\pi}(\eta, \zeta, \sigma^2_{sp}, \alpha | \boldsymbol{Z}) \propto \frac{\pi(\boldsymbol{Z}|\eta, \boldsymbol{Y})\pi(\boldsymbol{Y}|\alpha, \zeta, \sigma^2_{sp})\pi(\zeta)\pi(\alpha)\pi(\sigma^2_{sp})\pi(\zeta)}{\pi_G(\boldsymbol{Y}|\alpha, \eta, \zeta, \sigma^2_{sp}, \boldsymbol{Z})}, \quad (40)$$

$\pi_G(\boldsymbol{Y}|\alpha, \eta, \zeta, \sigma^2_{sp}, \boldsymbol{Z})$ is Gaussian approximation from Taylor series expansion of full conditional

$$\pi_G(\boldsymbol{Y}|\eta, \zeta, \sigma^2_{sp}, \boldsymbol{Z}) \propto (2\pi)^{n/2} \det(\Sigma(\theta))^{1/2} \exp\Bigg(-\frac{1}{2}(\boldsymbol{Y})^t \Sigma^{-1}(\theta)\boldsymbol{Y}$$

$$+ \sum_{s_i, t} f(\mu(s_i, t))(Y(s_i, t)) + 1/2k(\mu(s_i, t))(Y(s_i, t))^2\Bigg) \quad (41)$$

Issues

Evaluate $\pi_G(\boldsymbol{Y}|\eta, \zeta, \sigma^2_{sp}, \boldsymbol{Z})$ at mode $\boldsymbol{Y} = \boldsymbol{\mu}^*$:

$$\log(\pi(\boldsymbol{\theta}|\boldsymbol{Z})) \propto \log \pi(\boldsymbol{Z}|\eta, \boldsymbol{\mu}^*) + \log \pi(\boldsymbol{\mu}^*|\zeta, \sigma^2_{sp}, \alpha) - 1/2 \log |\Sigma(\boldsymbol{\theta})| + \log \pi(\boldsymbol{\theta}) + 1/2 \log |\Sigma(\boldsymbol{\theta}) + \text{diag } k(\mu(s_i, t)^*)| \tag{42}$$

- Even for small T, $\pi(\theta|\boldsymbol{Z})$ appears to be Gaussian
- $\pi_G(\boldsymbol{Y}|\alpha, \eta, \zeta, \sigma^2_{sp}, \boldsymbol{Z})$, not a good approximation
- σ^2_{sp} in particular appears biased (low)
- Bias increases as η or σ^2_{sp} increases

Likelihood

Full likelihood has latent dimensionality that increases as n increases:

$$L(\eta, \alpha, \zeta, \sigma^2_{sp} | \boldsymbol{Z}) \propto \int_{\boldsymbol{\Omega}_y} \prod_{i=1}^{n} \prod_{t=1}^{T} \exp(-\eta Z(s_i, t-1) - \exp(Y(s_i, t))) \times (\eta Z(s_i, t-1) + \exp(Y(s_i, t)))^{Z(s_i,t)} \, d\mu_{\boldsymbol{Y}}. \quad (43)$$

Given data, $\boldsymbol{Z}$:

$$L(\eta, \alpha, \zeta, \sigma^2_{sp} | \boldsymbol{Z}) \propto \prod_{t=1}^{T} \int_{\boldsymbol{\Omega}_{y_t}} \prod_{i=1}^{n} \exp(-\eta Z(s_i, t-1) - \exp(Y(s_i, t))) \times (\eta Z(s_i, t-1) + \exp(Y(s_i, t)))^{Z(s_i,t)} \, d\mu_{\boldsymbol{Y_t}} \quad (44)$$

Still high dimensional, intractable

Standard Laplace Approximation

$$\int_{\Omega_{y_t}} \prod_{i=1}^{n} \exp(-\eta Z(s_i, t-1) - \exp(Y(s_i, t)))$$
$$\times (\eta Z(s_i, t-1) + \exp(Y(s_i, t)))^{Z(s_i, t)} \, d\mu_{\boldsymbol{Y_t}}$$

Let

$$g(\boldsymbol{Y_t}) = \log \pi(\boldsymbol{Y_t} | \boldsymbol{\theta}) + \log(\boldsymbol{Z_t} | \boldsymbol{\theta}, \boldsymbol{Z_{t-1}}) \tag{45}$$
$$\frac{\partial g}{\partial Y(s_i, t)} = g_i \tag{46}$$
$$\frac{\partial g}{\partial Y(s_i, t) \partial Y(s_j, t)} = g_{ij} \tag{47}$$
$$g_{\boldsymbol{YY}} = \text{Hessian matrix of } g \tag{48}$$
$$g^{ij} = (i, j)\text{th element of inverse Hessian} \tag{49}$$
$$\hat{g}^{ij} \text{ evaluated at mode of } \pi(\boldsymbol{Y}|.) \tag{50}$$

Standard Laplace

$$M_t = \int \exp(-g(\boldsymbol{Y}_t))dY_t$$

- Multivariate Taylor Series expansion of g about unique minimum followed by Taylor Series expansion of exp about zero yields

$$\boldsymbol{M} = \exp(-\hat{g})\left|\frac{\hat{g}_{\boldsymbol{YY}}}{2\pi}\right| E\left[1 - \frac{1}{3!}\hat{g}_{i,j,k}U(s_1,t)U(s_2,t)U(s_3,t)\right.$$
$$\left. -\frac{1}{4!}\hat{g}_{i,j,k,l}U(s_1,t)U(s_2,t)U(s_3,t)U(s_4,t) - \cdots\right] \tag{51}$$

- $\boldsymbol{U} \sim \text{Gau}(\boldsymbol{0}, \hat{g}_{\boldsymbol{YY}})$
- $\hat{g}_{i,j,k,l} = 0$ unless $i = j = k = l$, $E[U(s_1,t)^4] = 3(g^{ii})^2$

First Three Truncated Terms

$$\frac{1}{4!}\hat{g}_{i,j,k,l}E[U(s_1,t)U(s_2,t)U(s_3,t)U(s_4,t)] = \frac{1}{8}\sum_i \hat{g}_{iiii}(\hat{g}^{ii})^2 \tag{52}$$

$$= -\frac{1}{72}\sum_{i,j\leq i}\hat{g}_{iii}\hat{g}_{jjj}\left(6\left(\hat{g}^{ij}\right)^3 + 9\hat{g}^{ii}\hat{g}^{jj}\hat{g}^{ij}\right) \tag{53}$$

$$= \frac{1}{48}\sum_i \hat{g}_{iiiiii}(\hat{g}^{ii})^4 \tag{54}$$

Issues when $\hat{g}^{ii} > 1$
Terms increase as η and σ^2_{sp} increase

Extended Laplace Approximation

(Shun & McCullagh, 1995) and (Evangelou et al. 2011)

$$M_t = \int \exp(-g(\boldsymbol{Y}_t))dY_t$$

- Taylor series expansion of $g(.)$, $\boldsymbol{M_t} = \exp(-\hat{g})E\exp(-\hat{g}_i Y(s_i) - \hat{g}_{ij} Y(s_i)Y(s_j)/2! - \hat{g}_{ijk} Y(s_i)Y(s_j)Y(s_k)/3! - \cdots)$
- $\log M_t$ is joint cumulant-generating function of $Y(s_i)$, $Y(s_i)Y(s_j)$, $Y(s_i)Y(s_j)Y(s_k)$, etc.
- Using Shun & McCullagh (1995) for $\log M$

$$\log M \propto -\hat{g} - \frac{1}{2}|\hat{g}_{\boldsymbol{YY}}| - \sum_t \sum_i \frac{1}{8}\hat{g}_{iiii}(\hat{g}^{ii})^2 - \sum_t \sum_i \frac{1}{48}\hat{g}_{iiiiii}(\hat{g}^{ii})^4 + \frac{1}{72}\sum_t \sum_{i,j \leq i} \hat{g}_{iii}\hat{g}_{jjj}\left(6\left(\hat{g}^{ij}\right)^3 + 9\hat{g}^{ii}\hat{g}^{jj}\hat{g}^{ij}\right)$$

$$\tilde{\pi}(\eta, \zeta, \sigma_{sp}^2, \alpha | \boldsymbol{Z}) \propto \frac{\pi(\boldsymbol{Z}|\eta, \boldsymbol{Y})\pi(\boldsymbol{Y}|\alpha, \zeta, \sigma_{sp}^2)\pi(\zeta)\pi(\alpha)\pi(\sigma_{sp}^2)\pi(\zeta)}{\pi_E(\boldsymbol{Y}|\alpha, \eta, \zeta, \sigma_{sp}^2, \boldsymbol{Z})}, \quad (55)$$

Algorithm 1 pseudocode for the calculation of marginal density of $\boldsymbol{\theta}$

1: Fix $\boldsymbol{\theta} = \boldsymbol{\theta_m}$ (Near Moment Based Estimates of $\boldsymbol{\theta}$)
2: Repeatedly solve $\left(\Sigma^{-1}(\boldsymbol{\theta}) + \text{diag } g^{ii}(\boldsymbol{\mu}_n)\right) \mu_{n+1} = f(\boldsymbol{\mu}_n)$, yielding $\boldsymbol{\mu}^*$ mode of $\pi_G(\boldsymbol{Y}|\boldsymbol{\theta})$
3: Evaluate $g_{iii}, g_{iiii}, g_{iiiiii}, g^{ii}, g^{ij}$ at $\boldsymbol{\mu}^*$ and $\log \pi(\boldsymbol{\theta})$ to approximate $\log M$
4: Numerically estimate Hessian of $\boldsymbol{\theta}$
5: Conduct Newton-Raphson update of $\boldsymbol{\theta}$
6: Repeat until convergence

When it Works

10×10 Lattice, 100 Time Points, $\zeta = .245, \alpha = 0$

	$\eta = .1$, $\sigma^2_{sp} = .4$	$\eta = .4$, $\sigma^2_{sp} = .6$
Relative Bias in LA(1)	.09	.25
Time to Fit LA(1) (min.)	2	3
Extended LA Without 6th Order	.03	.07
Extended LA With 6th Order	.03	.02
Time to Fit Extended LA	3	5
MCMC	.02	.01
Time to Fit MCMC	50-65	50-65

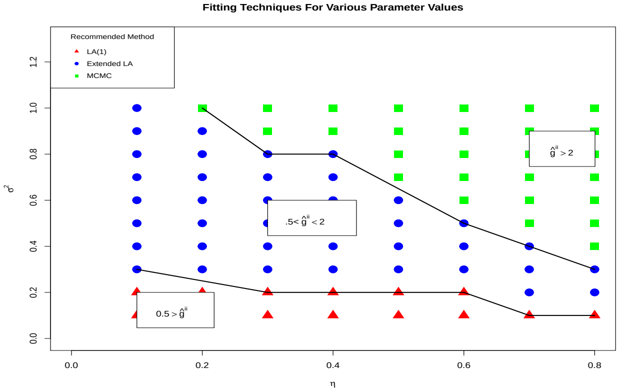
When it Doesn't

	$\eta = .7$, $\sigma^2_{sp} = 1$
Relative Bias in LA(1)	.32
Time to Fit LA(1) (min.)	5
Extended LA Without 6th Order	.14
Extended LA With 6th Order	.25
Time to Fit Extended LA	5
MCMC	.04
Time to Fit MCMC	75-244

- Examine $\frac{1}{48}\sum_i \hat{g}_{iiiiii}(\hat{g}^{ii})^4$, $\frac{1}{8}\sum_i \hat{g}_{iiii}(\hat{g}^{ii})^2$
- Higher order terms $(\hat{g}^{ii})^k$ for $k > 2$
- $\sigma^2_{sp} = 1, \eta = .7$ yield $\hat{g}^{ii} > 1$

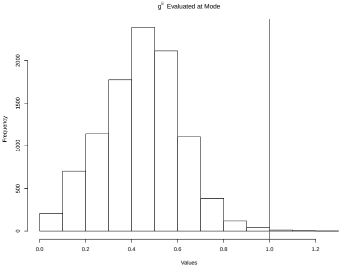
When it Works and When it Doesn't

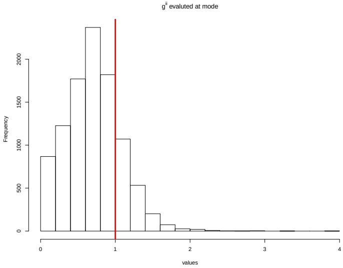
95% CI Cover Parameter	$0 < \hat{g}^{ii} < .5$	$.5 < \hat{g}^{ii} < 2$	$2 < \hat{g}^{ii}$
LA(1)	12/14	0/39	0/37
Extended LA	13/14	36/39	5/37



When it Works and When it Doesn't

$\hat{g}^{ii}$ terms for $\eta = .4, \sigma^2_{sp} = .6$ and $\eta = .7, \sigma^2_{sp} = 1$

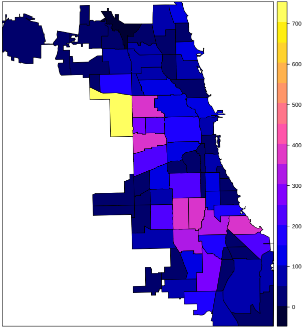




$\frac{1}{48}\sum_i \hat{g}_{iiiiii}(\hat{g}^{ii})^4 = -41$ $\qquad$ $\frac{1}{48}\sum_i \hat{g}_{iiiiii}(\hat{g}^{ii})^4 = -440$

Violent Crime in Chicago Aggregated By Neighborhood

Weekly from (December 28 2014 - January 2, 2016)



Model

SPINGARCH(0,1) or Spatially Correlated Self-Exciting model

$$Z(s_i, t) \sim \text{Pois}(\lambda(s_i, t)) \quad (56)$$
$$E[Z(s_i, t)] = \lambda(s_i, t) \quad (57)$$
$$\boldsymbol{\lambda_t} = \exp(\boldsymbol{Y_t}) + \eta \boldsymbol{Z_{t-1}} \quad (58)$$
$$\boldsymbol{Y_t} \sim \text{Gau}(\boldsymbol{\alpha_t}, (I_{n_d,n_d} - C)^{-1} M) \quad (59)$$

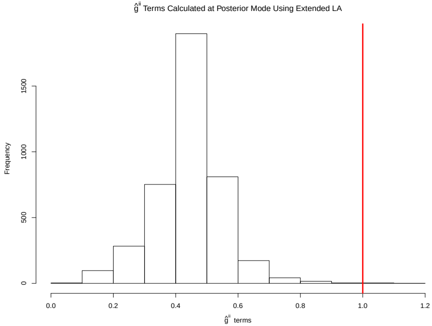
- Assume both spatial and temporal covariates
  $\alpha(s_i, t) = \beta_0 + \beta_{temp}\,\text{Temp}(t) + \beta_{pop}\,\text{Pop}(s_i)$
- Vague proper priors for all parameters

Posterior Median

Point Estimates	σ^2_{sp}	ζ	η	β_0	β_1	β_2
LA	.38	.180	.50	-5.6	.17	.50
Extended LA	.52	.179	.50	-5.6	.18	.49
MCMC	.50	.179	.50	-5.6	.18	.49

- Extended LA fit without 6th order term
- Stan - 3 chains, 15000 samples, no evidence of non-convergence, 3 hours run in parallel
- Extended LA/LA - 10 minutes

$\hat{g}^{ii}$ Terms at Posterior Median



Comparison of 95% Credible Intervals

	σ^2_{sp}	ζ	η
Extended LA	(.43,.61)	(.176,.182)	(.47,.53
MCMC	(.42,.59)	(.176,.182)	(.47,.53)

	β_0	β_1	β_2
Extended LA	(-6.3,-4.9)	(.09,.27)	(.42,.55)
MCMC	(-6.3,-5.0)	(.09,.27)	(.42,.56)

Summary

- Essential to examine magnitude of $\hat{g}^{ii}$ terms, if large, LA will have non-negligible issues
- Extending Laplace approximation removes bias over wide range of parameter space and credible intervals comparable to MCMC
- Pay price in front end coding and derivations

Dissertation Contributions

- Models for violence that are consistent with sociological theories on how violence spreads
- New class of models that extends existing INGARCH models to spatial-temporal problems and accurately capture beliefs on how violence and crime spreads
- Methodology for inference that is quick and relatively accurate

Remaining Gaps

- Extended LA for SPINGARCH(1,1)
- Theory for RDSEM models (INGARCH with spatio-temporally correlated latent structure)
- Impacts of aggregation (SPDE Approach)

Achievements

- Chapter 1 - Best Paper Competition Winner, Accepted in AoAS
- Omar Bradley Fellowship
- Graduate College Emerging Leader
- Invited Speaker - NSA
- Contributed Talk - JSM
- Invited Revision for Hidden Markov Model behavior of Mountain Goats
- Current manuscript on multivariate HMM