A Hierarchical Point Process Model for Storm Cells

W. John Braun (with Alisha Albert-Green and Charmaine Dean)

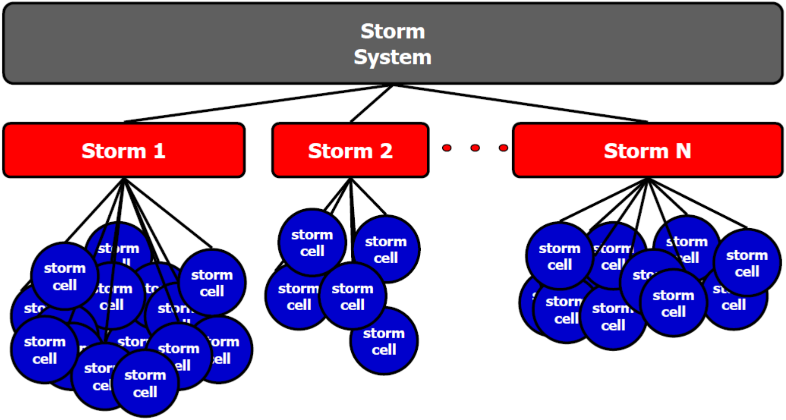
September 29, 2018

Introduction

- Severe thunderstorms hit Manitoba, Canada on September 5, 1996:
  - brought down electricity transmission line towers
  - wind and tension from failed towers caused cascading tower failure
  - electricity supply to networks in North Dakota was interrupted for the two weeks following
- Manitoba Hydro subsequently funded a project for the modelling and prediction of the failure of transmission lines caused by high-intensity winds
- **Objective:** Model storm cells and characterize storms and storm systems

Storm Cells

- **Storm cell**: smallest unit of a storm producing system



(Mohee & Miller, 2010)

Storm Cell Data (Bismarck, North Dakota)

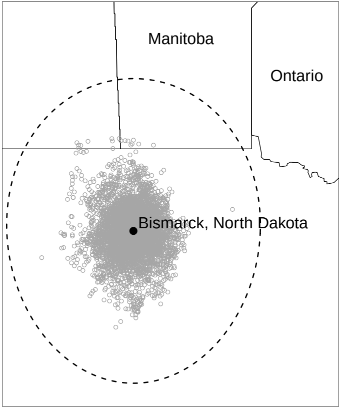
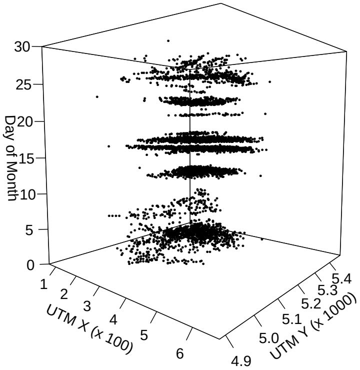


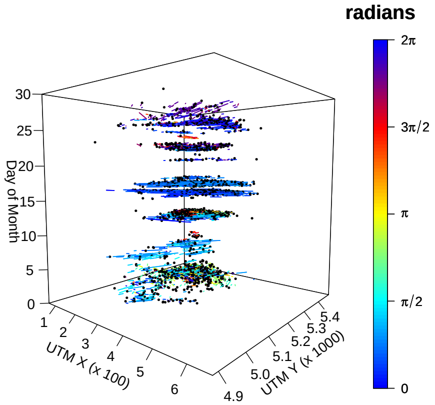
Figure: May 2003 storm cells.

(ernest borgnine was here)

Storm Cells



(a) Detected storm cells.



(b) Storm cell trajectories.

1. Log-Gaussian Cox Process (LGCP)

Unobserved Gaussian Process:

- $\mathcal{Z} = \left\{ Z(u) : u \in \mathbb{R}^d \right\}$ is a Gaussian process with:

$$\begin{aligned} E[Z(u)] &= \mu \\ \text{Cov}[Z(u_1), Z(u_2)] &= \sigma^2 \rho(u_2 - u_1; \phi) \end{aligned}$$

{Observed Events:}

- $\mathcal{X} = \{X_i, i = 1, 2, \ldots, N\}$ observations in $A \subset \mathbb{R}^d$ are conditionally an inhomogeneous Poisson process (IPP) with:

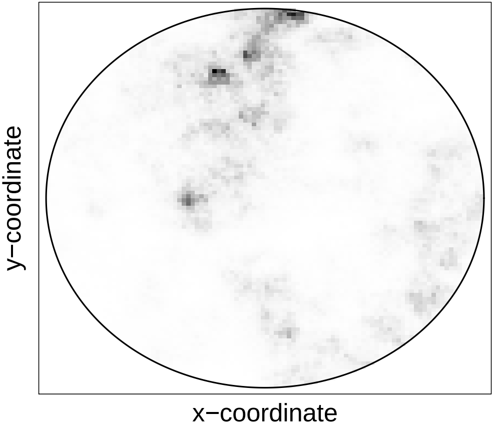
$$\{X_i, i = 1, 2, \ldots, N\} \mid Z(u) \sim \text{IPP}[\Lambda(u)]$$

where:

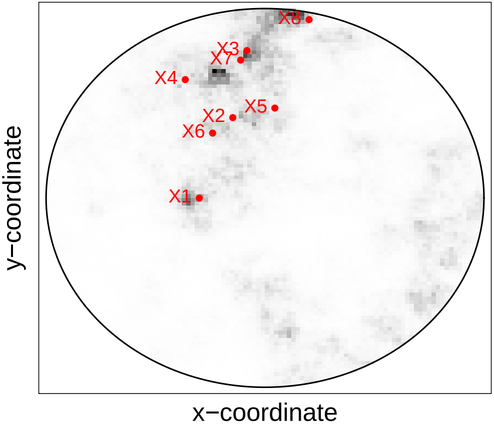
$$\Lambda(u) = \exp\{Z(u)\}$$

(Møller et al., 1998)

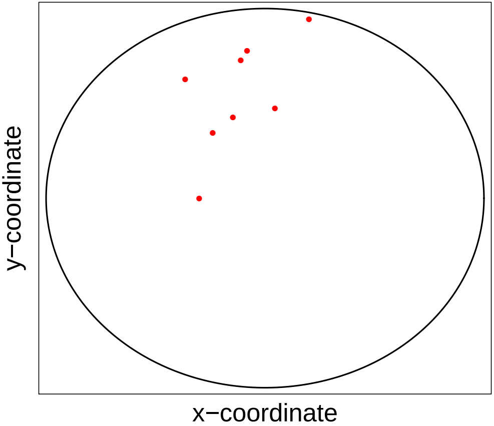
1. Log-Gaussian Cox Process (LGCP)



1. Log-Gaussian Cox Process (LGCP)



1. Log-Gaussian Cox Process (LGCP)



2. Neyman-Scott Process

Unobserved Parent Process:

- $\mathcal{P} = \{P_j,\ j = 1, 2, \ldots\} \in \mathbb{R}^d$ is the parent process

$$\{P_j,\ j = 1, 2, \ldots\} \sim \text{HPP}(\lambda_p)$$

{Observed Offspring Process:}

- $\mathcal{X} = \{X_{ij},\ i = 1, 2, \ldots, N_j\} \subset A$ is the offspring process

$$\{X_{ij},\ i = 1, \ldots, N_j\} \mid \{P_j,\ j = 1, 2, \ldots\} \quad \sim \quad \text{IPP}[\Lambda_{X|P}(x)]$$

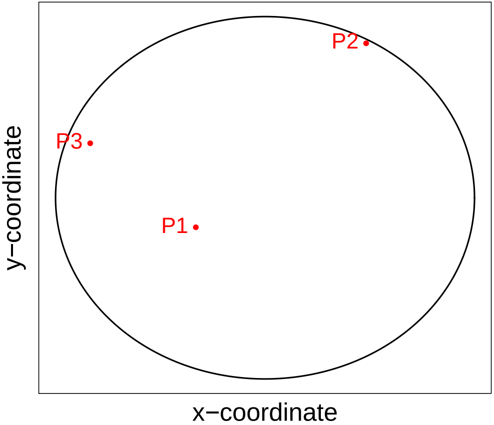
where:

$$\Lambda_{X|P}(x) \quad = \quad \sum_j \alpha f(x - P_j) \text{ and}$$

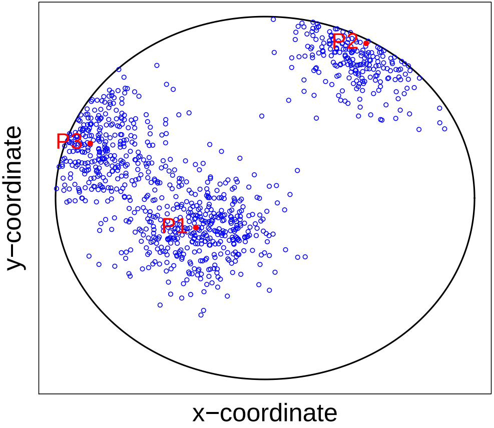
- α is the average number of offspring per parent
- $f(\cdot)$ is distribution of the displacement from offspring to the parent

(Møller, 2003)

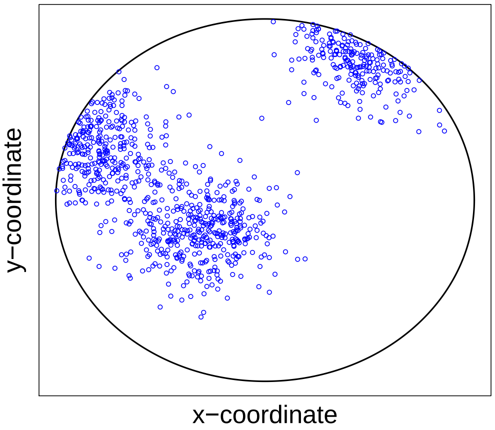
2. Neyman-Scott Process



2. Neyman-Scott Process



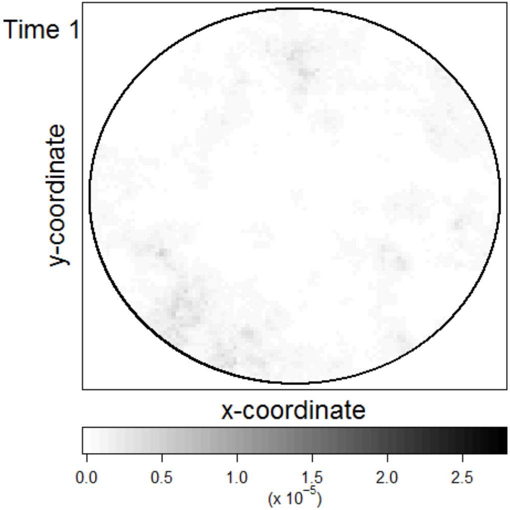
2. Neyman-Scott Process



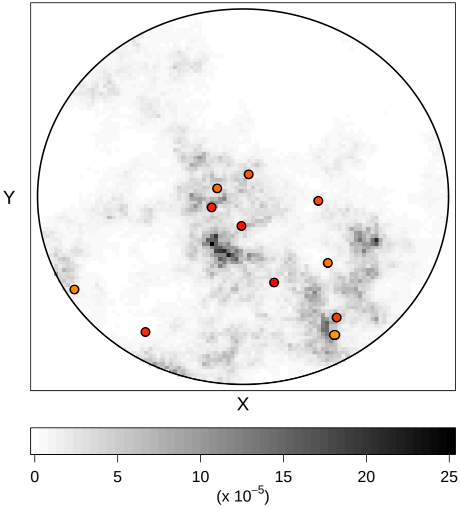
2. Neyman-Scott Process

Generalize Neyman-Scott process by allowing the parent process to follow an LGCP

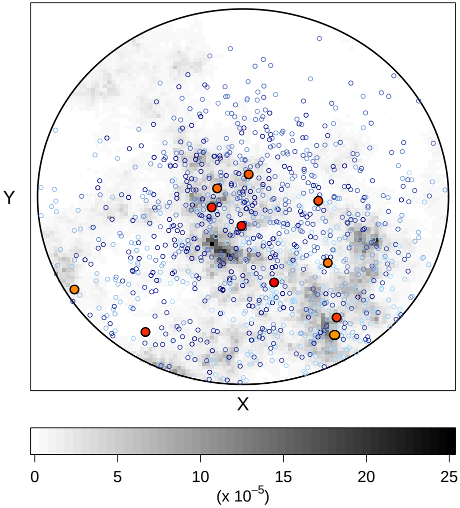
Conceptual Model Illustration



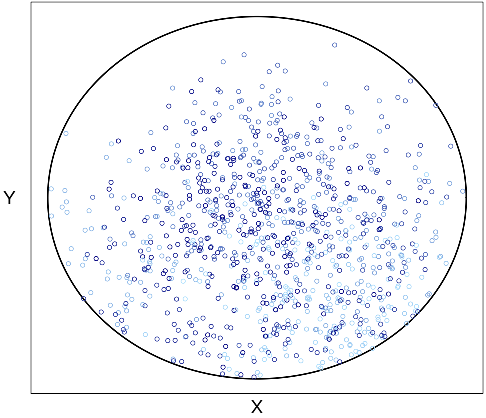
Conceptual Model Illustration



Conceptual Model Illustration



Conceptual Model Illustration



Hierarchical Cluster Process: Parent Process

For the jth storm with centre at $(u, v) \in \mathbb{R}^2 \times \mathbb{R}$:

$$\{P_j, j = 1, \ldots\} \mid Z(u, v) \quad \sim \quad \mathrm{IPP}[\Lambda_P(u, v)]$$

where:

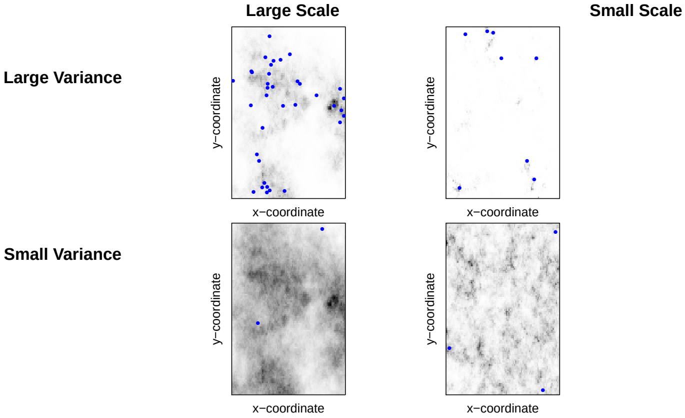
$$\begin{aligned} \Lambda_P(u, v) &= \exp\{\beta + Z(u, v)\} \\ E[Z(u, v)] &= -0.5(\sigma_s^2 + \sigma_t^2) \\ \mathrm{Cov}[Z(u_1, v_1), Z(u_2, v_2)] &= \sigma_s^2 e^{-(||u_2 - u_1||/\phi_s)} + \sigma_t^2 e^{-(|v_2 - v_1|/\phi_t)} \end{aligned}$$

and:

- β is an intercept parameter
- σ_s^2 and σ_t^2 are spatial and temporal variance parameters
- ϕ_s and ϕ_t are spatial and temporal scale parameters

Hierarchical Cluster Process: Parent Process

$$\Lambda_P(u,v) = \exp\{\beta + Z(u,v)\}$$
$$\mathrm{Cov}[Z(u_1,v_1), Z(u_2,v_2)] = \sigma_s^2 e^{-(||u_2-u_1||/\phi_s)} + \sigma_t^2 e^{-(|v_2-v_1|/\phi_t)}$$



Hierarchical Cluster Process: Offspring Process

For the ith storm cell from the jth storm at $x = (s, t) \in S \times T$:

$$\{X_{ij}, i = 1, \ldots, N_j\} \mid \{P_j, j = 1, \ldots\} \quad \sim \quad \text{IPP}[\Lambda_{X|P}(x)]$$

where

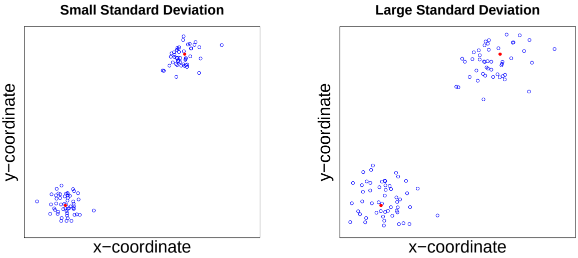
$$\Lambda_{X|P}(x) \quad = \quad \sum_j \alpha f_s\left(s - u_j; \omega_s^2\right) f_t\left(t - v_j; \omega_t^2\right)$$

and:

- α is the average number of storm cells per storm
- $f_s(\cdot)$ is the spatial displacement between the storm cell and the storm centre, $N_2(0, \omega_s^2 I_2)$
- $f_t(\cdot)$ is the temporal displacement between the storm cell and the storm centre, $N(0, \omega_t^2)$

Hierarchical Cluster Process: Offspring Process

$$\Lambda_{X|P}(x) \quad = \quad \sum_j \alpha f_s\left(s - u_j; \omega_s^2\right) f_t\left(t - v_j; \omega_t^2\right)$$



Intensities for Hierarchical Cluster Process

First-Order Intensity:

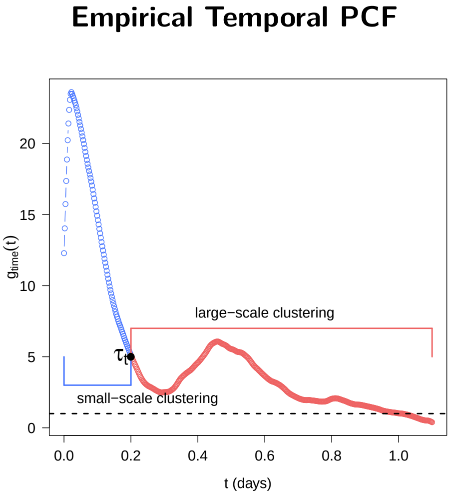
$$\begin{aligned}\lambda(x) &= E[\Lambda_{X|P}(x)] \\ &= \alpha \exp(\beta) \\ &= \lambda\end{aligned}$$

Pair Correlation Function (PCF):

- Probability of observing a pair of points separated by a distance of $x_2 - x_1$ relative to what you would expect from a Poisson process

$$\begin{aligned}g(x_2 - x_1) &= \frac{\lambda^{(2)}(x_1, x_2)}{\lambda(x_1)\lambda(x_2)} \\ &= \frac{f * f(x_2 - x_1)}{\exp(\beta)} + \\ &\quad \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \exp\left\{\sigma^2 \rho(u_2 - u_1)\right\} f(x_1 - u_1) f(x_2 - u_2) \mathrm{d}u_1 \mathrm{d}u_2\end{aligned}$$

PCF in Hierarchical Cluster Processes



Parameter Estimation via Minimum Contrast

First-Order Intensity:

$$\hat{\lambda} = \frac{N}{|S \times T|}$$

Clustering Parameters:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \int_0^r \left(g(u)^{1/4} - \hat{g}(u)^{1/4} \right)^2 \mathrm{d}u$$

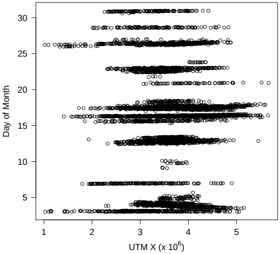
where:

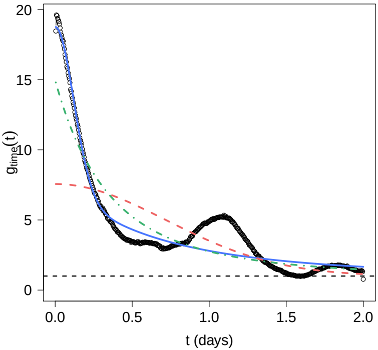
- $\boldsymbol{\theta}$ represents the clustering parameters
- $\hat{g}(u)$ is an empirical estimate of the PCF

Projection Processes (Prokešová and Dvořák, 2014):

- $\mathcal{X}_s = \{s : (s, t) \in \mathcal{X} \cap (S \times T)\}$
- $\mathcal{X}_t = \{t : (s, t) \in \mathcal{X} \cap (S \times T)\}$

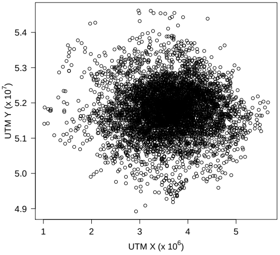
Results: Temporal Projection Process

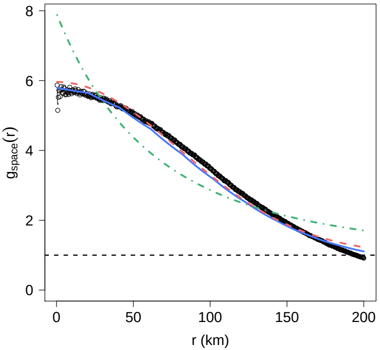




Right figure: Empirical temporal PCF (points) and fitted PCFs for the hierarchical cluster process (lines), the Neyman-Scott cluster process (dashed lines) and the LGCP (dotted-dashed lines).

Results: Spatial Projection Process





Right figure: Empirical spatial PCF (points) and fitted PCFs for the hierarchical cluster process (lines), the Neyman-Scott cluster process (dashed lines) and the LGCP (dotted-dashed lines).

Marked Point Processes

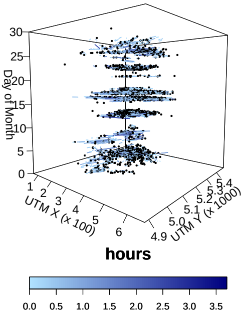
- **Mark:** random variable associated with a point process event (e.g. magnitude in earthquake models)

- Model joint distribution of $\mathcal{X}$ (point process) and $\mathcal{M}$ (mark process):

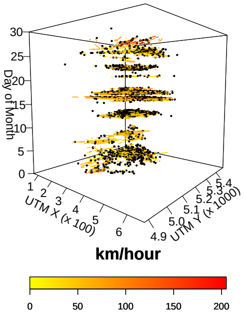
$$[\mathcal{X}, \mathcal{M}]$$

- Modelling Framework: $[\mathcal{X}, \mathcal{M}] = [\mathcal{X}][\mathcal{M} \mid \mathcal{X}]$

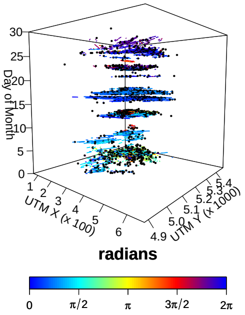
Mark Process for Storm Cell Trajectories



(a) Duration.

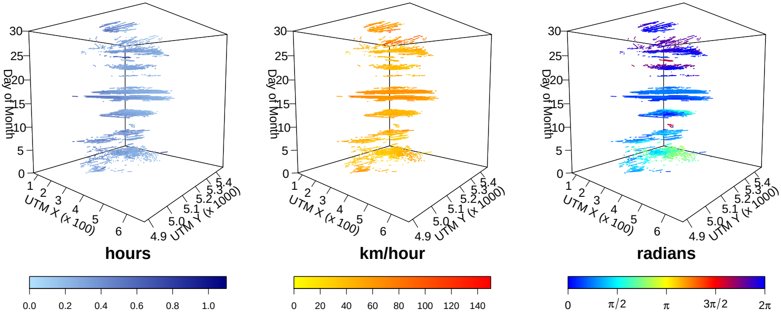


(b) Speed.



(c) Direction.

Results: Mark Process for Storm Cell Trajectories



(a) Duration. (b) Speed. (c) Direction.

Ad Hoc SemiParametric Procedure

Projection/Minimum Contrast Method Fitting Procedure is Limited:

- Very slow
- Uses too much memory
- Identifiability of parameter estimates is compromised by projection

Simple and Quick Alternative:

- Apply mean-shift clustering algorithm (Fukinaga and Hostetler, 1975) to find storm centers
- Apply LGCP directly to storm centers

In order for this to work, we need to ensure that the mean-shift algorithm can accurately find the cluster centers.

Data Sharpening (Choi and Hall, 1999)

- Introduced as a method to reduce bias in density estimation.
- Given raw data $x_1, x_2, \ldots, x_n$ in R^d with unknown density $f(x)$ one can use a symmetric pdf as a kernel with scale parameter h (the bandwidth). to estimate $f(x)$.
- By performing local constant regression of x on x, one obtains sharpened data, on which the kernel density estimate will have reduced bias.
- This method can be iterated.

Data Sharpening

- Choi and Hall (1999) advocated a few iterations of their algorithm in order to attain a sizeable bias reduction. At each iteration, the sharpened data move closer to local modes.
- The iteration converges (Braun and Woolford, 2007):

Theorem 1: For fixed h and for any initial vector of observations $\mathbf{x_0}$, the data sharpening algorithm of Choi and Hall (1999) converges to a unique vector $\mathbf{x}^*$.

- And this is equivalent to mean-shift clustering.

A Simulated Data Set

```
##      JULIAN            UTM.X             UTM.Y
##  Min.   : -4.991   Min.   :-387.75   Min.   :-330.7
##  1st Qu.: 10.488   1st Qu.:  84.64   1st Qu.: 190.5
##  Median : 87.893   Median : 251.46   Median : 317.8
##  Mean   : 73.150   Mean   : 256.71   Mean   : 308.8
##  3rd Qu.:135.279   3rd Qu.: 432.56   3rd Qu.: 432.2
##  Max.   :154.922   Max.   : 936.04   Max.   : 930.8
```

This is somewhat like the 2003 storm cell data set. The true values of ω are 60, 60 and 0.05. We expect 120 storms.

Estimated Cluster Centers

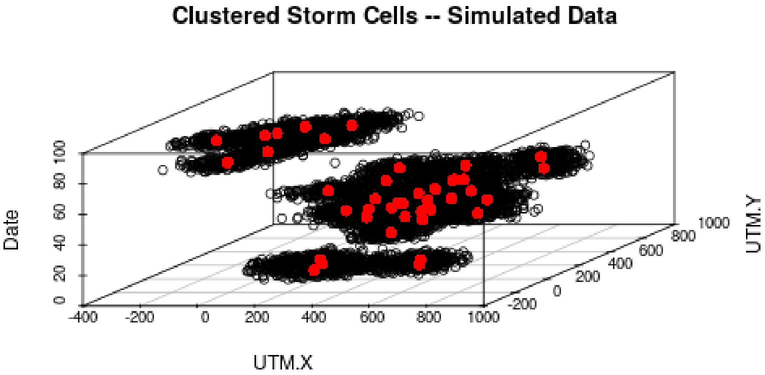
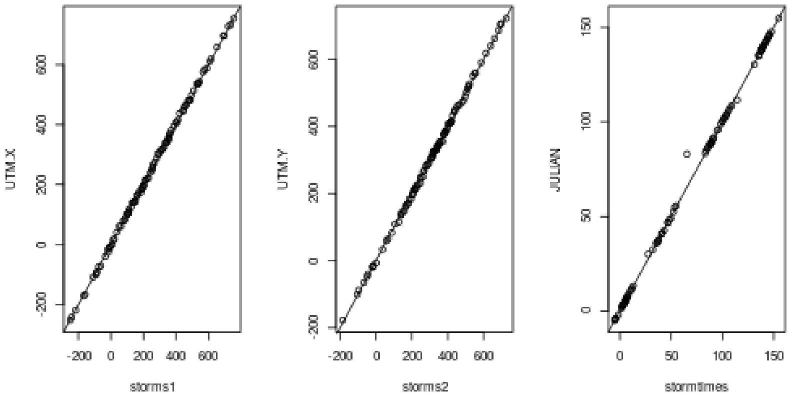


Figure 1: Black: simulated storm cells; Red: estimated cluster centers.

Comparison of Estimated Cluster Centers with Truth




```
## [1] "Estimated No. of Clusters: 107"

## [1] "Actual No. of Clusters: 110"

## [1] "Number of Storm Cells: 22091"
```

Cluster Parameter Estimates

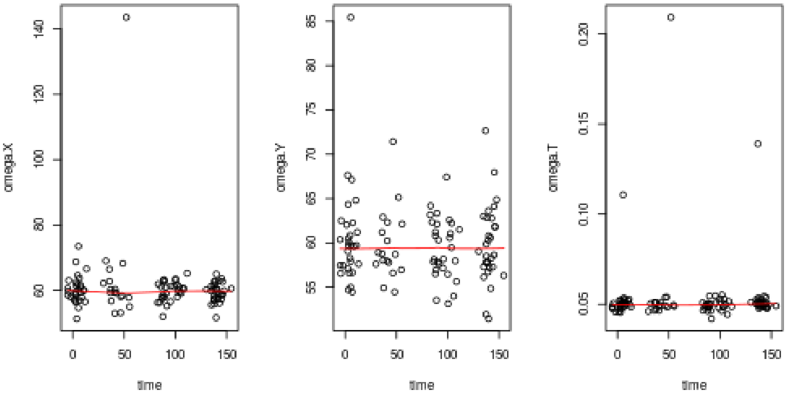


Figure 2: Estimates of cluster scale parameters.

Estimated Cluster Sizes

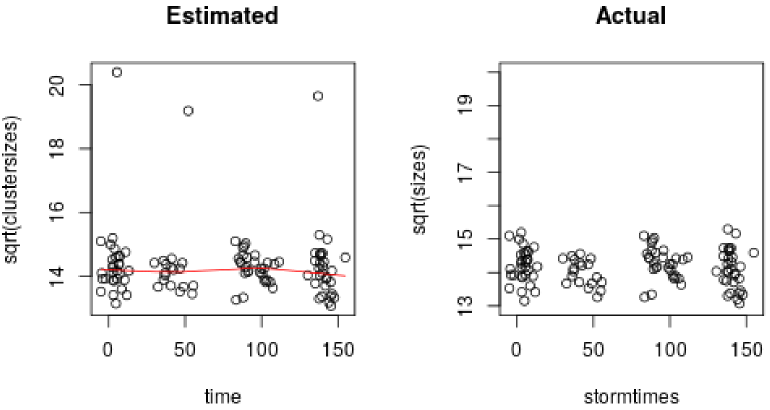


Figure 3: Estimated cluster sizes are almsot, but not always, accurate.

What about the 2003 North Dakota data?

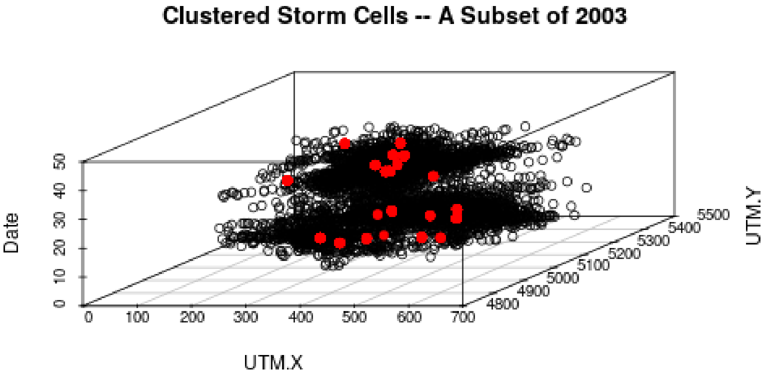


Figure 4: Black: storm cells; Red: estimated cluster centers.

Chararacteristics of Estimated Clusters

```
## [1] "Estimated Number of Clusters: 114"

## [1] "Number of Storm Cells: 29583"
```

Cluster Parameter Estimates - 2003

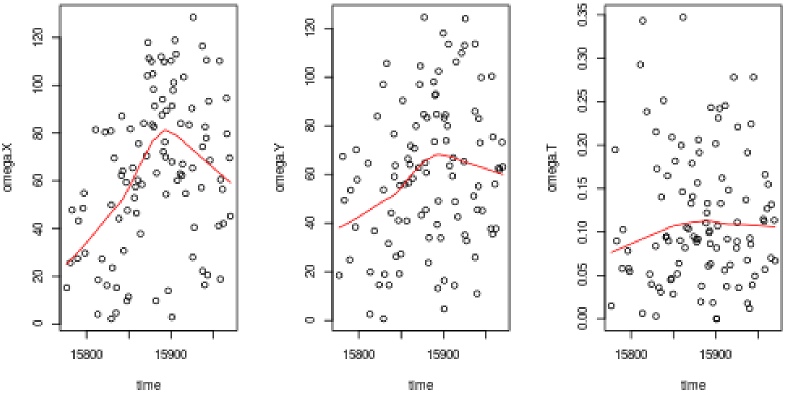


Figure 5: Cluster scale parameter estimates.

Estimated Cluster Sizes - 2003

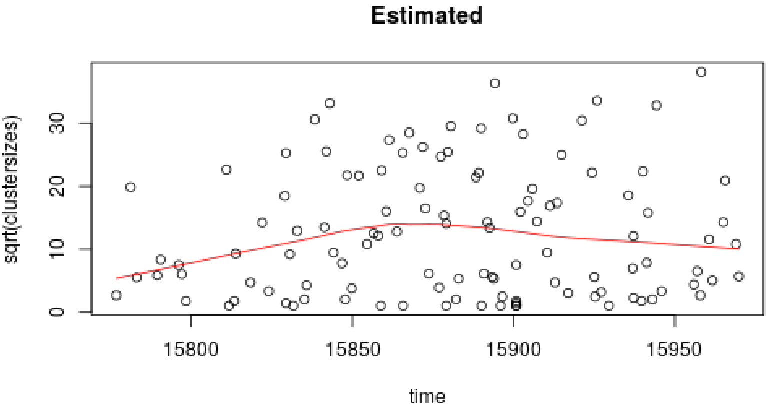
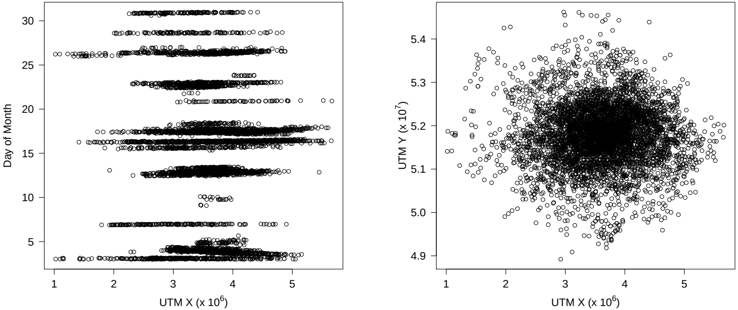


Figure 6: Cluster size etimates.

Discussion

- Ad hoc procedure looks promising, but not perfect, since the cluster centers are reasonably well predicted
- Clustering method is quick and simple
- When employing LGCP, measurement error should be incorporated
- Clustering method can also be used as a diagnostic check on the other method; e.g. it has provided evidence that observed differences in the clustering mechanism over the season are real.
- Clustering method indicates that parameter identifiability issue for other method might be induced by use of projection.
- Clustering method could be used to provide good starting guesses for the other method.

Future Work



1. Composite likelihood approach to parameter estimation of a spatio-temporal point process (Guan, 2006)
2. Joint estimation for spatio-temporal point processes with evolving marks (Renshaw & Sarkka, 2001 and Sarkka & Renshaw, 2006)

Acknowledgements

- Patrick Brown (St. Michael's Hospital, University of Toronto)
- Valerie Isham (University College London)
- Reg Kulperger (University of Western Ontario)
- Financial support from CANSSI and NSERC is gratefully acknowledged:

john owes patrick a pint

Sponsors