

# Model Assessment for Space-time Point Processes Using Super-thinning

Robert Alan Clements  
Rick Schoenberg  
Alejandro Veen

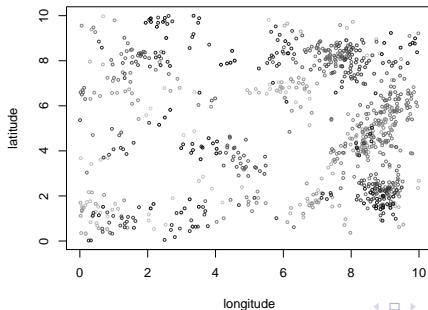
August 3, 2010

# Part I

## Background

# Point processes

- random collections of points in some metric space
- (locally)  $\sigma$ -finite: finitely many points within any bounded set
- simple: points at distinct locations and times
- examples: earthquakes, wildfires, crime locations, etc.



# The conditional intensity

- the conditional intensity  $\lambda(x, y, t)$  is defined as the frequency with which events are expected to occur in a space  $S$  around a specific point, time and mark (for marked point processes), conditional on the prior history  $\mathcal{H}_t$  (Daley and Vere-Jones, 2003)
- $\lambda$  uniquely characterizes a simple point process

$$\lim_{\Delta x, \Delta y, \Delta t \downarrow 0} \frac{E[N\{(x, x+\Delta x) \times (y, y+\Delta y) \times (t, t+\Delta t)\} | \mathcal{H}_t]}{\Delta x \Delta y \Delta t}$$

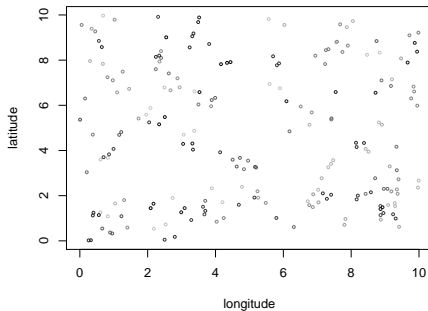
# Hawkes process

- useful for modeling earthquakes (Ogata 1988, 1998)
- cluster process
  - **background points** (immigrants) trigger **offspring points** (1st generation descendants)
  - **offspring points** trigger their own **offspring points** (2nd generation descendants)
- modeled by

$$\lambda(x, y, t) = \mu(x, y, t) + \sum_{\{i: t_i < t\}} g(x - x_i, y - y_i, t - t_i)$$

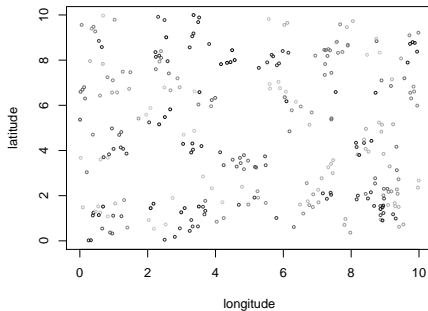
# Hawkes process

$$\lambda(x, y, t) = 0.02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha\beta}{\pi} e^{-\alpha t - \beta(x^2 + y^2)}$$



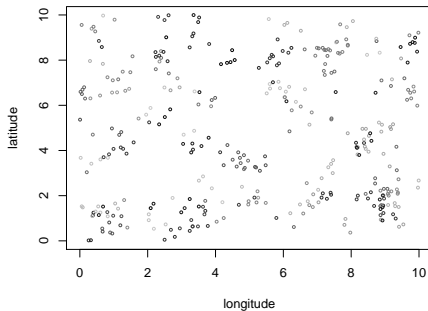
# Hawkes process

$$\lambda(x, y, t) = 0.02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha\beta}{\pi} e^{-\alpha t - \beta(x^2 + y^2)}$$



# Hawkes process

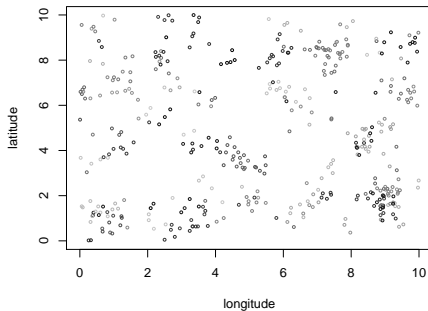
$$\lambda(x, y, t) = 0.02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha\beta}{\pi} e^{-\alpha t - \beta(x^2 + y^2)}$$





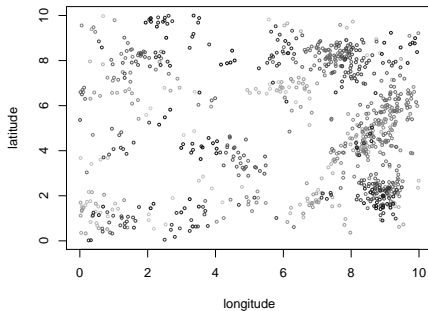
# Hawkes process

$$\lambda(x, y, t) = 0.02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha\beta}{\pi} e^{-\alpha t - \beta(x^2 + y^2)}$$



# Hawkes process

$$\lambda(x, y, t) = 0.02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha\beta}{\pi} e^{-\alpha t - \beta(x^2 + y^2)}$$



## Part II

# Residual Analysis

## Superposed residuals

- simulate points in  $S$  with probability

$$c - \hat{\lambda}(x, y, t) \text{ where } c = \sup\{\hat{\lambda}(x, y, t)\} \text{ over } S$$

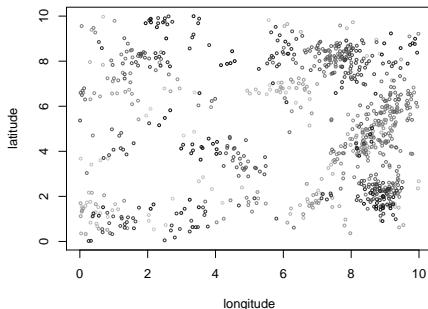
- essentially, superimpose a point process onto the existing point process, with conditional intensity

$$\hat{\lambda}_S(x, y, t) = c - \hat{\lambda}(x, y, t)$$

- the result, called *superposed* residuals, are homogeneous Poisson, with rate  $c$ , if fitted model is correct (Brémaud, 1981)

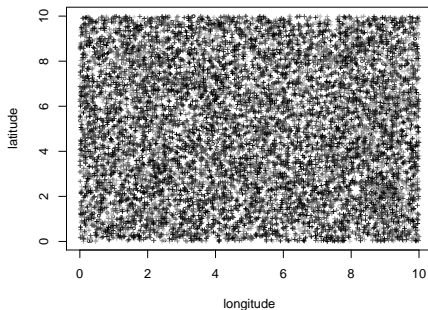
# Superposed residuals

$$\hat{\lambda}(x, y, t) = .02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha}{2\sigma^2\pi} \exp \left\{ -\alpha t - \frac{1}{2} \left( \frac{\sqrt{x^2 + y^2}}{\sigma} \right)^2 \right\}$$



# Superposed residuals

$$\hat{\lambda}(x, y, t) = .02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha}{2\sigma^2\pi} \exp \left\{ -\alpha t - \frac{1}{2} \left( \frac{\sqrt{x^2 + y^2}}{\sigma} \right)^2 \right\}$$



## Thinned residuals

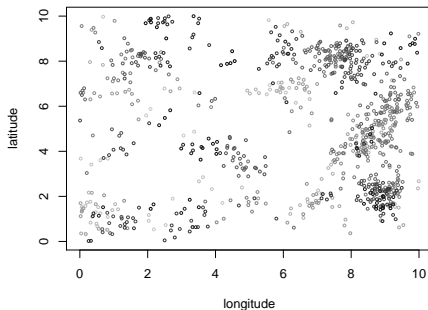
- keep each observation in  $S$  with probability

$$\frac{b}{\hat{\lambda}(x_i, y_i, t_i)} \text{ where } b = \inf\{\hat{\lambda}(x, y, t)\} \text{ over } S$$

- remaining points, called *thinned* residuals, are homogeneous Poisson, with rate  $b$ , if fitted model is correct (Schoenberg, 2003)

# Thinned residuals

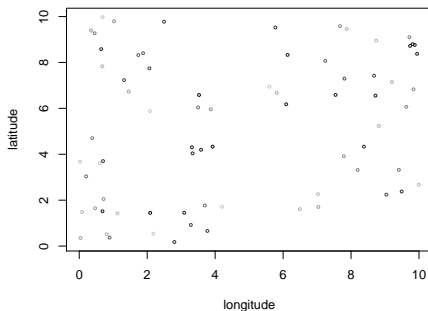
$$\hat{\lambda}(x, y, t) = .02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha}{2\sigma^2\pi} \exp \left\{ -\alpha t - \frac{1}{2} \left( \frac{\sqrt{x^2 + y^2}}{\sigma} \right)^2 \right\}$$





# Thinned residuals

$$\hat{\lambda}(x, y, t) = .02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha}{2\sigma^2\pi} \exp \left\{ -\alpha t - \frac{1}{2} \left( \frac{\sqrt{x^2 + y^2}}{\sigma} \right)^2 \right\}$$

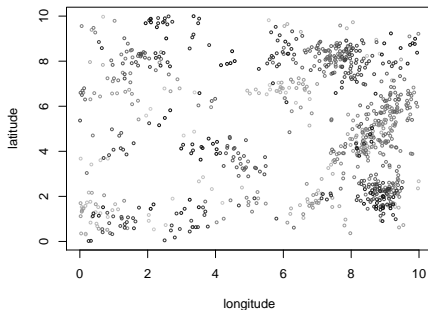


## Super-thinned residuals

- thin points if  $\hat{\lambda}(x_i, y_i, t_i) \geq k$ 
  - keep each point with probability  $\frac{k}{\hat{\lambda}(x_i, y_i, t_i)}$
- simulate points if  $\hat{\lambda}(x, y, t) < k$ 
  - simulate points with probability  $k - \hat{\lambda}(x, y, t)$
- the result, called *super-thinned* residuals, is homogeneous Poisson, with rate  $k$ , if fitted model is correct

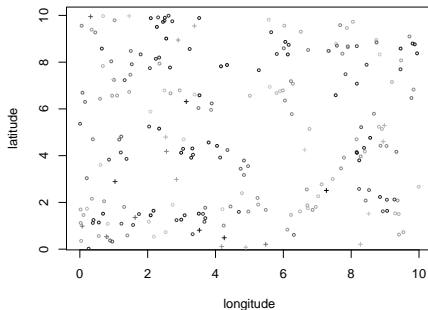
# Super-thinned residuals

$$\hat{\lambda}(x, y, t) = .02 + \sum_{\{i: t_i < t\}} K_0 \frac{\alpha}{2\sigma^2\pi} \exp \left\{ -\alpha t - \frac{1}{2} \left( \frac{\sqrt{x^2 + y^2}}{\sigma} \right)^2 \right\}$$



# Super-thinned residuals

$$k = \frac{250}{|S|} = 0.0833$$

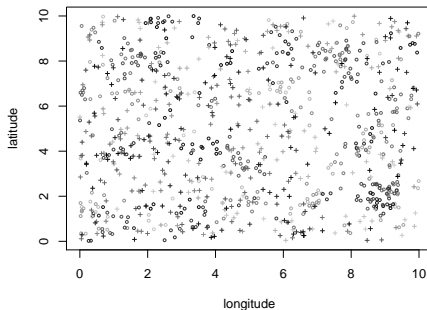


## Super-thinned residuals

- some choices of  $k$  seem to be more powerful than others
- how should we choose  $k$ ?
- common sense choices:
  - choose  $k$  such that the same number of points are thinned and superposed (the mean of  $\lambda$ )
  - choose  $k$  such that the fewest number of points are thinned and superposed (the median of  $\lambda$ )

# Super-thinned residuals

$$k = \frac{889}{|S|} = 0.2963$$



# Summary

- superposition may result in too many points, which is difficult to interpret
- thinning may result in too few points
- super-thinning is a more powerful alternative due to the tuning parameter,  $k$
- open question: which value of  $k$  is most powerful?

Thank you for attending!

Contact info:  
Robert Clements  
Department of Statistics, UCLA  
[clements@stat.ucla.edu](mailto:clements@stat.ucla.edu)