# Small Area Spatiotemporal Crime Rate Forecasting

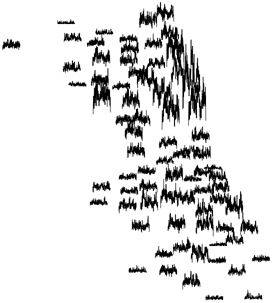
William Herlands

Carnegie Mellon University

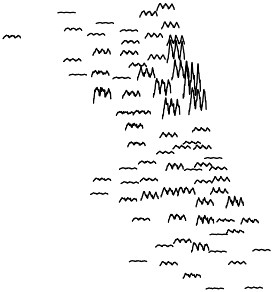
November 19, 2015

# Road map

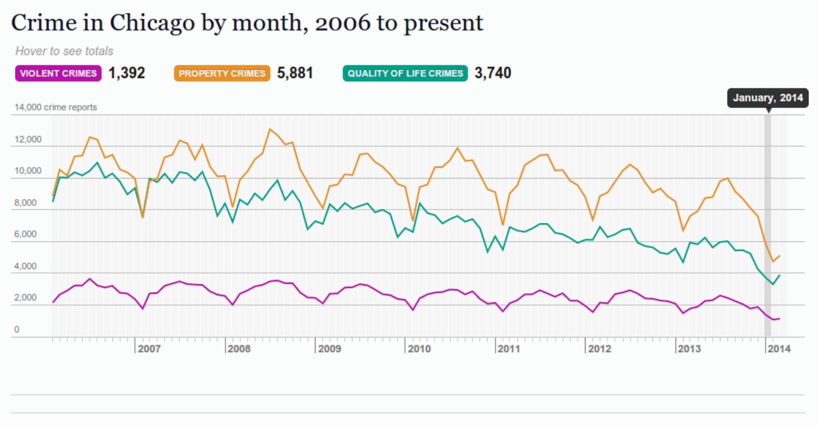
- Forecasting and prediction of crime rates
- Bayesian modeling framework
- Theft in Chicago
- Experimental results



$\Rightarrow$

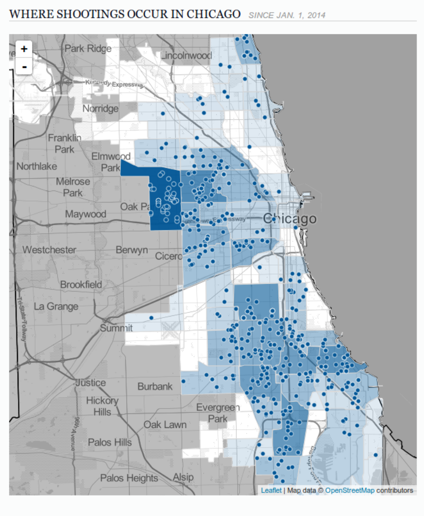


# Crime rates: temporal




source: crime.chicagotribune.com

# Crime rates: spatial



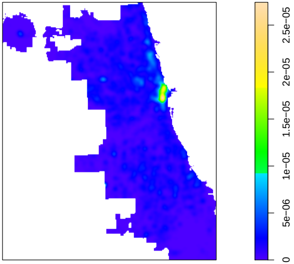

source: crime.chicagotribune.com/chicago/shootings

# Previous Work

- Univariate time series models [Gorr, Olligschlaeger, Thompson 2003]

# Previous Work

- Univariate time series models [Gorr, Olligschlaeger, Thompson 2003]
- "Heat maps" [Groff and La Vigne, 2002, many others], "risk terrain modeling" [Caplan and Kennedy, 2011]



# Our goals

- Fully probabilistic framework

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies
  - Long-term and seasonal trends

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
  - Model variance, not just mean

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
  - Model variance, not just mean
- Prediction

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
  - Model variance, not just mean
- Prediction
  - Good in-sample performance (MSE)

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
  - Model variance, not just mean
- Prediction
  - Good in-sample performance (MSE)
  - Not a black box

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
  - Model variance, not just mean
- Prediction
  - Good in-sample performance (MSE)
  - Not a black box
  - Produce heat maps, interpretable by policy makers

# Our goals

- Fully probabilistic framework
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
  - Model variance, not just mean
- Prediction
  - Good in-sample performance (MSE)
  - Not a black box
  - Produce heat maps, interpretable by policy makers
- Forecasting: better out-of-sample performance (MSE) compared to existing methods, with forecast intervals

# Modeling framework: Gaussian Processes

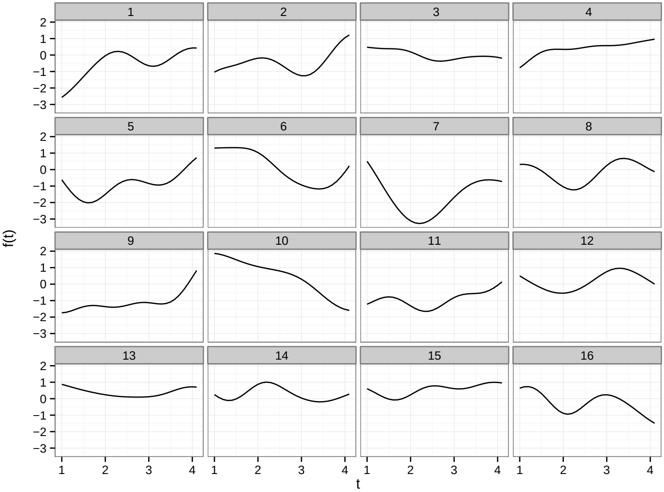
Bayesian framework for specifying priors over functions

# Modeling framework: Gaussian Processes

Bayesian framework for specifying priors over functions
$f(t) \sim \mathcal{GP}(0, k(t, t'))$

# Modeling framework: Gaussian Processes

Bayesian framework for specifying priors over functions
$f(t) \sim \mathcal{GP}(0, k(t, t'))$

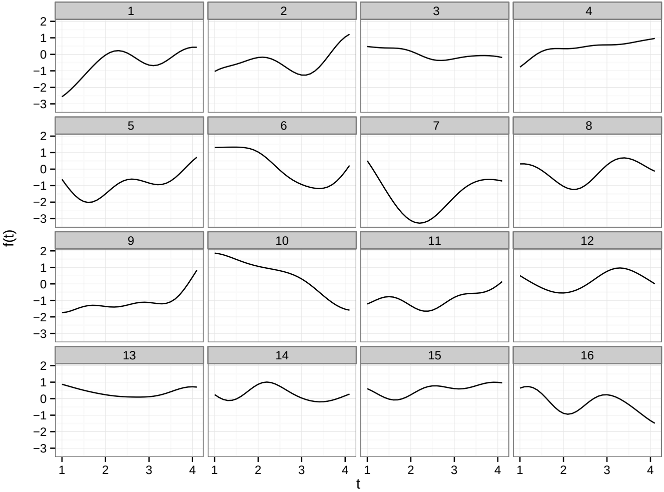


# Gaussian Processes

- Covariance function specifies smoothness of function:
$$f(t) \sim \mathcal{GP}(0, k(t, t')),\ k(t, t') = \exp(-|t - t'|^2)$$

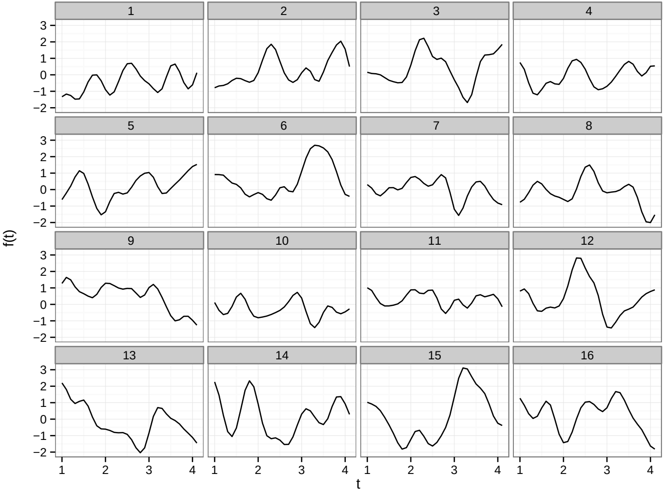
# Gaussian Processes

- Covariance function specifies smoothness of function:
$$f(t) \sim \mathcal{GP}(0, k(t, t')),\ k(t, t') = \exp(-|t - t'|^2)$$



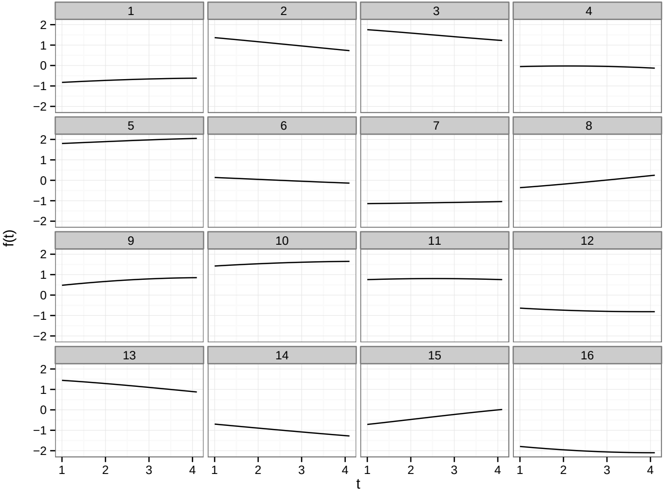
# Gaussian Processes

- Covariance function specifies smoothness of function:
$$f(t) \sim \mathcal{GP}(0, k(t,t')),\ k(t,t') = \exp(-\tfrac{1}{0.01}|t-t'|^2)$$



# Gaussian Processes

- Covariance function specifies smoothness of function:
  $$f(t) \sim \mathcal{GP}(0, k(t,t')),\ k(t,t') = \exp(-\frac{1}{10}|t-t'|^2)$$



# Gaussian Processes for Count Data

- At space-time location $(s, t)$:

$$n_{s,t} \sim \text{Poisson}(\lambda(s,t))$$

# Gaussian Processes for Count Data

- At space-time location $(s, t)$:

$$n_{s,t} \sim \mathrm{Poisson}(\lambda(s, t))$$

- Rate $\lambda$ varies in space and time:

$$\lambda(s, t) = \exp(f(s, t))$$

# Gaussian Processes for Count Data

- At space-time location $(s, t)$:

$$n_{s,t} \sim \text{Poisson}(\lambda(s, t))$$

- Rate $\lambda$ varies in space and time:

$$\lambda(s, t) = \exp(f(s, t))$$

- Place a GP prior on the log-intensity:

$$f(s, t) \sim \mathcal{GP}(0, K)$$

# Recap and Preview: Gaussian Processes

- Fully Bayesian framework: uncertainty intervals for all parameters, predictions, and forecasts

# Recap and Preview: Gaussian Processes

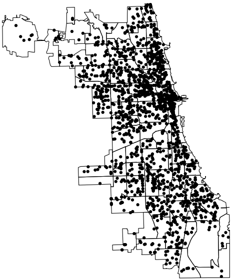
- Fully Bayesian framework: uncertainty intervals for all parameters, predictions, and forecasts
- Flexible and interpretable models for spatial and temporal dependencies

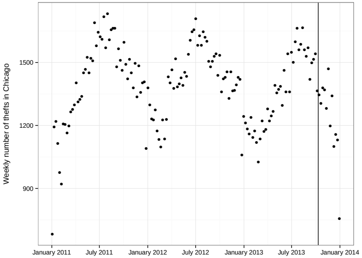
# Recap and Preview: Gaussian Processes

- Fully Bayesian framework: uncertainty intervals for all parameters, predictions, and forecasts
- Flexible and interpretable models for spatial and temporal dependencies
- Generalizes spatial approaches (heat maps) and temporal approaches (autoregressive models, periodic models)

# Application: Theft in Chicago

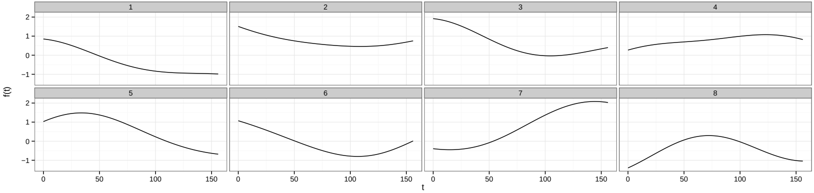
| week (t) | neighborhood (s) | # of thefts |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 2 | 7 |
| 2 | 1 | 0 |
| 2 | 2 | 3 |
| ⋮ | ⋮ | ⋮ |





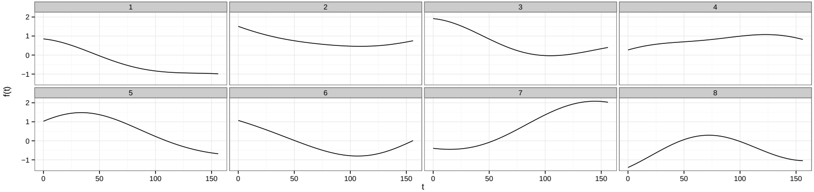
# Time Component

- $k_{\text{smooth}}(t, t') = \sigma_1^2 \exp(-\frac{1}{\ell_1^2}(t - t')^2)$

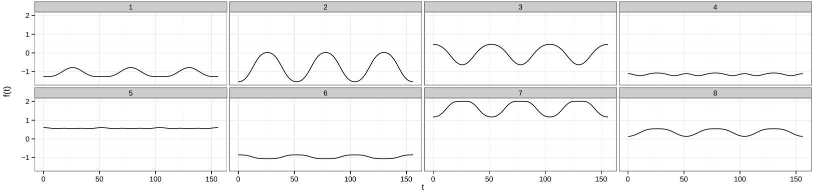


# Time Component

- $k_{\text{smooth}}(t, t') = \sigma_1^2 \exp(-\frac{1}{\ell_1^2}(t - t')^2)$

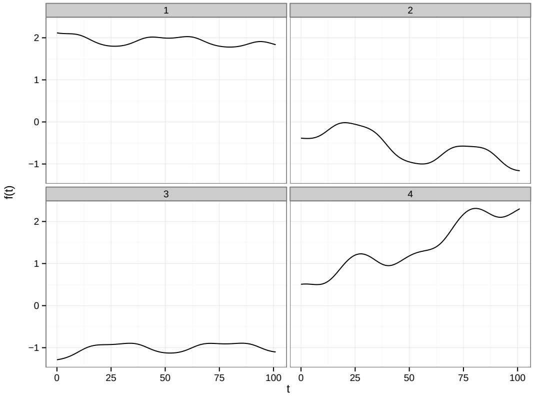


- $k_{\text{periodic}}(t, t') = \sigma_2^2 \exp\left(-\frac{1}{\ell_2^2} \sin^2\left(\frac{(t-t')\pi}{52}\right)\right)$



# Time Component

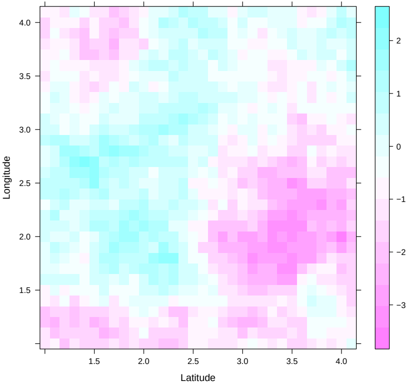
$$k(t, t') = k_{\text{smooth}}(t, t') + k_{\text{periodic}}(t, t')$$



# Spatial Component

Given locations $\{s_1 \ldots, s_n\}$, specify Matérn covariance:

$$k(s, s') = \sigma^2 \left(1 + \frac{\|s - s'\|\sqrt{3}}{\ell}\right) \exp\left(-\frac{\|s - s'\|\sqrt{3}}{\ell}\right)$$



# Theft in Chicago

Hyperparameters:

$$\ell_1, \ldots, \ell_3, \sigma_1, \ldots, \sigma_5 \sim \text{Student-t}(\nu = 4)$$

Parameters:

$$k_{\text{space}}(s, s') = \text{Matern}_{\ell_1, \sigma_1^2}(s, s')$$

$$k_{\text{time}}(s, s') = \ell_2 \exp(-\frac{1}{\sigma_2^2}\|s - s'\|^2)$$

$$k_{\text{periodic}}(s, s') = \text{Periodic}_{\ell_3, \sigma_3^2}(s, s')$$

$$k_{\text{space-periodic}}((s, t), (s', t')) = \ell_4 \exp(-\frac{1}{\sigma_4^2}\|s - s'\|^2) \cdot \text{Periodic}_{1, \sigma_5^2}(t, t')$$

Latent Risk Surface:

$$f(s, t) \sim \mathcal{GP}(0, k_{\text{space}} + k_{\text{time}} + k_{\text{periodic}} + k_{\text{space-periodic}})$$

Data:

$$n_{s,t} \sim \text{Poisson}(\exp(f(s, t)))$$

# Experiments and Results

- Fit full spatiotemporal model to week-neighborhood counts of theft from January 2011 to September 2013

# Experiments and Results

- Fit full spatiotemporal model to week-neighborhood counts of theft from January 2011 to September 2013
- Forecast October - December 2013

# Experiments and Results

- Fit full spatiotemporal model to week-neighborhood counts of theft from January 2011 to September 2013
- Forecast October - December 2013
- Perform posterior predictive checks for predictions, variances

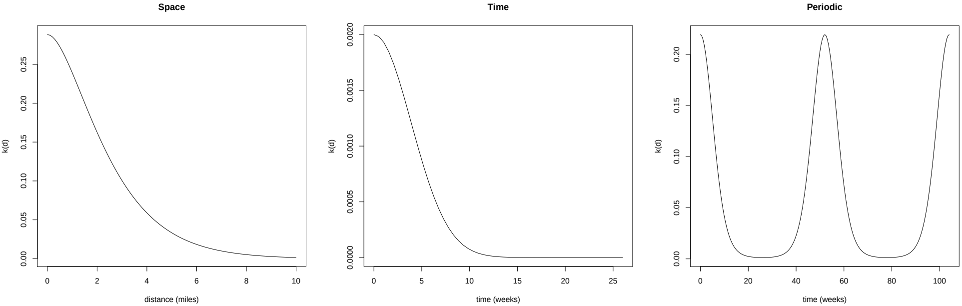
# Experiments and Results

- Fit full spatiotemporal model to week-neighborhood counts of theft from January 2011 to September 2013
- Forecast October - December 2013
- Perform posterior predictive checks for predictions, variances
- Calculate mean squared error of predictions (in-sample) and forecasts (out-of-sample)

# Experiments and Results

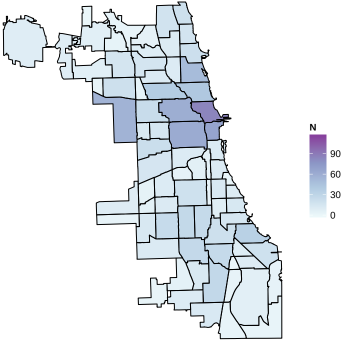
- Fit full spatiotemporal model to week-neighborhood counts of theft from January 2011 to September 2013
- Forecast October - December 2013
- Perform posterior predictive checks for predictions, variances
- Calculate mean squared error of predictions (in-sample) and forecasts (out-of-sample)
- Compare to competing methods

# Results

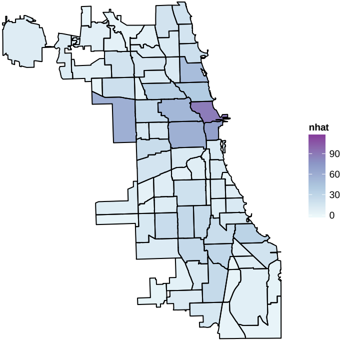

Space
k(d)
0.00
0.05
0.10
0.15
0.20
0.25
0
2
4
6
8
10
distance (miles)
Time
k(d)
0.0000
0.0005
0.0010
0.0015
0.0020
0
5
10
15
20
25
time (weeks)
Periodic
k(d)
0.00
0.05
0.10
0.15
0.20
0
20
40
60
80
100
time (weeks)

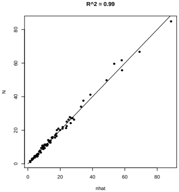
# Predictions: January - June 2011

Observed



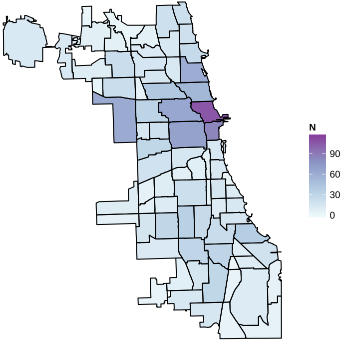

Predicted



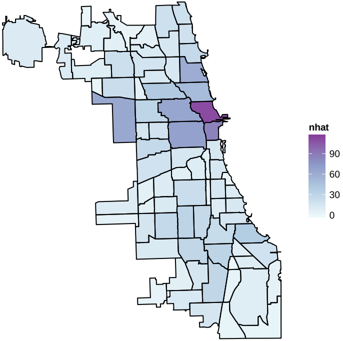



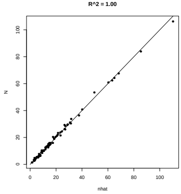
# Predictions: July - December 2011

Observed



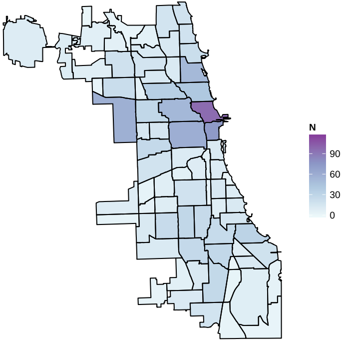

Predicted



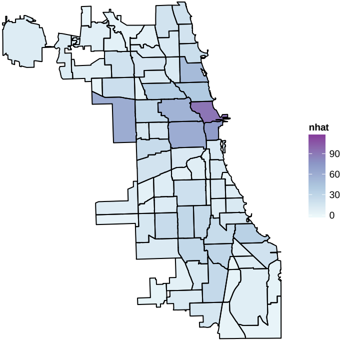



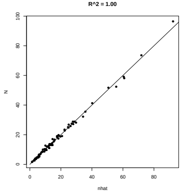
# Predictions: January - June 2012

Observed



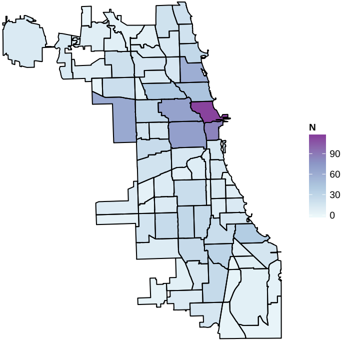

Predicted



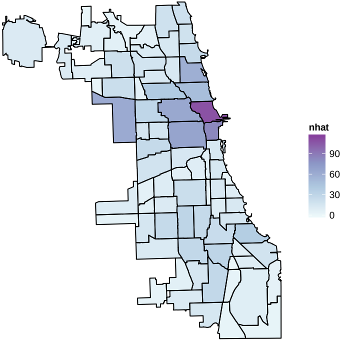



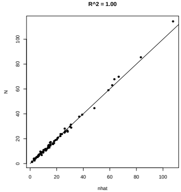
# Predictions: July - December 2012

Observed



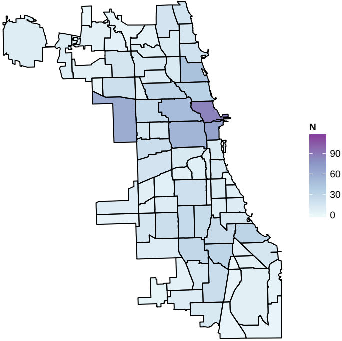

Predicted



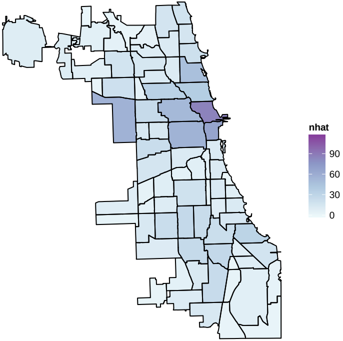



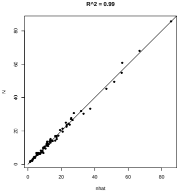
# Predictions: January - June 2013

Observed



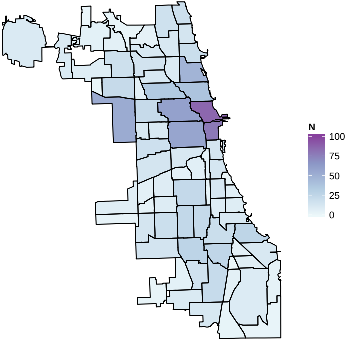

Predicted



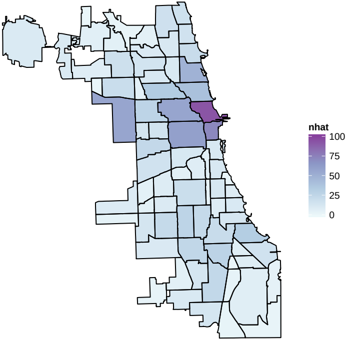



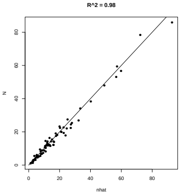
# Forecasts: October - December 2013

## Observed

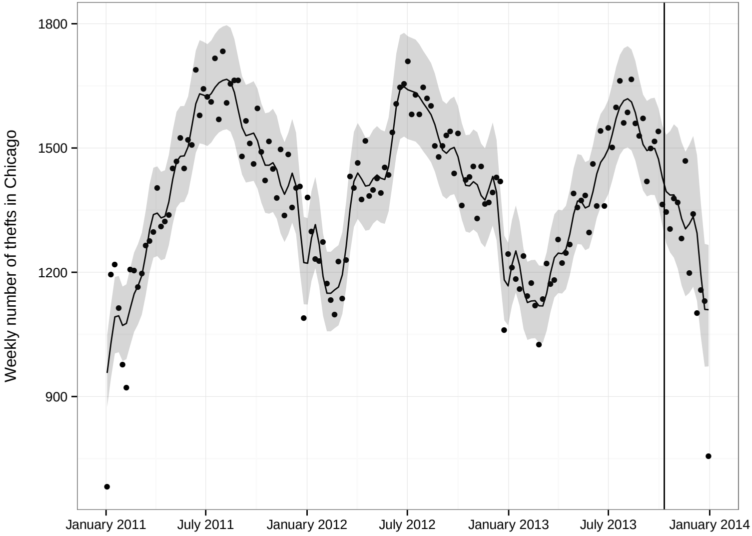



## Predicted

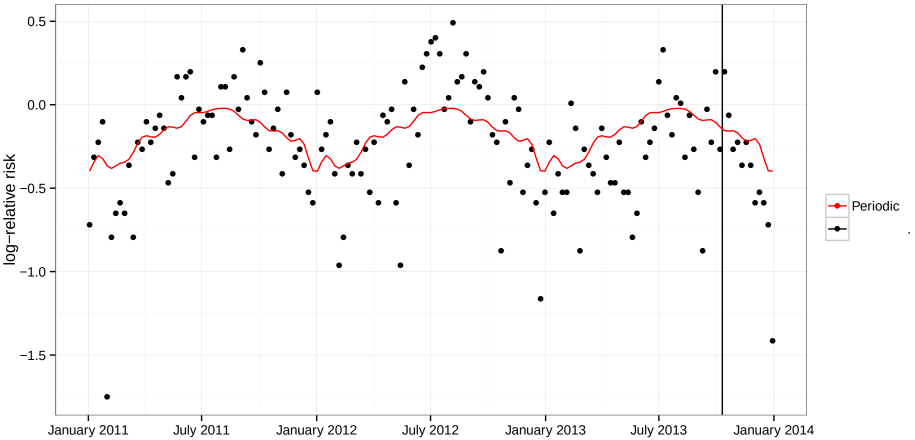


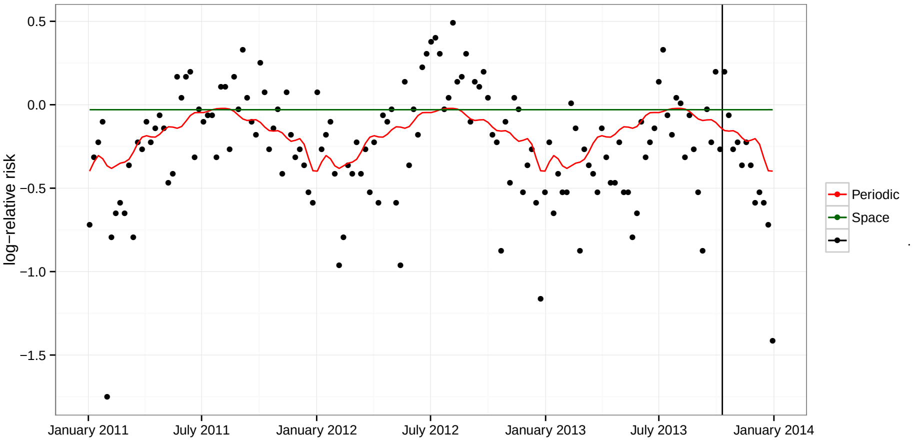
# Results: Time



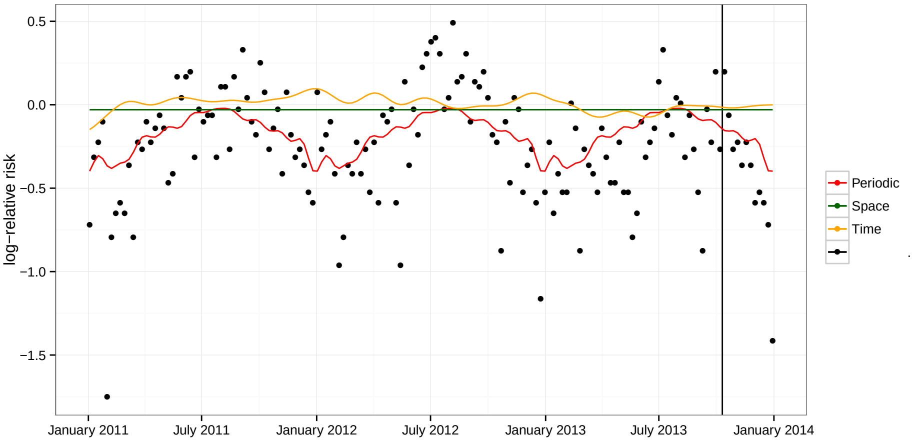
# Results: Time Decomposition



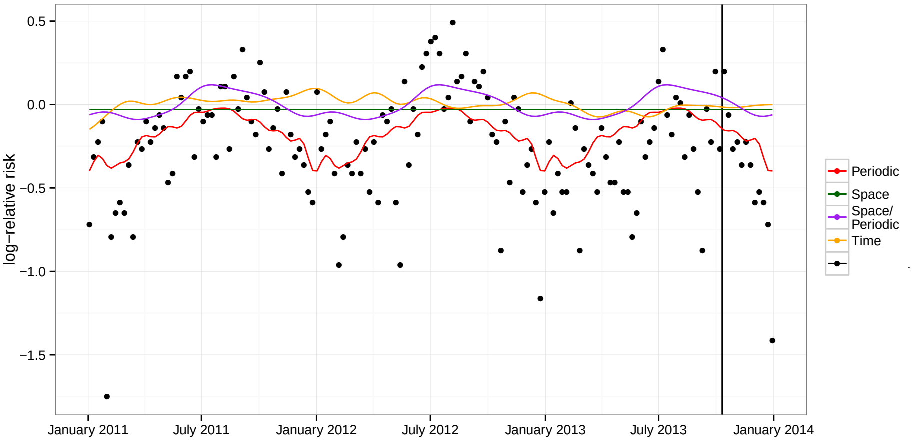
# Results: Time Decomposition



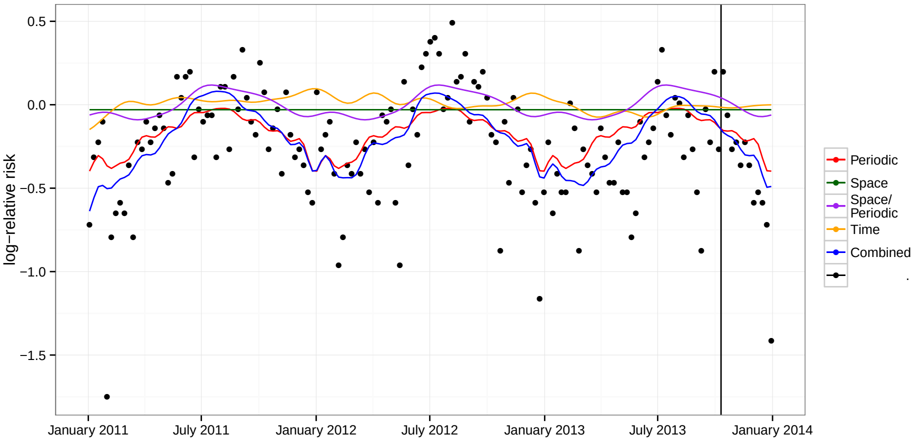
# Results: Time Decomposition



# Results: Time Decomposition



# Results: Time Decomposition



# Forecasting (Out-of-Sample) MSE

Baseline :

# Forecasting (Out-of-Sample) MSE

Baseline :

- No change: 52.19

# Forecasting (Out-of-Sample) MSE

Baseline :

- No change: 52.19

Competitors :

# Forecasting (Out-of-Sample) MSE

Baseline :

- No change: 52.19

Competitors :

- Kernel intensity estimation (heat maps): 47.70.

# Forecasting (Out-of-Sample) MSE

Baseline :

- No change: 52.19

Competitors :

- Kernel intensity estimation (heat maps): 47.70.
- AR(1): 37.98.

# Forecasting (Out-of-Sample) MSE

Baseline :

- No change: 52.19

Competitors :

- Kernel intensity estimation (heat maps): 47.70.
- AR(1): 37.98.
- Holt-Winters: 46.99.

# Forecasting (Out-of-Sample) MSE

Baseline :

- No change: 52.19

Competitors :

- Kernel intensity estimation (heat maps): 47.70.
- AR(1): 37.98.
- Holt-Winters: 46.99.

Our model: 25.81

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling
  - Spatial and temporal dependencies

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling
  - Spatial and temporal dependencies
  - Long-term and seasonal trends

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
- Prediction

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
- Prediction
  - Good in-sample performance (MSE), not a black box, interpretable output

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
- Prediction
  - Good in-sample performance (MSE), not a black box, interpretable output
- Forecasting: better out-of-sample performance (MSE) compared to existing methods

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
- Prediction
  - Good in-sample performance (MSE), not a black box, interpretable output
- Forecasting: better out-of-sample performance (MSE) compared to existing methods
- And there's more: see Flaxman et al [ICML, 2015] for a more scalable, finer-grained model fit to 8 years of data with predictions 2 **years** into the future.

# Recap and Conclusions

- Proposed a flexible GP/Poisson framework for spatiotemporal modeling
  - Spatial and temporal dependencies
  - Long-term and seasonal trends
- Prediction
  - Good in-sample performance (MSE), not a black box, interpretable output
- Forecasting: better out-of-sample performance (MSE) compared to existing methods
- And there's more: see Flaxman et al [ICML, 2015] for a more scalable, finer-grained model fit to 8 years of data with predictions 2 **years** into the future.
- Future work: continuous changepoint models, more extensive comparisons to existing methods, jointly fitting different types of crimes, more applied work to understand how well-calibrated forecasts are

# Thanks!