

MODELOS DE PREDICCIÓN DE CRIMEN

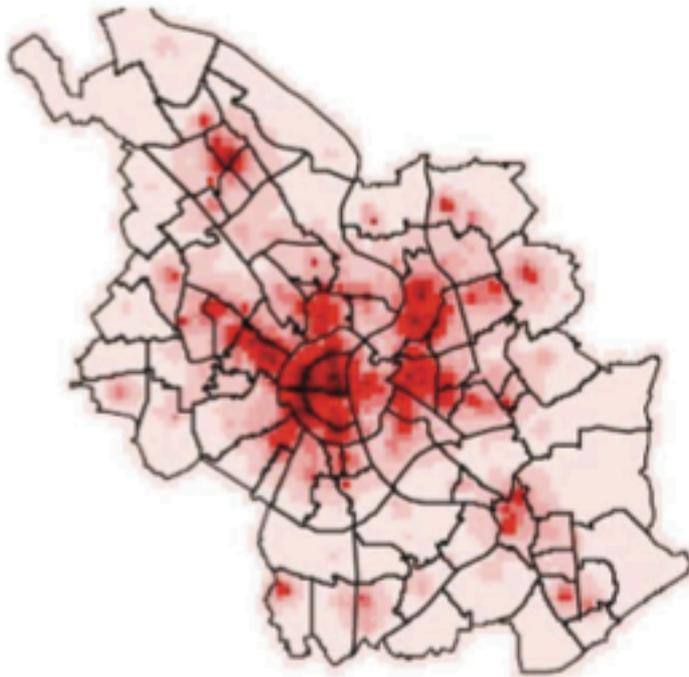
Una introducción a las metodologías de
Geographic Profiling

Geographic Profiling - Objetivos

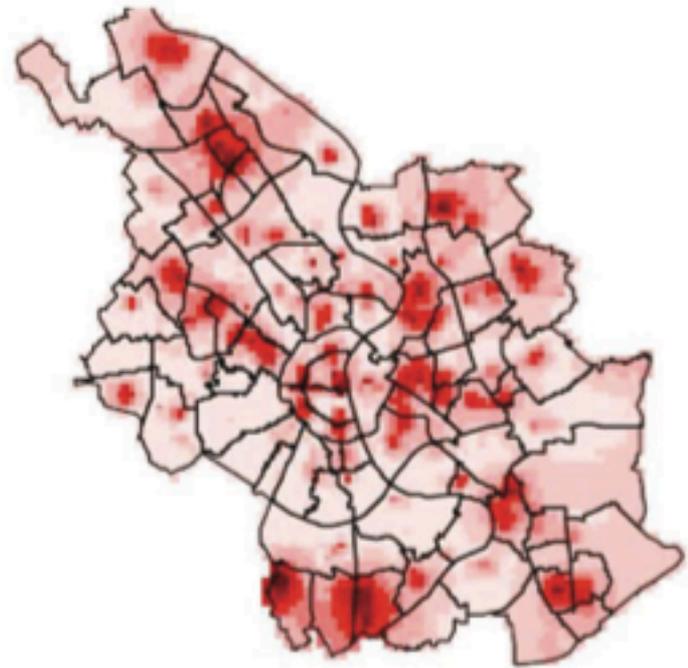
- Encontrar el lugar desde donde operan grupos criminales.
- Predecir dónde ocurrirán futuros crímenes.
- Encontrar factores geográficos que incentivan o desincentivan la criminalidad.
- Arrojar un orden de prioridades a partir del cual hacer patrullaje en una ciudad.

Identificación de HotSpots

- Covering Ellipses
- Single and Dual Kernel Estimation: Se tiene en cuenta la densidad poblacional y la población en riesgo



Single kernel density of assault and battery incidents



Dual kernel density of assault and battery incidents relative to at-risk population

Identificación de HotSpots

- Se puede integrar el tiempo tal como se mostrará en la implementación de O'Leary.

Identificación de HotSpots

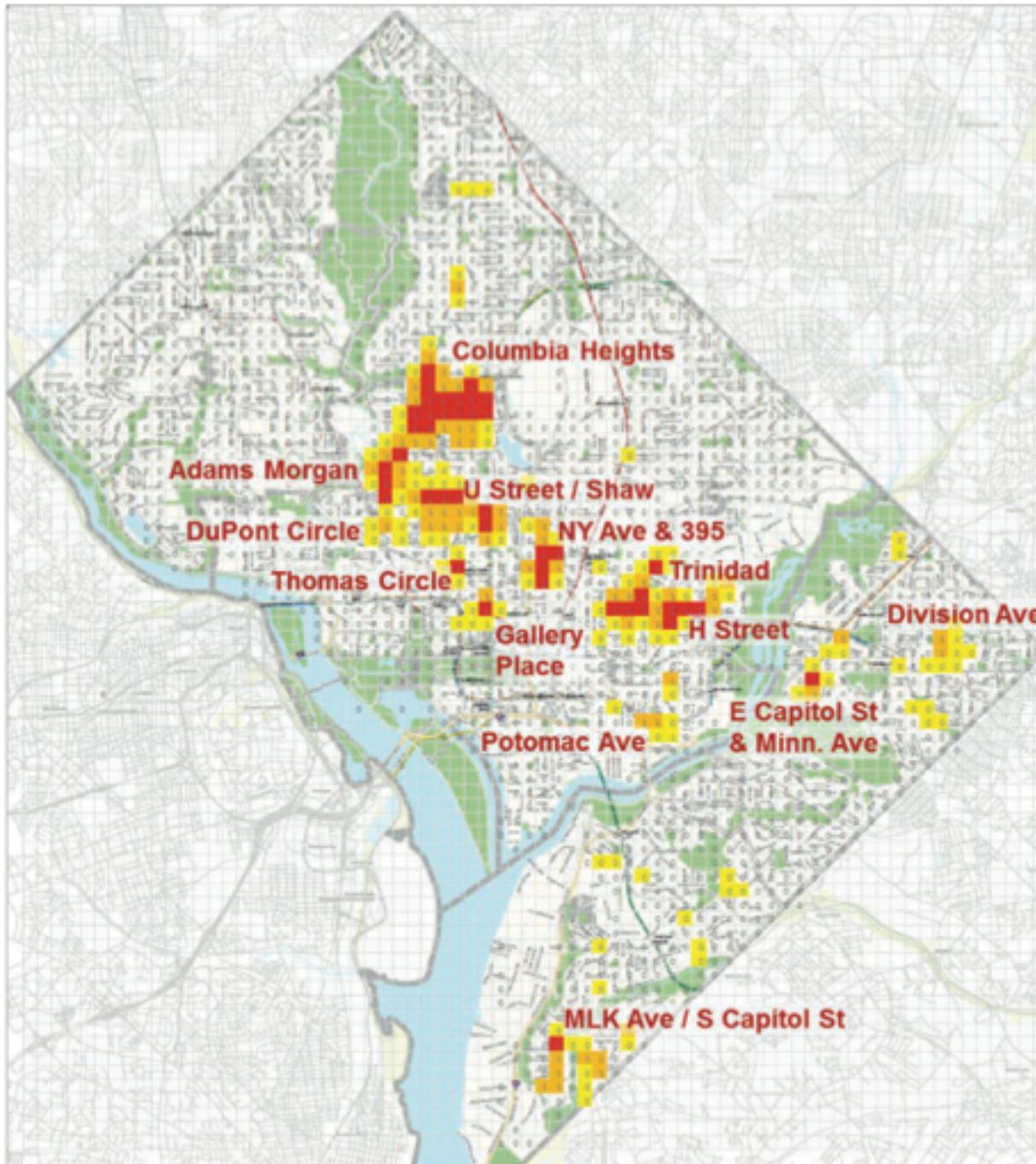
- Se puede modelar la probabilidad de que ocurra un crimen usando una regresión local. Para esto se divide el plano en grids y se estima un modelo de regresión.

Factores para tener en cuenta:

- Hora, día, ciclos.
- Iluminación, transporte
- Densidad poblacional
- Condiciones económicas

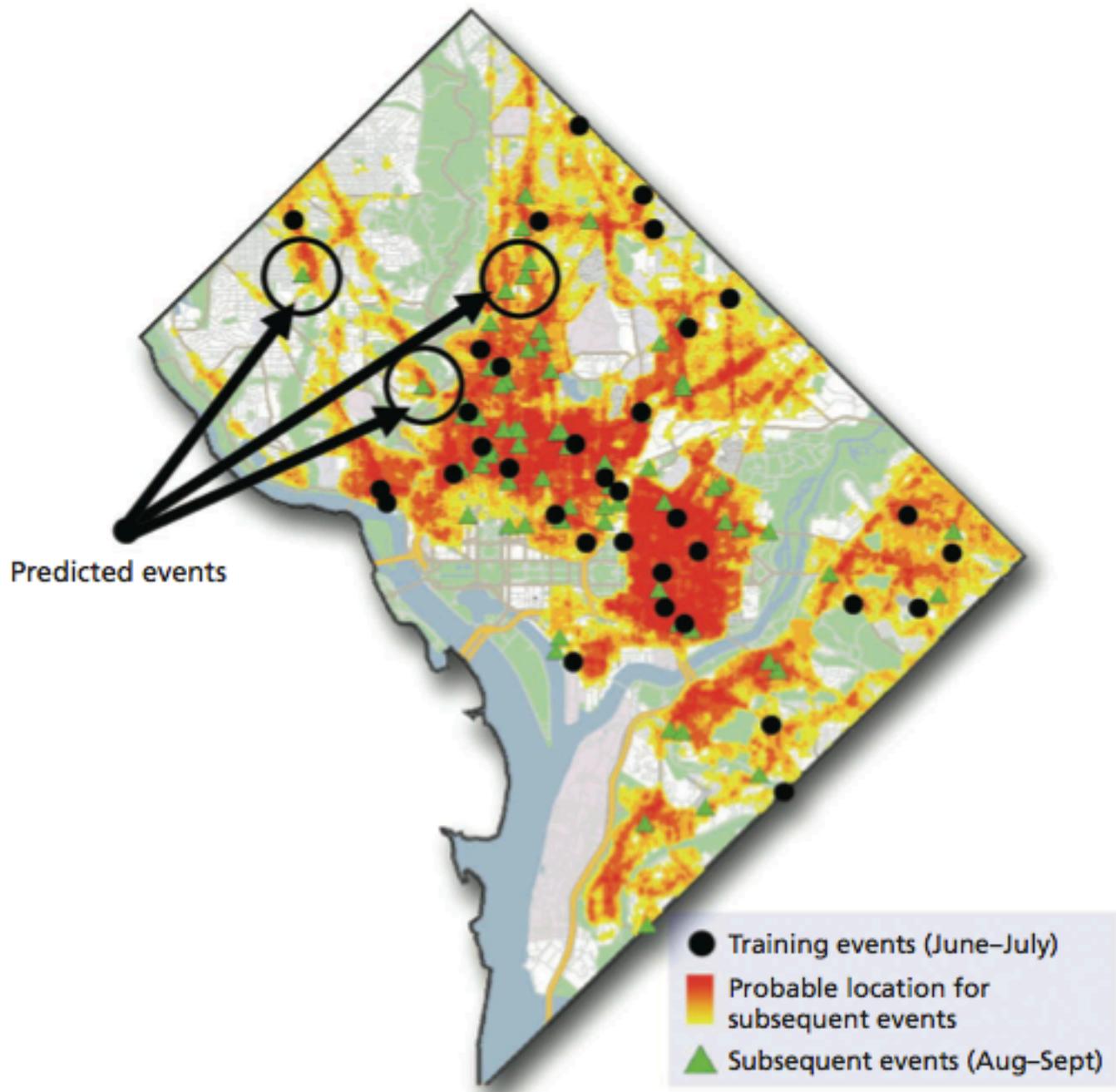
Opciones:

- Modelos Lineales Generalizados



Risk Terrain Modelling

- Realizar predicciones de crimen usando la geografía de la ciudad.
- Tienen un poder predictivo mucho mayor que los hot-spots.
- Permiten deducir una serie de reglas generales.



Métodos Preliminares

- Suponga que un conjunto de crímenes x_1, x_2, \dots, x_n ha sido georeferenciado. Cada uno de estos crímenes es un vector $\mathbf{x} = (x^{(1)}, x^{(2)})$
- Suponga además que existe un único criminal cuyo centro de operaciones está denotado por z .
- Bajo el supuesto de que el criminal no quiere cometer crímenes lejos de su centro de operaciones, tiene sentido encontrar z minimizando la función

$$S(\mathbf{y}) = \sum_{i=1}^n f(d(\mathbf{x}_i, \mathbf{y})) = f(d(\mathbf{x}_1, \mathbf{y})) + \dots + f(d(\mathbf{x}_n, \mathbf{y}))$$

Métodos Preliminares

$$S(\mathbf{y}) = \sum_{i=1}^n f(d(\mathbf{x}_i, \mathbf{y})) = f(d(\mathbf{x}_1, \mathbf{y})) + \dots + f(d(\mathbf{x}_n, \mathbf{y}))$$

- Se proponen varias métricas: Euclidiana y Manhattan.
- A la función f se le conoce como función de decaimiento.

Linear: $f(d) = A + Bd$;

Negative exponential: $f(d) = Ae^{-\beta d}$;

Normal: $f(d) = A(2\pi S^2)^{-1/2} \exp[-(d - \bar{d})^2/2S^2]$;

Lognormal: $f(d) = A(2\pi d^2 S^2)^{-1/2} \exp[-(\ln d - \bar{d})^2/2S^2]$

Métodos Preliminares

Desventajas:

- Existe un único criminal.
 - No existe un modelo predictivo para el siguiente crimen.
 - La solución es muy sensible a la calibración de parámetros.
-
- Los algoritmos PST y H-PST presentan alternativas similares, que intentan minimizar la energía del criminal, pero tienen los mismos problemas.

Método Probabilístico

- Se supone ahora que el lugar donde se cometerá el siguiente crimen es una medida de probabilidad que depende del centro de operaciones y de la distancia promedio que se aleja un criminal α .
- La densidad será denotada por $P(\mathbf{x}|\mathbf{z}, \alpha)$
- De nuevo, se puede hacer una estimación puntual de los parámetros \mathbf{z} y α maximizando la función de verosimilitud.

$$L(\mathbf{y}, a) = \prod_{i=1}^n P(\mathbf{x}_i|\mathbf{y}, a) = P(\mathbf{x}_1|\mathbf{y}, a) \cdots P(\mathbf{x}_n|\mathbf{y}, a)$$

- De nuevo, esto trae problemas.

Método Probabilístico

- Usando la Regla de Bayes se tiene que

$$P(\mathbf{z}, \alpha | \mathbf{x}) = \frac{P(\mathbf{x} | \mathbf{z}, \alpha) \pi(\mathbf{z}, \alpha)}{P(\mathbf{x})}$$

- Suponiendo independencia entre el centro de operaciones y la distancia recorrida promedio, se llega a que

$$\pi(\mathbf{z}, \alpha) = H(\mathbf{z}) \pi(\alpha)$$

- Así que

$$P(\mathbf{z}, \alpha | \mathbf{x}) \propto P(\mathbf{x} | \mathbf{z}, \alpha) H(\mathbf{z}) \pi(\alpha)$$

Método Probabilístico

- El supuesto de independencia se puede hacer si la estructura de la ciudad es homogénea (zona rural).
- Es necesario hacer una distinción entre crímenes: robos vs. Violencia sexual.
- Asumiendo además independencia entre crímenes después de controlar por z y α

$$P(\mathbf{z}, \alpha | \mathbf{x}_1, \dots, \mathbf{x}_n) \propto P(\mathbf{x}_1 | \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n | \mathbf{z}, \alpha) H(\mathbf{z}) \pi(\alpha)$$

- Con lo que

$$P(\mathbf{z} | \mathbf{x}_1, \dots, \mathbf{x}_n) \propto \int P(\mathbf{x}_1 | \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n | \mathbf{z}, \alpha) H(\mathbf{z}) \pi(\alpha) d\alpha.$$

- Ojo que aquí se calcula la densidad de z .

Método Probabilístico

- Finalmente, se asume que después de determinado el centro de operaciones y la distancia promedio que viajará el criminal, realizar o no un crimen dependerá de ciertas preferencias geográficas (como lo son la ubicación de estaciones de policía, las rutas de patrullaje, etc.)

$$P(\mathbf{x}|\mathbf{z}, \alpha) = D(d(\mathbf{x}, \mathbf{z}), \alpha)G(\mathbf{x})N(\mathbf{z}).$$

- Donde D modela el efecto de la distancia, G las preferencias y N es una función normalizadora

Método Probabilístico

- En la literatura se asume que D es una función de densidad normal o es exponencial negativa.

$$D(d, \alpha) = \frac{1}{4\alpha^2} \exp\left(-\frac{\pi}{4\alpha^2} d^2\right)$$

$$D(d, \alpha) = \frac{2}{\pi\alpha^2} \exp\left(-\frac{2}{\alpha} d\right)$$

Método Probabilístico

- Se debe estimar qué tan atractiva es una ubicación geográfica para el ladrón, es decir, se debe calibrar la función G .
- Se podría usar información del pasado y estimar G por Kernels de forma no paramétrica.

$$G(\mathbf{x}) = \sum_{i=1}^N K(\mathbf{x} - \mathbf{c}_i | \lambda)$$

- También se podría identificar factores que incentiven o desincentiven el robo y ajustar los anchos de banda de forma no paramétrica.
- Es importante hacer una distinción entre tipo de crimen pues G puede variar.

Método Probabilístico

- Se debe calibrar además la distribución a priori de la distancia promedio y de la ubicación de los centros de operación.
- Para la ubicación de los centros de operación $H(\mathbf{z})$ se puede usar información de la policía junto con densidad poblacional

$$H(\mathbf{z}) = \sum_{i=1}^{N_{\text{blocks}}} p_i K(\mathbf{z} - \mathbf{q}_i | \sqrt{A_i})$$

- Para estimar la función de distancia promedio $\pi(\alpha)$ se pueden usar técnicas no paramétricas con ayuda de la muestra.

Método Probabilístico

- Si no se busca ubicar los centros de operaciones sino los lugares donde ocurrirán crímenes, se tiene que

$$P(\mathbf{x}_{\text{next}}|\mathbf{x}_1, \dots, \mathbf{x}_n) = \iiint P(\mathbf{x}_{\text{next}}|\mathbf{z}, \alpha)P(\mathbf{z}, \alpha|\mathbf{x}_1, \dots, \mathbf{x}_n) dz^{(1)} dz^{(2)} d\alpha$$

- Así que

$$P(\mathbf{x}_{\text{next}}|\mathbf{x}_1, \dots, \mathbf{x}_n) \propto$$

$$\iiint P(\mathbf{x}_{\text{next}}|\mathbf{z}, \alpha)P(\mathbf{x}_1|\mathbf{z}, \alpha) \cdots P(\mathbf{x}_n|\mathbf{z}, \alpha)H(\mathbf{z})\pi(\alpha) dz^{(1)} dz^{(2)} d\alpha$$

Método Probabilístico

Ventajas:

- Genera una distribución de probabilidad a partir de la cual se puede priorizar la acción policial.
- Permite identificar centros de operación, además de predecir la ocurrencia de crímenes.

Desventajas:

- Predicción de parámetros.
- No participa el tiempo.

Modelo con Tiempo

- Para cada momento del tiempo se define una densidad de probabilidad dada por

$$P(\mathbf{x}_{\text{next}} | \mathbf{x}_1, \dots, \mathbf{x}_n) = \sum_{i=1}^n K(\mathbf{x}_i, \lambda) \tau(t_i)$$

- Donde t_i indica el tiempo transcurrido desde el crimen y τ es una función de decaimiento.

Modelo con Tiempo

- Se realiza una combinación de los dos modelos, de tal forma que la nueva distribución de probabilidad estará dada por

$$P = \chi P_1 + (1 - \chi) P_2$$

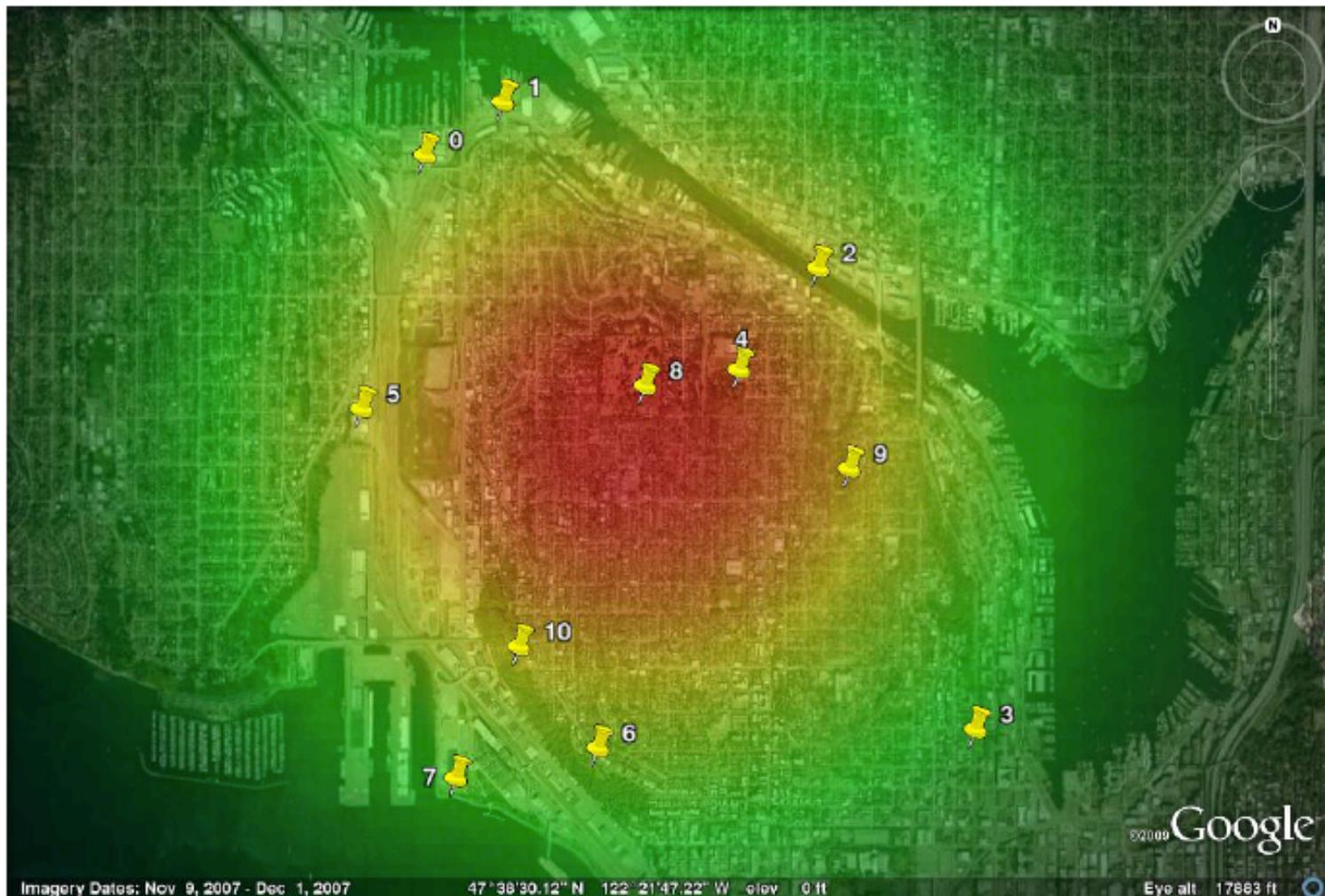


Figure 4: Prediction of next crime location using anchor-distance-decay

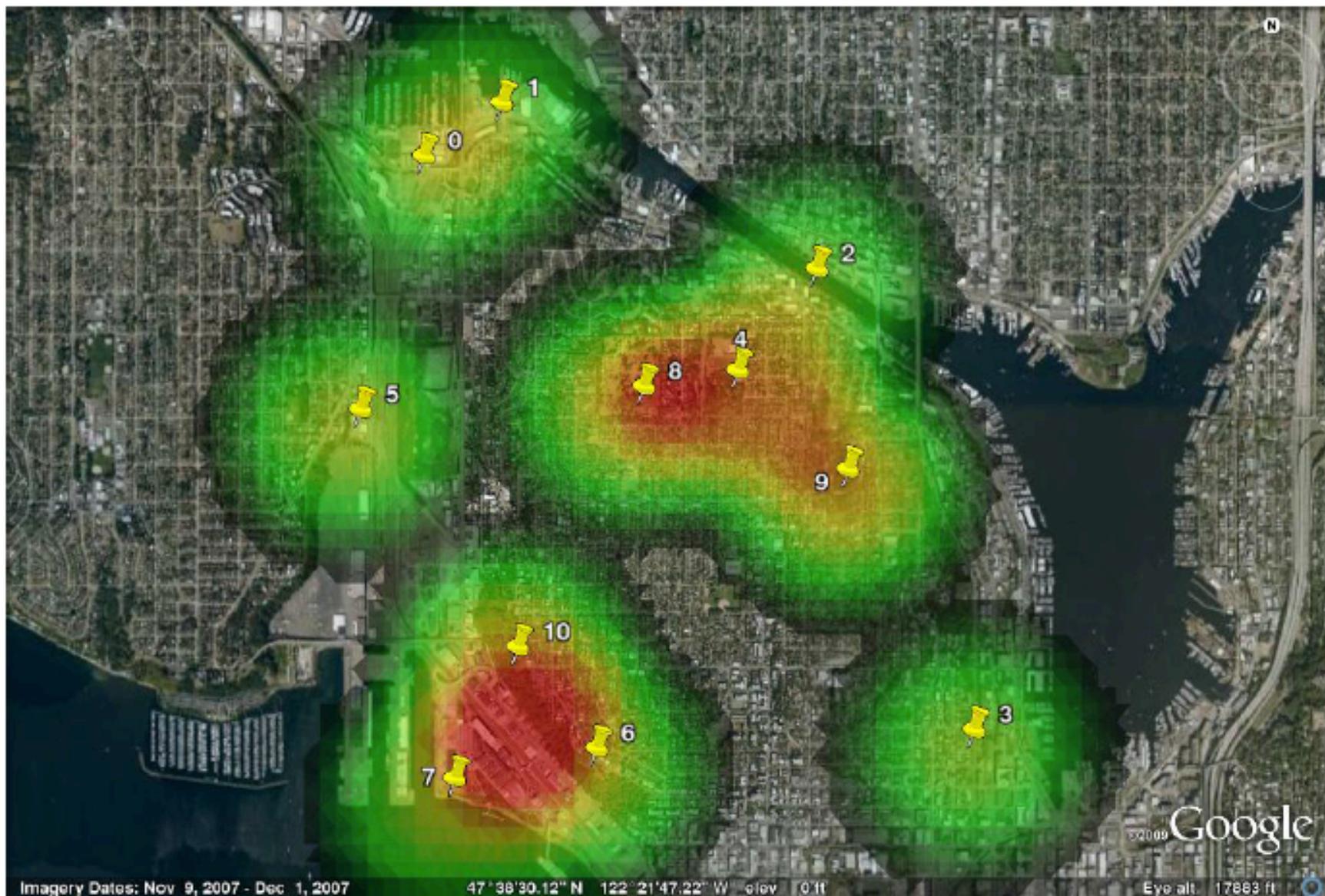


Figure 5: Prediction of next crime location using hot-spotting



Figure 6: Prediction of next crime location using the combined model

Análisis de Correlación

- Se desea determinar si dos procesos espacio-temporales están correlacionados o no.
- No se hace un análisis independiente del tiempo y el espacio, pues esto puede llevar a relaciones espurias: Población vs. Estacionalidad.
- Los tests de Mantel, Knox y Diggle permiten construir estadísticos no paramétricos para hacer análisis de correlación: tablas de contingencia y random shuffling.
- Problema: Establecer thresholds

	near in space	far in space
near in time	X	n_1
far in time	n_2	n_3

Análisis de Correlación

- Estos tests presentan problemas, así que un nuevo acercamiento es testear si para un proceso espacio-temporal las coordenadas de tiempo y de espacio están correlacionadas.
- Para esto, se testea si las dos distribuciones $P(X,Y)$ y $P(X)P(Y)$ son cercanas o no. El proceso puntual es sólo 1
- El criterio usado es MMD

$$\text{MMD} = \sup_f \left(E_{x \sim p}[f(x)] - E_{y \sim q}[f(y)] \right)$$

- La versión muestral está dada por

$$\widehat{\text{MMD}} = \sup_f \left(\frac{1}{n} \sum_i f(x_i) - \frac{1}{n} \sum_j f(y_j) \right)$$

Análisis de Correlación

- Si nos restringimos a funciones con un producto interno, entonces el MMC2 será igual a

$$\widehat{\text{HSIC}} = \frac{1}{n^2} \sum_{i,j} k(x_i, x_j)l(y_i, y_j) - \frac{2}{n^3} \sum_{i,j,q} k(x_i, x_j)l(y_i, y_q) + \frac{1}{n^4} \sum_{i,j,q,r} k(x_i, x_j)l(y_q, y_r)$$

- Para calcular la distribución dada la hipótesis nula se usa randomización.
- Qué pasa cuando el proceso puntual es doble?

Análisis de Correlación - Aplicaciones

- Análisis de correlación entre crímenes violentos y llamadas al 911: se busca establecer medidas preventivas.
- Análisis de concentración de crímenes: Existen zonas calientes o no? Prostitución vs. Ofensas Sexuales.

Propuestas

- Estimar la probabilidad de un crimen en el tiempo usando datos y estimación por kernels.
- Explicar esta probabilidad usando la geografía del espacio. Ajustar los parámetros de un nuevo kernel (amplitud de banda e intensidad). Se debe hacer una división por tipo de crimen.
- También se podría usar el método de grids que se usa en los softwares policiales.
- Usar metodologías no paramétricas para relacionar la probabilidad de crimen con la presencia de ciertos elementos en la geografía (regiones de alta probabilidad vs. Baja probabilidad).
- Determinar para qué tipo de crímenes ocurren agrupaciones y usar esta información en las otras metodologías. (Viajes cortos vs. Viajes largos del criminal).