

COMPARACIÓN DE DIFERENTES MODELOS PARA LA PREDICCIÓN DEL CRIMEN EN BOGOTÁ*

*Francisco Barreras
Carlos Díaz
Álvaro J. Riascos
Mónica Ribero*

Introducción

El mapeo de puntos calientes es una herramienta que se ha venido utilizando desde hace bastante tiempo por académicos y departamentos de policía, con el fin de identificar las zonas en las que hay alta probabilidad de crimen, de tal forma que se pueda asignar de manera eficiente los limitados recursos policiales para su prevención. Existen diversas técnicas

- * Agradecemos a Polkan García por sus comentarios y a Sergio Camelo por su excelente colaboración en la investigación. También agradecemos a Daniel Mejía, secretario de Seguridad de Bogotá, por su apoyo y dirección durante la realización del estudio y el apoyo financiero del Centro de Estudios sobre Seguridad y Drogas (CESED) de la Universidad de los Andes en Bogotá. Este trabajo no hubiese sido posible sin la información compartida por la Policía Metropolitana de Bogotá. Todos los errores u opiniones son responsabilidad exclusiva de los autores y no comprometen a la Policía Metropolitana, a la Alcaldía de Bogotá o la Universidad de los Andes.

de mapeo y estas han evolucionado para ser visualmente concisas y de fácil interpretación. La escogencia del modelo, sin embargo, tiene implicaciones sobre la forma y definición de los puntos calientes y, por tanto, en la eficacia en la prevención del crimen. A diferencia de otros mapas de calor, un mapa de crimen debe servir como herramienta para orientar el patrullaje y la disposición de recursos de prevención de crimen, los cuales están restringidos al espacio público y, en particular, a las calles. Los diferentes modelos pueden implicar definir los puntos de calor como intersecciones de calles, polígonos no uniformes o áreas arbitrarias. En los modelos más específicos es claro cómo orientar el patrullaje, pero no lo es en el caso en el que los puntos calientes son formas grandes o irregulares.

En este estudio, se exploran cuatro técnicas de modelamiento diferentes para la prevención del crimen: modelo de puntos, de elipses espaciales, estimación de densidad por Kernel y un modelo espacio-temporal. Se explora el Índice de Precisión de Exactitud (PAI) (Chainey, Tompson y Uhlig, 2008), una medida similar a las medidas de discriminación predictiva en la ciencia de datos (p. ej. medida F, área bajo la curva ROC) como medida de eficacia que permite comparar la capacidad predictiva de los modelos fuera de la muestra. Por su naturaleza, esta medida permite entender el porcentaje de error tipo I y tipo II para diferentes tamaños de puntos calientes, contribuyendo a la discusión en criminología sobre el tamaño óptimo de los puntos (véanse Weisburd, Bernasco y Bruinsma, 2009). Los modelos se estimaron con la información de criminalidad de Bogotá, Colombia, durante el periodo 2011-2012.

Los resultados indican que el modelo espacio-temporal predice mejor los puntos calientes de crimen en comparación con los otros tres modelos. Estos resultados son robustos, pues se calculó el valor esperado de nuestra métrica de rendimiento, el PAI (véase su definición más adelante), promediando resultados de veinte semanas de prueba diferentes seleccionadas aleatoriamente.

Para los cuatro modelos que no incluyeron de manera clara la dimensión temporal, se obtuvieron resultados óptimos estimando diferentes parámetros por cada día de la semana en rangos de seis horas, lo que permitió capturar las dinámicas cambiantes del crimen en Bogotá.

Parte de la contribución de este trabajo fue la implementación del modelo de KDE (Estimación de Densidad de Kernel, por sus siglas en inglés) en un *software* para dispositivos móviles, destinado a su eventual uso por los patrulleros de policía de dicha ciudad. En la última sección se discuten sus especificaciones y su desarrollo.

Diferentes modelos para la predicción del crimen

El mapeo de puntos calientes ha sido una herramienta útil para predecir dónde el crimen tiende a ser más alto y su aplicación ha sido utilizada para ayudar a las patrullas de policía (Hough y Tilley, 1998; Lavigne y Wartel, 1998). Este tiene como base la premisa de que el crimen tiene más probabilidad de ocurrir en zonas geográficas en las que ha sucedido en el pasado, lo cual provee una oportunidad para explotar la gran cantidad de datos geocodificados de crimen, disponibles en las últimas décadas gracias a la sofisticada tecnología del GPS y los sistemas de geocodificación.

Hay diversas técnicas para agrupar y representar los datos espacio-temporales que han sido desarrolladas para facilitar la visualización y uso práctico en los departamentos de policía. No obstante, hay pocos estudios que comparan esas diferentes técnicas en su habilidad para predecir un crimen en el futuro (Chainey, Tompson y Uhlig, 2008). El propósito de este capítulo es utilizar la medida de precisión de los modelos de predicción del crimen basada en (Chainey, Tompson y Uhlig, 2008) que compara una “tasa de éxito” normalizada con el área cubierta para los diferentes modelos y así encontrar cuál de estos modelos es el más apropiado para predecir el crimen en Bogotá.

Algunos de los modelos de predicción de crimen, como el modelo de puntos discutido a continuación, son simplemente agregaciones y visualizaciones de los eventos ocurridos en el pasado. Sin embargo, modelos como el de elipses y el de KDE tienen parámetros que pueden calibrarse con base en una función de costo y, por ende, hacen parte de la familia de modelos conocidos como aprendizaje de máquinas (*machine learning*, en inglés). Dichos modelos, y otros de mayor complejidad, se están volviendo más populares en la última década por su flexibilidad para incorporar diversas fuentes de datos, para predecir con precisión variables como la ubicación geotemporal del crimen y otras variables relacionadas como la violencia o la reincidencia criminal. Para ampliar esta discusión véase (Bazzi *et al.*, 2017).

La diferencia en desempeño de los diferentes modelos puede ser explicada por el tamaño de los segmentos de datos que permiten identificar puntos calientes. Si la agrupación de los datos es granular, se hace difícil identificar los clústeres (es decir, *agrupaciones*) de crimen (Chainey y Ratcliffe, 2013), porque la visualización tendrá demasiadas zonas demarcadas como “puntos calientes”. Por otro lado, si los datos se agrupan de acuerdo con diferentes criterios geográficos habrá un *problema de unidad*

de área modificable (Openshaw, 1984), en el que las estadísticas resultantes son altamente sensibles a la escogencia arbitraria de los límites de agrupación de los datos.

En las últimas décadas, algunas medidas han sido propuestas para comparar diferentes técnicas de mapeo de puntos calientes y su capacidad para predecir la ocurrencia del crimen. Una de ellas es la “tasa de éxito” que cuenta cuántos futuros crímenes ocurrieron en las zonas marcadas por el modelo como puntos calientes. Sin embargo, esta medida es de poca utilidad, ya que no es normalizada por área. Así, un modelo puede tener una tasa de éxito del 100 % marcando toda el área de estudio como un punto caliente. Para hacer frente a este problema, Bowers, Johnson y Pease (2004) propusieron una medida llamada “tasa de eficiencia de búsqueda” definida como el número de eventos por kilómetro cuadrado en áreas marcadas como puntos calientes.

Como se ha señalado en (Chainey, Tompson y Uhlig, 2008), la “tasa de eficiencia de búsqueda” no considera la relación con el tamaño del área de estudio. Dos escenarios con una tasa de eficiencia de búsqueda de veinte crímenes pueden diferir en el porcentaje de crímenes presentados en los puntos calientes y en el tamaño de estos. Es importante resaltar el hecho de que los puntos calientes más específicos definen un modelo más útil al pensar en el despliegue de los recursos policiales para la prevención del crimen.

En (Chainey, Tompson y Uhlig, 2008) los autores presentan una medida que normaliza por el número total de crímenes en la evaluación del estudio del área, así como el área total de estudio, esto es el Índice de Precisión de Exactitud (PAI). Esta medida fue encontrada como la más conveniente para comparar entre modelos, no solo por las normalizaciones que permiten comparabilidad entre zonas de estudio y momentos del tiempo distintos, sino también porque refleja el compromiso entre errores tipo I y tipo II. Una medida que solo se enfoque en el error tipo I, por ejemplo, midiendo la efectividad en el área patrullada, beneficiaría modelos que tengan poca área patrullada aunque fracasen en prevenir crimen en el resto del área de estudio; por otra parte, una evaluación enfocada solo en el error tipo II podría puntuar muy bien un modelo que concluya que hay que patrullar toda el área de estudio, sin considerar las limitaciones de recursos. Las medidas que tienen en cuenta ambos errores resultan en modelos óptimos más balanceados y, por ende, más aplicables.

Modelo de puntos

El modelo de puntos es la estrategia de agregación más simple para detectar puntos calientes. Este se basa en la idea de que los puntos calientes son microáreas definidas como celdas del orden del largo de una calle o una cuadra en una manzana (Weisburd, Bushway, Lum y Ming Yang, 2004; Weisburd, Groff y Ming Yang, 2012). Entonces, la agregación sobre un campo de tamaño fijo de esta magnitud permite la detección de esos puntos calientes, cuando los crímenes concentrados en un solo segmento excedan cierto umbral.

Este modelo es muy sencillo de implementar, los únicos parámetros a calibrar son el tamaño del campo y el umbral de puntos calientes, esto ha sido una herramienta práctica en la predicción de puntos calientes temporales pequeños (Gorr y Lee, 2015). No obstante, tiene algunas limitaciones resultantes de la discrecionalidad con la que se escojan las celdas y sus tamaños. Por ejemplo, un clúster de crimen puede pasar desapercibido al quedar “artificialmente” dividido en cuatro celdas contiguas, en la que cada una refleje un nivel bajo de crimen. Es más, dado que las celdas tienen un ancho fijo, hace imposible detectar concentraciones de crimen que conformen un punto caliente más grande que el tamaño de una celda, y, en consecuencia, puede limitar la visualización del efecto que un punto caliente tenga sobre áreas circundantes.

Modelo de elipses espaciales

El primer modelo de elipses comenzó como un programa para identificar “círculos calientes” usando datos espacio-temporales de criminalidad (Block, 1995). El desarrollo de estos modelos llevó a la superposición de círculos y su eventual degeneración en elipses. La idea principal continúa siendo la misma: identificar puntos de alta concentración en el mapa y ajustar una “elipse” en cada uno. El tamaño de las elipses y su orientación indican la distribución subyacente del crimen.

El modelo de elipses espaciales ha sido ampliamente utilizado gracias a su implementación en el *software* de análisis espacial y temporal del crimen (STAC). Sus aplicaciones son diversas y en (Bowers, 1999; Martin y Britt, 1998) se puede apreciar su popularidad y utilidad para delimitar las regiones en las que se debe tomar acción.

Sus beneficios incluyen (1) la habilidad de destacar las regiones sin usar límites predefinidos como los cuadrantes de policía o límites político-

administrativos (Martin y Britt, 1998), (2) la compatibilidad con múltiples aplicaciones GIS y (3) los pocos parámetros que necesita para la estimación. Sin embargo, este modelo ha sido criticado porque la distribución típica del crimen no sigue elipses y porque no permite un análisis significativo dentro de las regiones resaltadas (Chainey, Tompson y Uhlig, 2008).

El ajuste de una elipse de desviación estándar para un conjunto de puntos es directo. La magnitud de los dos ejes de cada elipse es dada por las siguientes ecuaciones:

$$SDE_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

$$SDE_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}}$$

Donde n corresponde al número de puntos en una celda predefinida, (x_i, y_i) son las coordenadas de cada punto y cada elipse está orientada en dirección a la máxima dispersión posible para cada celda de puntos. El otro parámetro por ser estimado es el radio de búsqueda, que es el parámetro subyacente en un algoritmo de aglomeración que asigna cada punto en el conjunto de datos a un clúster (p. ej. celda). El radio de búsqueda es la distancia máxima que puede separar dos puntos que pertenecen a la misma elipse.

Estimación de Densidad por Kernel (KDE)

KDE es una técnica de suavizamiento que, en este caso, ajusta distribuciones gaussianas centradas en cada punto de la muestra y las suma para producir un mapa suave sobre el área de estudio. Permite capturar la densidad y ocurrencia del crimen a lo largo del área de estudio sin necesidad de conformar áreas de formas predeterminadas como elipses o campos fijos (Chainey, Tompson y Uhlig, 2008). El radio de búsqueda (ancho de banda) para ajustar cada distribución gaussiana captura el nivel de agregación deseado para construir los puntos calientes. El resultado es representado como los valores de intensidad en una cuadrícula cuyas dimensiones también deberían ser calculadas y determinan la granularidad en la visualización de dichos puntos calientes. Sin embargo, no hay

una metodología universal para establecer estos parámetros (Chainey, Tompson y Uhlig, 2008).

Los puntos calientes pueden ser derivados de una estimación KDE escogiendo un límite para la intensidad del crimen en cada celda. Por ejemplo, un porcentaje de puntos en la parte de las celdas con un umbral mínimo de intensidad. Este procedimiento define claramente zonas geográficas que no están limitadas por divisiones arbitrarias como los cuadrantes o censos que puedan sesgar el análisis. Otras ventajas son la simplicidad visual, además de la creciente disponibilidad de su implementación en el *software* GIS (Chainey, Tompson y Uhlig, 2008). Existen varias aplicaciones de esta metodología en la literatura especializada (véanse Eck, Chainey, Cameron y Wilson, 2005; Goldsmith, McGuire y Mollenkopf, 1999).

Para el propósito de este estudio, la comparación más relevante de KDE y otros modelos está en (Chainey, Tompson y Uhlig, 2008), donde se prueba que el KDE supera en desempeño a otros modelos como el de elipses y mapeo de campo temático. Esto probablemente se debe a la flexibilidad de la forma que el KDE permite y a la captura de los efectos de contagio.

La Estimación de Densidad Kernel, distinto de los otros dos modelos, no requiere crear *clústeres*. Este simplemente consiste en ajustar la densidad gaussiana de la probabilidad sobre cada punto en el conjunto y, luego, el correspondiente mapa de intensidad es la suma de todas esas densidades sobrepuestas. Para cada punto, se ajusta la siguiente intensidad.

$$K(x) = \frac{1}{n} \sum_{i=1}^n K_i(x)$$

Donde

$$K_i(x) = \frac{1}{\sqrt{2\pi \det(\Sigma)}} \exp \left[-\frac{1}{2} (x - x_i) \Sigma^{-1} (x - x_i) \right]$$

Y la matriz de varianza-covarianza se supone la misma para todos los Kernel y se estima minimizando el error de predicción del modelo.

Dimensión temporal

La dimensión temporal desempeña un papel crucial en el análisis de los datos de criminalidad. Está claro que los sucesos criminales no tienen la

misma intensidad a través del tiempo. La actividad criminal depende de factores tales como víctimas potenciales, delincuentes motivados y capacidad de vigilancia (Gorr y Lee, 2015), los cuales dependen, a su vez, de la hora del día y del día de la semana. Además, muchas teorías han demostrado un comportamiento contagioso o recurrente del crimen (Johnson, 2008). Es así como los modelos temporales han sido enriquecidos al hacer un paralelo con la sismología, donde movimientos sísmicos son estructuralmente diferentes a sus réplicas (Mohler, Short, Brantingham, Schoenberg y Tita, 2012). En este artículo, algunas suposiciones en los modelos de estimación pretenden capturar estas relaciones para producir modelos más precisos.

Aunque los modelos anteriores no incluían de manera explícita una variable temporal, sí vale la pena aclarar que el tiempo desempeña un papel decisivo sobre su calibración y resultados. En todos ellos se utiliza una ventana de tiempo para decidir qué eventos se deben considerar para la construcción del mapa de calor. En principio, el modelo es más rico si cuenta con más datos, pero ir demasiado hacia atrás en el tiempo puede producir un modelo que refleje factores desactualizados y que ya no se generalizan al presente. La escogencia de la ventana de tiempo se debe hacer optimizando la medida de desempeño predictivo, pero dicha optimización no es relevante para las conclusiones de este capítulo. Para dichos modelos se empleó la misma ventana de tiempo, usando los crímenes comprendidos entre el 1.º de enero del 2012 y el 31 de diciembre del 2013.

El modelo propuesto por Mohler *et al.* (2012) ha sido de interés para este artículo. En este se identifican los clústeres del crimen en el tiempo y espacio usando KDE y el proceso de puntos con intensidad.

$$\lambda(t, x, y) = \nu(t) \mu(x, y) + \sum_{k: t_k < t} g(t - t_k, x - x_k, y - y_k)$$

Aquí, como en sismología, los datos son desagrupados en eventos antecedentes y réplicas. El primer grupo es modelado como un proceso de Poisson con intensidad $\mu(x, y)$. Las réplicas ocurren en el tiempo y el espacio de acuerdo con el Kernel g que pretende capturar la probabilidad decreciente en tiempo y espacio, siendo más alto este riesgo cuando se encuentra próximo a un evento antecedente.

Debido al tamaño de los datos necesarios para ajustar el Kernel g , se utiliza una simulación no paramétrica de Montecarlo que evita usar toda la información obtenida para ajustar el modelo. Sea $\{(t_k, x_k, y_k)\}_{k=1}^N$ una muestra del modelo, p_{ii} la probabilidad de que el evento i sea un

evento antecedente y p_{ij} la probabilidad de que el evento j sea una réplica del evento i .

Entonces

$$p_{ii} = \frac{\mu(t_i, x_i, y_i)}{\lambda(t_i, x_i, y_i)}$$

$$p_{ji} = \frac{g(t_i - t_j, x_i - x_j, y_i - y_j)}{\lambda(t_i, x_i, y_i)}$$

El siguiente algoritmo iterativo es usado para estimar ν , μ y g (para más detalles véase Mohler *et al.* (2012):

- (1) Use $P_{n1} = [p_{ij}]$ para muestrear eventos antecedentes $\{(t_k, x_k, y_k)\}^N k = 1$.
- (2) Estime ν_n, μ_n, g_n de acuerdo con las muestras anteriores.
- (3) Actualice P_n usando (4) y (5) con los nuevos valores de ν_n, μ_n, g_n .

En este estudio se utilizaron kernels gaussianos.

Datos y metodología

Los datos utilizados corresponden a los datos históricos de crimen de la zona urbana de Bogotá, Colombia. Bogotá es la ciudad más grande y más poblada del país, tiene un área de 1587 km² divididos en veinte localidades o distritos. La ciudad tiene una población urbana de 7878783 habitantes al 2015, dando como densidad 4310 habitantes por kilómetro cuadrado. En el 2014, la proporción de personas en situación de pobreza fue del 10,1% y el ingreso promedio fue de 999195 COP, equivalente a 1,62 salarios mínimos. Durante la década de los noventa, Bogotá fue considerada una de las ciudades más peligrosas en el mundo (Acero, 2006), no sorprende, entonces, que el crimen haya sido un tema muy popular en los debates de políticas públicas de las últimas décadas.

La información utilizada fue obtenida y geocodificada por la Policía Metropolitana de Bogotá desde el 2004 al 2014. Esta base de datos comprende 329793 observaciones individuales, en la que cada una tiene sus coordenadas, hora del evento y tipo de crimen. El área de estudio es la zona urbana de Bogotá y los modelos fueron realizados con datos del 2011 y evaluados fuera de muestra en veinte semanas del 2012, escogidas de manera aleatoria.



Mapa 1. Bogotá, Colombia

Fuente: mapa de Google Maps, 2016. Recuperado de <https://goo.gl/maps/6MufsuD6t9C2>

Los datos fueron codificados por tipo de crimen (tabla 1), siendo la categoría 2 un subconjunto redundante de la categoría 3. Las diferentes categorías se encuentran en una lista en la tabla 1 y su distribución en el gráfico 1. Como el propósito de este estudio no era comparar las dinámicas de los diferentes tipos de crimen, se hicieron modelos predictivos

sobre un grupo de crímenes tipo 1, 3, 5, 6, 7 y 9, los cuales tienen una víctima de robo o agresión. La escogencia del conjunto de datos se debió a la necesidad de tener suficiente información local, dado que esta fue segmentada por periodo de tiempo.

Tabla 1. Codificación por tipo de crimen

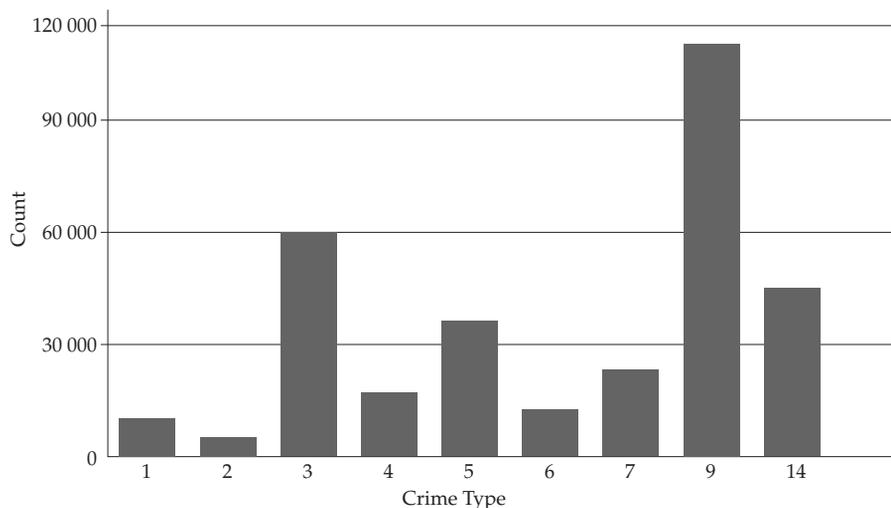
<i>Crimen</i>	<i>Código</i>
Homicidio	1
Lesiones (anteriores)	2
Lesiones	3
Robo de vivienda	5
Robo de motocicleta	6
Robo de automóvil	7
Atraco	9
Tráfico de drogas	14

Fuente: Policía Metropolitana de Bogotá.

Los modelos que no incorporan variables temporales para el mapeo de puntos calientes fueron estimados en periodos comparables de datos basados en el día de la semana y ciertos rangos de horas, con el propósito de comparar su desempeño a través del tiempo. La escogencia óptima de los rangos de horas fue tomada usando el PAI en un conjunto de prueba, equilibrando la disyuntiva entre un modelo específico para cada hora del día y la ausencia de datos suficientes para hacer predicciones acertadas en cada modelo, si la división era muy granular. El mejor desempeño se encontró con los modelos con rangos de seis horas por día.

Con el fin de validar los modelos, se realizó una metodología de conjunto de prueba y entrenamiento estándar (común en la mayoría de las aplicaciones de aprendizaje de máquinas), en la que los modelos se evalúan con datos que no fueron usados durante el ejercicio final. La evaluación de los modelos en los datos de entrenamiento puede llevar a conclusiones erróneas, porque los modelos tienden a “sobreajustar” la información, esto quiere decir que pueden identificar relaciones aparentes o ruidosas entre los datos de entrenamiento, pero que, en general, no se mantienen.

Los modelos fueron entrenados durante diferentes espacios de tiempo antes del 2012 y probados en veinte semanas del 2012, escogidas de



Gráfica 1. Frecuencia por tipo de crimen

Fuente: elaboración propia. Policía Metropolitana de Bogotá.

manera aleatoria. Realizar las pruebas en diferentes semanas es importante para evitar que fenómenos de crímenes específicos por temporada pudieran llevar a concluir que los modelos no son adecuados para realizar predicciones. Por ejemplo, la primera semana del año, Semana Santa o semanas con un fenómeno climático muy particular. El modelo usado para cada periodo de tiempo (día y hora) es utilizado en la semana de prueba correspondiente.

La medida utilizada para comparar los modelos es el Índice de Precisión de Exactitud o PAI, que puede ser calculado una vez el modelo marca ciertas áreas como puntos calientes. El PAI consiste en el cociente de dos medidas: la tasa de éxito y el porcentaje del área; para hacer más consistente esta medida, el PAI fue promediado durante las veinte semanas usadas en la prueba.

$$PAI = \frac{Tasa\ de\ éxito}{Porcentaje\ de\ eje}$$

$$Tasa\ de\ éxito = \frac{Crímenes\ incluidos\ en\ los\ puntos\ calientes}{Crímenes\ totales}$$

$$Porcentaje\ de\ área = \frac{Área\ de\ los\ puntos\ calientes}{Área\ total}$$

Todos los modelos estimados fueron ajustados usando el *software* estadístico R y las librerías “sp” y “rgeos”. Como se dijo el modelo de puntos es la estimación más sencilla, solo requiere especificar el tamaño de cada celda en el campo y luego contar cuántos eventos en cada marco de tiempo caen por celda. Para esta estimación, se escogieron intervalos de seis horas por cada día de la semana y las cuentas fueron por el mismo periodo en el 2011. Esto significa, por ejemplo, que un modelo fue creado para el lunes en el periodo de 00:00-06:00, agregando todos los eventos para el mismo periodo de tiempo para cada lunes del 2012. Las celdas en la cuadrícula fueron establecidas de 150 m cada una. Los puntos calientes fueron definidos como las celdas con el 2% superior en crímenes. Nótese que los puntos calientes son diferentes dependiendo del momento del día y de la semana, capturando así dinámicas de crimen.

Debido al bajo nivel de agregación que viene del modelo de puntos (no hay contagio de una celda a otra, entonces los puntos calientes son separados y claramente definidos), el PAI tiene un comportamiento inestable y no informativo (al dividir por valores cercanos a cero). Esto es porque el área definida como puntos calientes es pequeña (solo cerca del 10% de las celdas posee crímenes). Así, el PAI es a menudo cero cuando el modelo no acierta y tiene un valor muy elevado cuando el modelo predice correctamente un evento en un área. Una solución es utilizar la tasa de éxito o tasa de eficiencia de búsqueda (mencionada antes) en áreas comparables, evitando la inestabilidad del PAI en áreas pequeñas. Se compara el modelo de puntos solo con el KDE usando áreas muy similares (alrededor del 10% del área de estudio) y calculando la tasa de éxito. Debido a su comportamiento explosivo, cuando el modelo es específico, los autores recomiendan solo emplear el PAI en el momento que se tenga una serie extensa de eventos (o densa geográficamente) relativa al área de patrullaje.

El modelo de elipses fue ajustado en el mismo periodo de tiempo usando la biblioteca “aspace” de R. El radio de búsqueda fue definido como 250 m, ya que este es documentado como un radio eficiente (Levine, 2013).

Los únicos parámetros por establecer en el modelo KDE son el ancho de banda (representado por Σ) que determina qué tan dispersa está cada densidad y, por tanto, permite la sobreposición de las densidades y la formación de puntos calientes. El ancho de banda óptimo que fue usado se basó en el resultado del mejor PAI y fue establecido como 150 m. El otro parámetro es el tamaño de la celda de la cuadrícula de intensidades

emitidas que concierne a la visualización y aplicación, para este fue escogido un tamaño de celda de 150 m.

Los mapas 3, 4 y 5 muestran la visualización para tres de los modelos limitados a Usaquén, una localidad de Bogotá.

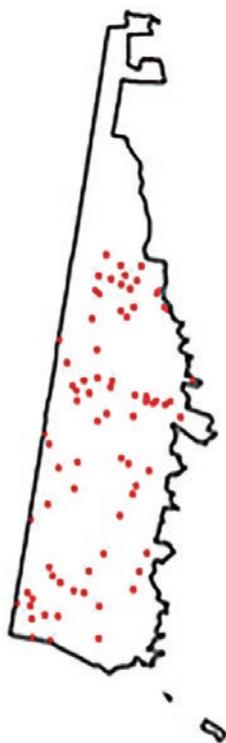


Mapa 2. Localidad de Usaquén, Bogotá

Fuente: mapa de Google Maps, 2016. Recuperado de <https://goo.gl/maps/TxyCbAZiQK92>

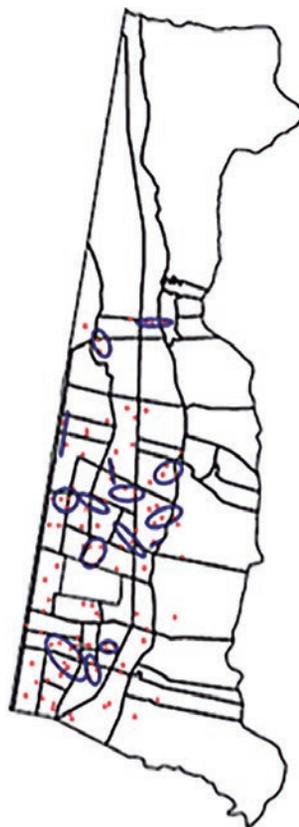
Resultados y conclusiones

A continuación, se presentan los resultados del PAI para los modelos de elipses y KDE. Los resultados son presentados para cada una de las semanas corridas y el promedio (el valor esperado del PAI para ese modelo).



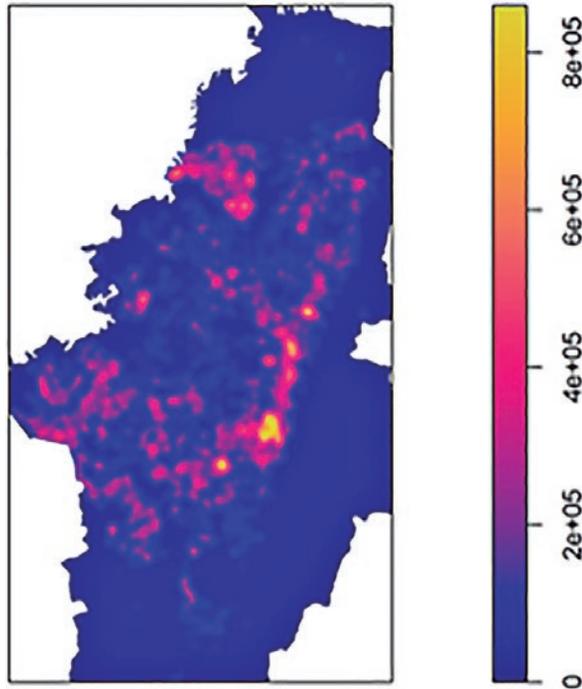
Mapa 3. Visualización del modelo de puntos en la localidad de Usaquéen al nororiente de Bogotá

Fuente: elaboración propia.



Mapa 4. Ejemplo del modelo de elipses espaciales en la localidad de Usaquéen al nororiente de Bogotá

Fuente: elaboración propia.



Mapa 5. Mapa de calor generado con el KDE de Bogotá, área metropolitana

Fuente: elaboración propia.

Tabla 2. Medición del PAI para distintos subconjuntos de prueba

	<i>Ellipses</i>	<i>kde1</i>	<i>kde2</i>	<i>Spatiotemp</i>
1	3,76	8,10	3,18	7,34
2	6,65	6,64	4,76	6,64
3	3,70	6,95	2,14	7,44
4	3,24	5,91	2,05	7,12
5	3,98	6,02	3,88	8,05
6	3,67	6,30	3,02	7,32
7	2,93	7,10	3,56	8,21
8	2,46	5,77	2,12	7,22
9	3,02	6,07	2,54	7,18
10	3,56	7,57	3,22	8,10

Tabla 2. (continuación)

	<i>Ellipses</i>	<i>kde1</i>	<i>kde2</i>	<i>Spatiotemp</i>
11	5,24	7,82	4,13	7,34
12	3,45	7,38	2,16	8,24
13	4,55	6,03	3,94	7,13
14	3,23	6,86	2,04	7,58
15	3,68	5,85	3,64	8,02
16	4,80	5,68	3,55	7,14
17	3,24	5,28	2,18	7,16
18	2,55	5,08	2,07	6,51
19	3,99	5,94	3,68	7,38
20	5,92	6,09	4,66	8,15
Promedio	3,8810	6,4220	3,1260	7,4635

Fuente: elaboración propia.

En la tabla 2 se observa la medida del PAI escogiendo veinte subconjuntos de prueba aleatorios, la medida promedio es más robusta y converge a la medida real. Esta técnica se conoce como “validación cruzada”. Ahora se comparan las tasas de éxito del modelo de puntos y el modelo KDE en un área comparable (aproximadamente el 10 %) para evaluar cuál de los dos modelos es más preciso (tabla 3).

Tabla 3. Comparación de la tasa de éxito para el modelo de puntos y KDE

<i>Semana</i>	<i>Tasa de éxito</i>	
	<i>Puntos</i>	<i>KDE</i>
1	0,09	0,41
2	0,11	0,74
3	0,12	0,77
4	0,10	0,66
5	0,10	0,67
6	0,11	0,70
7	0,08	0,36

Tabla 3. (continuación)

<i>Semana</i>	<i>Tasa de éxito</i>	
	<i>Puntos</i>	<i>KDE</i>
8	0,10	0,64
9	0,10	0,67
10	0,08	0,38
11	0,09	0,39
12	0,08	0,37
13	0,10	0,67
14	0,11	0,76
15	0,10	0,65
16	0,09	0,63
17	0,09	0,59
18	0,08	0,56
19	0,10	0,66
20	0,10	0,68
Promedio	0,09	0,59

Fuente: elaboración propia.

Como se puede apreciar, los modelos KDE superan en desempeño al de elipses y al modelo de puntos. Esto es consistente con los resultados en la literatura (Chainey, Tompson y Uhlig, 2008). Esto no solo se debe a la habilidad de KDE de agrupar a los eventos de manera flexible, una característica de la que el modelo de puntos carece, sino también a que puede definir puntos calientes de una manera muy ajustada, y no necesariamente en las geometrías restringidas del modelo de elipses.

Las diferencias tan grandes observadas entre el PAI del modelo de puntos y el del modelo KDE, se observan debido a que el PAI es más eficiente en la definición de puntos calientes restringiéndose apenas al tamaño de los *kernels* alrededor de los clústeres de crimen. Mientras que en cada celda del modelo de puntos que se asignó como punto caliente se “desperdicia” área (considere que el ancho de estas celdas es mucho mayor al de los *kernels* del modelo KDE), es decir, las mejoras vienen de que KDE es un modelo mucho más “ajustado”.

El modelo espacio temporal es el que tiene mejor desempeño, debido a la incorporación de tendencias temporales que reducen el área de patrullaje cuando se está fuera de dicha tendencia. Este resultado valida la hipótesis de la importancia del tiempo en la predicción de crimen y del efecto “réplica” en la dinámica criminal.

Software desarrollado

Como una aplicación del estudio realizado, los autores diseñaron una aplicación web que permite estimar el modelo KDE y sugerir la asignación de recursos policiales óptimos por día y hora en el área urbana de Bogotá. El propósito es dirigir los recursos para la prevención del crimen hacia donde más se necesitan. El modelo subyacente es el de modelo KDE estimado con una muestra de un año (el año anterior al periodo de predicción y uso de la aplicación).

Este *software* y sus pruebas en campo no se restringen al modelo KDE, sino que podrían realizarse con cualquier modelo predictivo que retorne puntos calientes bien definidos. Esta implementación es valiosa, porque las pruebas en el campo y la retroalimentación por parte de usuarios pueden enriquecer el análisis y producir nuevas preguntas de investigación.

Interfaz general

La aplicación fue desarrollada usando el paquete Shiny de RStudio (2013) que provee un ambiente compatible con HTML para programar en R. Este usa mapas livianos de interacción para mostrar al usuario la distribución del crimen y los segmentos recomendados para patrullar o desplegar los recursos para la prevención del crimen.

La aplicación requiere que el usuario final especifique el día de la semana y la hora del día y esta carga la matriz de intensidad de crimen apropiada, la cual es desplegada en la interfaz. Además, se requiere que el usuario ingrese en un cuadro de texto el nombre del cuadrante de policía en el que el mapa debe centrarse.

La aplicación es compatible con múltiples sistemas operativos y navegadores web, incluyendo teléfonos móviles con capacidad de procesar HTML (la mayoría de los teléfonos inteligentes) (figura 1 y 2).

Función de mapas de calor

La primera función es una interfaz que despliega un mapa de calor de intensidades de crimen sobre un mapa de Bogotá centrado en el cuadrante de policía deseado. Este muestra al usuario dónde se ubican las áreas de alta intensidad de crimen mostrándolas en una escala de color rojo proporcional a la intensidad en la matriz de intensidad de crimen calculada por el KDE.

El usuario puede hacer *zoom* y desplazarse alrededor del mapa para comparar las intensidades con otras áreas. La visualización de estos mapas de calor es una manera rápida para detectar patrones de crimen y puntos calientes. La aplicación muestra diferentes intensidades de crimen dependiendo de la hora y el día de la semana escogidos, obteniendo dinámicas de crimen temporales.

Función de patrullaje por segmentos

La segunda función permite hacer un despliegue más específico y apuntar a los segmentos de calle exactos a los que los recursos deben ser dirigidos. De acuerdo con la restricción actual de que los agentes de policía deben limitar sus patrullas a sus cuadrantes, esta función muestra los segmentos que se intersectan con las áreas de más alta intensidad del cuadrante.

Estas áreas de más alta intensidad son definidas como el 1% de las celdas con la más alta intensidad de crimen en la matriz de intensidad de crimen subyacente generada por el KDE. Como antes, el patrullaje recomendado puede cambiar dependiendo del día y la hora que el usuario ingrese. Cuando el usuario escoge un cuadrante determinado, la pantalla muestra unos segmentos en línea azul delimitando el cuadrante especificado y en línea roja los segmentos señalando las calles a patrullar como se muestra en la figura 2.

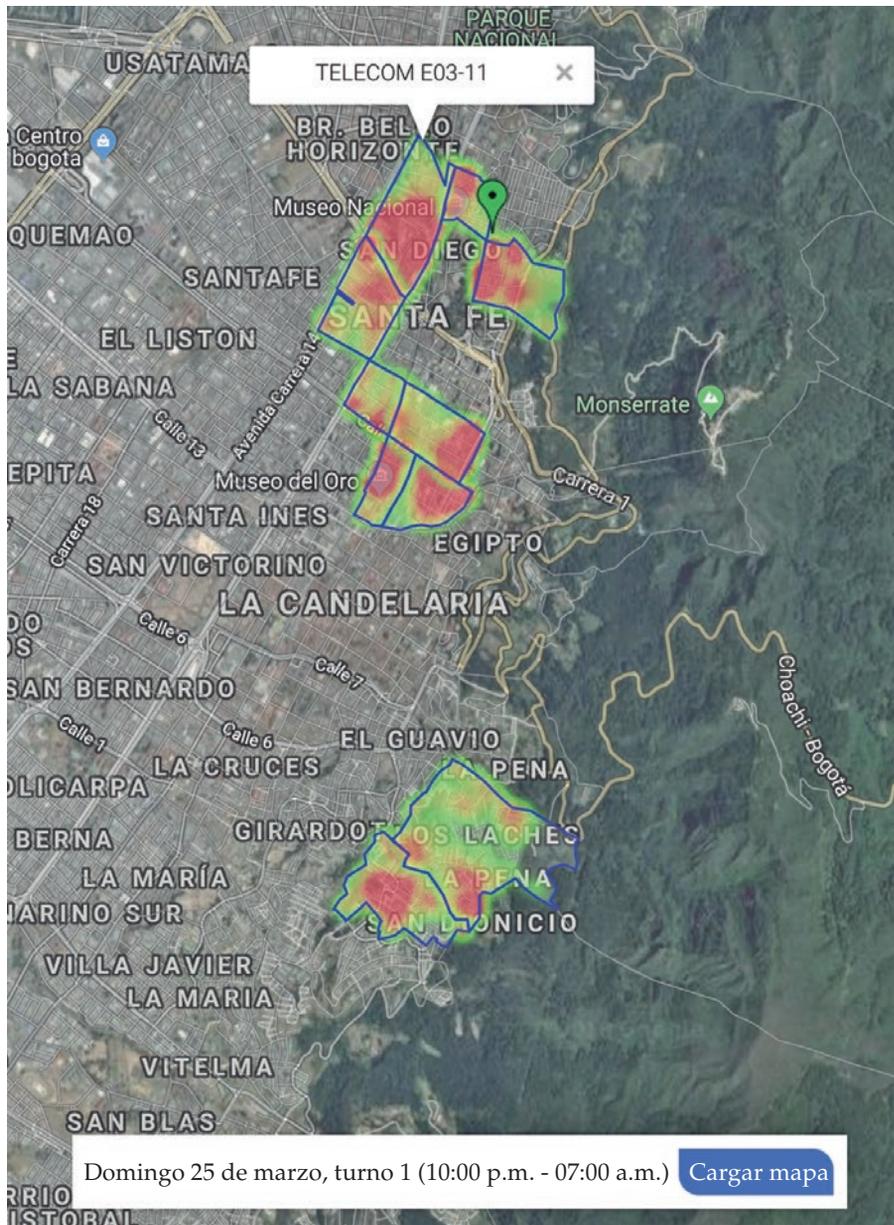


Figura 1. Función de mapas de calor

Fuente: elaboración propia.

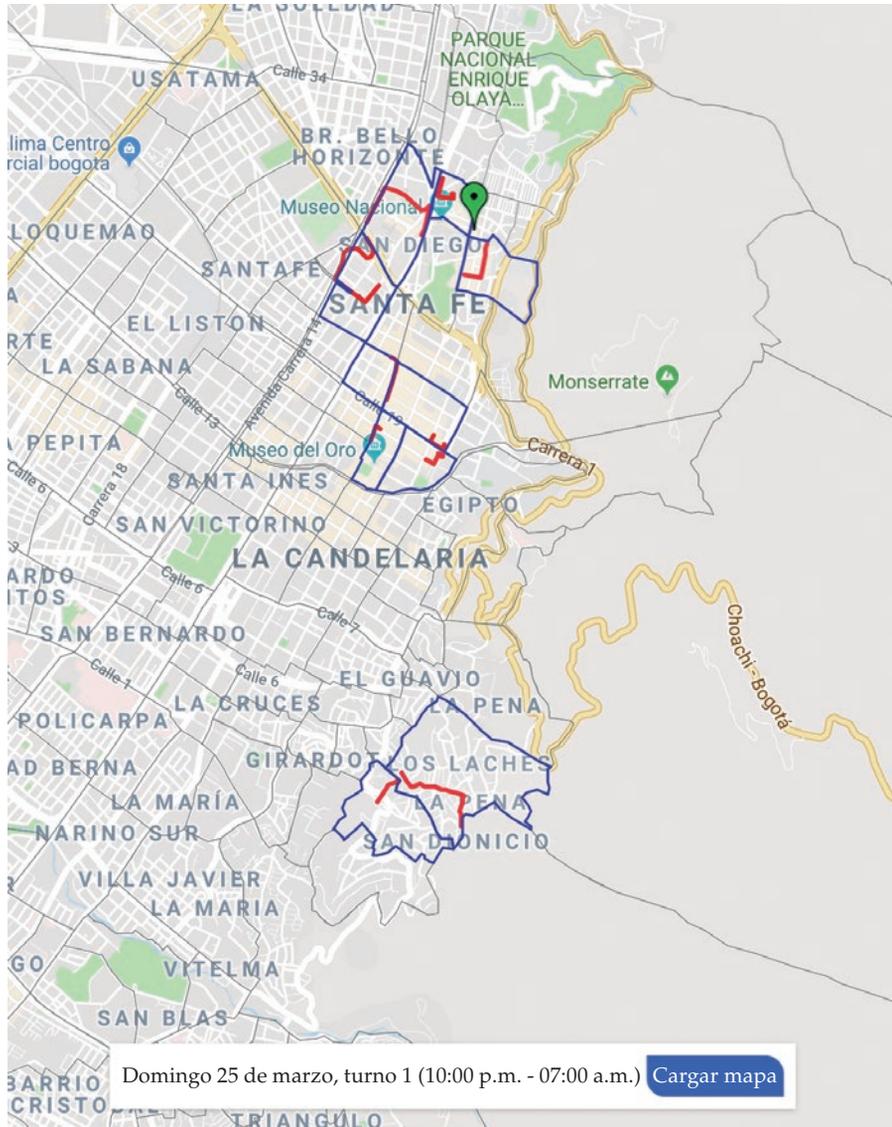


Figura 2. Función de patrullaje por segmentos

Fuente: elaboración propia.

Referencias

- Acero, H. (2006). *Bogota's success story, comunidad segura*. Recuperado de www.comunidadessegura.org/?q:en/node/31203
- Bazzi, S., Blair, R., Blattman, C., Dube, O., Gudgeon, M. y Peck, R. (2017). *What can prediction teach us about violence? Machine learning applications in Indonesia and Colombia*. San Diego: NBER.
- Block, C. R. (1995). Stac hot - spot areas: A statistical tool for law enforcement decisions. En C. R. Block, M. Dabdoub y S. Fregly (eds.), *Crime analysis through computer mapping* (pp. 15-32). Washington: Police Executive Research Forum.
- Bowers, K. (1999). Exploring links between crime and disadvantage in northwest England: An analysis using geographical information systems. *International Journal of Geographical Information Science*, 13(2), 159-184.
- Bowers, K., Johnson, S. y Pease, K. (2004). Prospective hot - spotting the future of crime mapping? *British Journal of Criminology*, 44(5), 641-658.
- Chainey, S. y Ratcliffe, J. (2013). *GIS and crime mapping*. Londres: John Wiley & Sons.
- Chainey, S., Tompson, L. y Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, 21(1), 4-28.
- Easy web applications in R*. (2013). Recuperado de www.rstudio.com/shiny.
- Eck, J., Chainey, S., Cameron, J. y Wilson, R. (2005). *Mapping crime: Understanding hotspots*. Washington: National Institute of Justice.
- Goldsmith, V., McGuire, P., Mollenkopf, J. y Ross T. (1999). *Analyzing crime patterns: Frontiers of practice*. California: Sage Publications.
- Gorr, W. y Lee, Y. (2015). Early warning system for temporary crime hot spots. *Journal of Quantitative Criminology*, 31(1), 25-47.
- Hough, M. y Tilley N. (1998). *Getting the grease to the squeak: Research lessons for crime prevention*. Crime Prevention and Detection Paper 85. Londres: Home Office Police Research Group.
- Johnson, S. (2008). Repeat burglary victimisation: a tale of two theories. *Journal of Experimental Criminology*, 4(3), 215-240.
- Lavigne, N. y Wartell, J. (1998). *Crime mapping case studies: Successes in the field*. Washington: Police Executive Research Forum.

- Levine, N. (2013). *CrimeStat: A spatial statistics program for the analysis of crime incident locations*. Washington: National Institute of Justice.
- Martin, B. y Britt, D. (1998). *The multiple impacts of mapping it out; police, geographic information systems (GIS) and community mobilization during devil's night in Detroit, Michigan*. En N. Lavigne y J. Wartell (eds.), *Crime mapping case studies: Successes in the field* (pp. 3-13). Washington: Police Executive Research Forum.
- Mohler, G., Short, M., Brantingham, J., Schoenberg, F. y Tita, G. (2012). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493), 100-108.
- Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geobooks.
- Weisburd, D., Bushway, S., Lum, C. y Ming Yang, S. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of Seattle. *Criminology*, 42(2), 283-322.
- Weisburd, D., Bernasco, W. y Bruinsma, G. (2009). *Putting crime in its place. Units of analysis in geographic criminology*. Nueva York: Springer-Verlag.
- Weisburd, D., Groff, E. y Ming Yang, S. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Nueva York: Oxford University Press.