
Un indicador de criminalidad procesada en Colombia mediante análisis de componentes principales ponderados geográfica y temporalmente

Prosecuted crime index for Colombia through geographically and temporally weighted principal component analysis

Tulia Eva Salcedo Palacios^a
tuliasalcedo@usantotomas.edu.co

Mario José Pacheco López^b
mariopacheco@usta.edu.co

Resumen

El análisis de la criminalidad es esencial para desarrollar políticas de seguridad efectivas. Este trabajo se enfoca en la construcción de un indicador para la evaluación de la criminalidad procesada en Colombia a nivel municipal durante el período 2018-2023. Se integraron variables de los delitos tipificados penalmente, junto con la densidad poblacional, para identificar patrones espaciotemporales de criminalidad y focos de alto riesgo mediante el análisis de componentes principales ponderados geográfica y temporalmente. Se implementó una herramienta robusta para una mejor explicación del fenómeno criminal, subrayando la relevancia de la autocorrelación y la heterogeneidad espaciotemporal en su análisis. Se realizó una comparación de metodologías de análisis multivariado espacial, evaluando diversas funciones kernel y anchos de banda para optimizar los modelos. Los resultados evidencian una marcada heterogeneidad espacial en la criminalidad, lo que sugiere la presencia de patrones no aleatorios determinados por factores geográficos y temporales. Este hallazgo proporciona una comprensión más precisa de la distribución espacial del crimen. Finalmente, se construye un indicador espaciotemporal que permite el análisis y la comparación entre regiones y periodos, lo cual facilita la implementación de medidas preventivas y correctivas más eficaces en las áreas más afectadas.

Palabras clave: índice de criminalidad, GWPCA, GTWPCA, análisis espaciotemporal.

Abstract

^aUniversidad Santo Tomás

^bUniversidad Santo Tomás

Crime analysis is essential to put in place effective security policies. This study focuses on creating an index to assess borough prosecuted crime in Colombia from 2018 to 2023. The index integrates variables of criminal offences along with population density to identify spatiotemporal patterns of criminality and high-risk hotspots through geographically and temporally weighted principal component analysis. The implementation of a robust tool allowed a better explanation of the criminal phenomenon, highlighting the relevance of autocorrelation and spatial-temporal heterogeneity in its analysis. A comparison of spatial multivariate analysis methodologies was carried out by evaluating different kernel functions and bandwidths to optimise the models. The results signal a strong spatial heterogeneity in crime, which suggest the presence of non-random patterns determined by geographical and temporal elements. These findings provide a more precise understanding of the spatial distribution of crime. Finally, a spatiotemporal indicator is created to allow for analysis and comparison across regions and time periods, facilitating the implementation of more effective preventive and corrective measures in the most affected areas.

Keywords: crime rate, GWPCA, GTWPCA, spatiotemporal analysis.

1. Introducción

Después de la pandemia COVID-19, se han afectado las condiciones de desigualdad, inseguridad alimentaria y disminución de oportunidades económicas de los ciudadanos. De acuerdo con el informe de 2023 de la *Global Initiative Against Transnational Organized Crime*, los grandes cambios políticos y las dificultades económicas repercuten en casi todos los aspectos de la vida cotidiana, impulsando directamente la dinámica del crimen a nivel mundial. En este marco, los intereses criminales tienden a explotar la inestabilidad, la vulnerabilidad y la escasez de productos básicos.

Las altas tasas de criminalidad en Colombia representan costos sociales y económicos significativos (Gómez Pinto, 2022). Estos incluyen la victimización letal y no letal, la pérdida de ingresos cesantes de la población privada de libertad, los gastos en seguridad y los altos costos fiscales debido a las inversiones en organizaciones del Estado dedicadas a la protección ciudadana. En este contexto, la implementación de metodologías sólidas que evalúen diversos escenarios es crucial para obtener un panorama claro de la situación delictiva y sus implicaciones en la seguridad ciudadana (Cadena-Urzúa et al., 2023). Las actividades criminales no solo afectan negativamente la calidad de vida, sino que también repercuten económicamente al desalentar la inversión extranjera, incrementar los costos de seguridad y deteriorar la percepción internacional del país (Pardo, 2018). Desde 2015, el Gobierno Nacional de Colombia ha implementado nuevas disposiciones en seguridad y convivencia ciudadana para mejorar los registros sobre criminalidad en zonas urbanas y rurales, y analizar la evolución de los patrones delictivos.

La prevención en este ámbito es crucial para mejorar la percepción de seguridad

en las sociedades modernas, tanto en áreas rurales como urbanas, dado los elevados costos derivados de las movilizaciones por inconformidad en diversas regiones. En Colombia, el crimen representa una de las principales preocupaciones públicas. Por lo tanto, cuantificar correctamente el impacto de los factores principales que influyen en la criminalidad en una región permitirá diseñar políticas públicas correctivas, enfocadas en las poblaciones más afectadas (de Justicia y del Derecho, 2017).

La recopilación y análisis de datos sobre delitos penalizados, como homicidios, hurtos, secuestros, entre otros, proporcionan una base objetiva para la toma de decisiones gubernamentales y la asignación eficiente de recursos. Los registros administrativos relacionados a las denuncias de diversos delitos se organizan en el Sistema de Información Estadístico Delincuencial, Contravencional y Operativo (SIEDCO). A partir de estos registros se elaboran los reportes y análisis criminológicos que sirven como base para interpretar los cambios en los patrones de criminalidad desde una perspectiva temporal y territorial (Cubides y Céspedes, 2016).

Los reportes oficiales a menudo suelen subregistrar delitos debido a la poca confianza en las entidades protectoras, lo que limita la efectividad de los recursos legales disponibles para cuantificarlos. Además, estos reportes generalmente se limitan a la generación de estadísticas descriptivas sobre las tasas de algunos delitos específicos, priorizando los homicidios debido a su significativo impacto social (Quintero Cuello et al., 2008), dejando en segundo plano otros delitos que también afectan gravemente la seguridad ciudadana, tales como los delitos sexuales y el hurto. Este enfoque tiene limitaciones importantes, ya que tiende a centrarse únicamente en la delincuencia y la inseguridad sin abordar el impacto profundo que el crimen tiene en la calidad de vida cotidiana (Cohen, 2008).

Desde esta perspectiva, los análisis criminológicos adquieren una relevancia particular al destacar la necesidad de centrarse en la medición de procesos judiciales activos. Esto se debe a que solo estos procesos, alineados con los principios de legalidad y debido proceso, permiten validar la existencia de un delito de manera formal. En consecuencia, conceptos como la criminalidad aparente, basada en denuncias, y la criminalidad oculta, que refiere a hechos no conocidos por las autoridades, carecen de valor jurídico para ofrecer una evaluación precisa y confiable de la criminalidad (García, 2022). En esta investigación se emplean exclusivamente datos de delitos procesados judicialmente, por ser una fuente fiable y jurídicamente válida, lo que garantiza conclusiones basadas en información objetiva y verificable, evitando especulaciones de cifras no corroboradas.

La situación en Colombia presenta desafíos únicos para el análisis de la criminalidad debido a la interacción de factores como el conflicto armado interno, el narcotráfico y la violencia urbana, que generan dinámicas complejas y cambiantes. Estas dinámicas incluyen desplazamientos masivos forzados, actividades delictivas de actores armados y patrones heterogéneos en áreas urbanas y rurales, dificultando la consolidación de datos consistentes para un análisis efectivo (Posada, 1999). Además, los patrones delictivos varían significativamente entre regiones, influen-

ciados directamente por factores como la densidad poblacional, que condiciona la concentración de eventos criminales en el espacio (Kim y Hipp, 2021).

Existen varios índices de criminalidad que se utilizan para medir y analizar los niveles de delincuencia en una sociedad. Las tasas de delitos han sido ampliamente utilizadas para comparar la criminalidad de manera geográfica y temporal, al normalizar la incidencia delictiva por cada 100.000 habitantes (Tornero y Andrade, 2020). Sin embargo, estas tasas asumen que los delitos se distribuyen de manera homogénea en una región y en un período, lo cual es una simplificación que no refleja la complejidad real del fenómeno delictivo. Para implementar políticas públicas uniformes se debe emplear herramientas analíticas capaces de considerar las dependencias espaciales y temporales entre regiones, ya que los delitos en una zona pueden estar correlacionados con los de zonas adyacentes (Cetina et al., 2022).

Por ello, es crucial desarrollar indicadores que integren estas características espacio-temporales para capturar de manera más precisa las dinámicas delictivas, superando las restricciones de las tasas simples. Aunque se han implementado enfoques espaciales para el análisis del crimen, la geografía del daño delictivo ha sido poco explorada (Curtis-Ham y Walton, 2017a) a pesar de ser un posible factor clave en el reconocimiento de focos delincuenciales. Incorporar factores como la ponderación geográfica para ajustar diferencias interregionales o la ponderación temporal para reflejar la evolución de los delitos en el tiempo podría mejorar significativamente la precisión de los indicadores y aportar un entendimiento más integral de la criminalidad.

En este marco, resulta esencial adoptar metodologías de medición capaces de explorar los impactos de la variación geográfica en los patrones socioeconómicos y revelar las dinámicas espacio-temporales de los procesos geográficos (Demšar et al., 2012). Este planteamiento destaca la relevancia de herramientas analíticas avanzadas, como el análisis de componentes principales ponderados geográficamente (GWPCA, por sus siglas en inglés: *Geographically Weighted Principal Component Analysis*) y el análisis de componentes principales ponderados geográfica y temporalmente (GTWPCA, por sus siglas en inglés: *Geographically and Temporally Weighted Principal Component Analysis*) que constituyen técnicas multivariadas espaciales diseñadas para abordar las limitaciones asociadas a los enfoques tradicionales en el análisis de datos espaciales y espacio-temporales.

En el presente trabajo, se construyó un índice de criminalidad procesada en Colombia por municipios para el período comprendido entre 2018 y 2023, considerando los delitos penales en Colombia que afectan la seguridad ciudadana, tales como delitos sexuales, hurto a personas y propiedad privada, homicidio, secuestro, violencia intrafamiliar, etc. La construcción se realizó mediante un análisis de componentes principales ponderados geográfica y temporalmente, con el objetivo de evaluar el impacto regional y focalizar territorios de alto riesgo. Adicionalmente, se incluye la densidad poblacional, ya que este factor es relevante en la configuración de la seguridad ciudadana.

2. Antecedentes

Algunos investigadores han empleado distintas técnicas para analizar los patrones de criminalidad en diversas regiones del mundo, proporcionando una comprensión más profunda de los patrones delictivos y las variables asociadas. Bello et al. (2014) y Ayoola et al. (2015) aplicaron un análisis de componentes principales (ACP) para estudiar la victimización y las tasas de criminalidad en Nigeria, basados en datos de una entrevista personal cara a cara en el hogar utilizando un procedimiento de selección aleatoria estratificada de múltiples etapas. Estas investigaciones iniciales, centradas en ACP, proporcionaron una visión preliminar de las áreas propensas al crimen.

Posteriormente Summers y Caballero (2017) incorporaron la dependencia espacial en sus análisis, demostrando la autocorrelación espacial entre robos y ubicaciones específicas, mediante la implementación de una prueba de significancia basada en el método de Monte Carlo utilizando datos de robos callejeros en Austin, Texas. Otros autores como Mojica et al. (2017) y Erick et al. (2019) han empleado tanto ACP como análisis de autocorrelación espacial para estudiar las causas de reincidencia delictiva y las tasas de criminalidad en áreas como Metro Manila y Nairobi, respectivamente.

Enfoques más recientes como los de Aguirre-Rodríguez et al. (2021) y Cadena-Urzuá et al. (2023) avanzaron en esta línea, identificando patrones delictivos en Perú y Chile mediante el uso de modelos espaciales y temporales analizando registros de denuncias y faltas. Estos estudios subrayan la utilidad de técnicas espaciales para comprender la criminalidad y diseñar políticas de prevención basadas en evidencia, destacando la importancia del análisis espacial y la identificación de clústeres en la formulación de estrategias de seguridad.

Actualmente, numerosos países han desarrollado indicadores específicos para medir y analizar la criminalidad, adaptándolos a sus contextos sociales y metodológicos, lo que proporciona antecedentes clave sobre cómo se abordan estas dinámicas a nivel global. En Europa, el *European Sourcebook of Crime and Criminal Justice Statistics* presenta indicadores de criminalidad basados en tasas delictivas, segregados por infracción penal y categorías de delitos. En el Reino Unido, la Oficina de Estadísticas Nacionales (ONS, por sus siglas en inglés) utiliza el Índice de Severidad del Crimen (CSS, por sus siglas en inglés), el cual pondera los delitos según la severidad de las sentencias impuestas, permitiendo la comparación entre distintas áreas locales. En Canadá, el Índice de Gravedad del Delito (CSI, por sus siglas en inglés) mide anualmente los cambios en la criminalidad asignando mayor peso a los delitos más graves y se actualiza con datos de la Encuesta Integrada de Tribunales Penales (ICCS, por sus siglas en inglés). En Chile, la Fundación Paz Ciudadana elabora el Índice de Victimización, basado en encuestas sobre la prevalencia de delitos y la percepción de inseguridad, cuyos resultados son utilizados para formular políticas públicas. En Honduras, el Observatorio de la Violencia de la Universidad Nacional Autónoma de Honduras (OV-UNAH) publica el Índice de Violencia y Criminalidad, utilizando datos de instituciones de seguridad y salud, analizando

la incidencia de homicidios y otros delitos a través de sistemas de información geográfica (SIG), proporcionando una evaluación detallada de la violencia en el país.

Diversos autores han propuesto índices innovadores para medir la criminalidad que, aunque no se encuentran implementados actualmente, ofrecen enfoques valiosos para el análisis del impacto del crimen. En países como Dinamarca (Andersen y Mueller-Johnson, 2018), Nueva Zelanda (Curtis-Ham y Walton, 2017b) y Suecia (Kärrholm et al., 2020) se han desarrollado índices basados en el daño social causado por diferentes delitos, diseñados para proporcionar una medida relativa del impacto delictivo. Estos indicadores buscan optimizar la asignación de recursos policiales limitados, permitiendo enfocar los esfuerzos en áreas que puedan maximizar la reducción del daño y facilitar la medición de los efectos de dichas acciones. Sin embargo, ninguno de estos métodos integra el análisis espacio-temporal.

En Colombia, se han diseñado diversos indicadores enfocados en la medición de la criminalidad, los cuales ofrecen aproximaciones relevantes para comprender este fenómeno en el contexto nacional. Quintero Cuello et al. (2008) proponen el primer indicador que integra diversos delitos entre 1994 y 2007, utilizando metodologías como promedios móviles, logaritmo y valor z , ponderadas según la pena impuesta. Norza-Céspedes y Espino-Duque (2011) realizan un análisis estadístico y espacial preliminar de los delitos en 2010, pero se limitan a tasas específicas sin abordar todas las categorías delictivas. Ortega et al. (2015) desarrollan un índice para Bogotá, Medellín, Cali y Barranquilla, que identifica “hotspots” delictivos, aunque sin considerar una posible heterogeneidad temporal. Cetina et al. (2022) desarrollaron un índice multivariado de medición del conflicto interno colombiano utilizando una metodología mixta, basado en una revisión de estudios previos y un ACP seguido de un modelo de regresión múltiple que abarca el periodo de 2014 a 2021, aunque no considera los efectos espacio-temporales en su construcción. Actualmente, el Ministerio de Justicia utiliza estadísticas oficiales, pero estas se limitan a tasas desagregadas, careciendo de la robustez necesaria para un análisis integral de la criminalidad y la formulación de políticas de seguridad basadas en datos actualizados.

La integración de dimensiones espaciales y temporales resulta crucial para comprender las dinámicas delictivas en diferentes regiones y períodos de tiempo. La literatura sugiere que la ausencia de un indicador que combine estas dos perspectivas limita una comprensión integral de los patrones delictivos, lo que a su vez afecta la formulación de políticas públicas más efectivas para la prevención y control del crimen.

3. Marco Teórico

El marco teórico se estructura en cinco secciones: análisis de componentes principales ponderados geográficamente (GWPCA), que introduce la extensión del ACP a contextos espaciales para capturar la variabilidad local, que es fundamental en

la identificación de patrones geográficos de criminalidad; análisis de componentes principales ponderados geográfica y temporalmente (GTWPCA), que añade la dimensión temporal para analizar la no estacionariedad espaciotemporal, que es esencial para reflejar cambios en las dinámicas delictivas a lo largo del tiempo; selección de parámetros, centrada en optimizar el ancho de banda y el balance espacial-temporal, garantizando así la precisión de los resultados; construcción de un indicador, donde se integran las variables relevantes en un índice ponderado, permitiendo la comparación espaciotemporal de la criminalidad; y la prueba de Monte Carlo, que asegura la significancia estadística de las variaciones en los autovalores, fortaleciendo la validez del modelo.

3.1. Análisis de Componentes Principales Ponderados Geográficamente (GWPCA)

En el ámbito de los SIG, la reducción del número de capas facilita la modelización y simplifica la interpretación de las relaciones entre variables, un objetivo que se logra frecuentemente mediante el ACP. Sin embargo, en aplicaciones espaciales, el ACP estándar no considera los efectos espaciales, lo que puede limitar su capacidad para capturar la autocorrelación inherente a los procesos geográficos. Para abordar esta limitación, el GWPCA es una alternativa más adecuada, ya que permite analizar la heterogeneidad espacial y estudiar la dinámica geográfica de las tasas delictivas y la densidad poblacional asociada. Esta metodología, al capturar la variación local, ofrece una herramienta más efectiva para revelar las estructuras espaciales cambiantes en el análisis de los principales delitos (Demšar et al., 2012).

Según Harris et al. (2011), en el GWPCA se asume que un vector de variables observadas x_i en la ubicación espacial i sigue una distribución normal multivariada con un vector medio μ y una matriz de varianza-covarianza Σ , es decir, $x_i \sim N(\mu, \Sigma)$. Si la ubicación espacial i tiene coordenadas (u, v) , entonces el ACP con efectos geográficos locales considera a x_i como condicional en u y v , haciendo que μ y Σ sean funciones de estas coordenadas; es decir, $x_i|(u, v) \sim N(\mu(u, v), \Sigma(u, v))$. De este modo, tanto μ como Σ dependen de u y v , lo que implica que cada elemento de $\mu(u, v)$ y $\Sigma(u, v)$ también es función de estas coordenadas. Por lo tanto, $\mu(u, v)$ y $\Sigma(u, v)$ corresponden al vector medio y la matriz de varianza-covarianza ponderados geográficamente (PG) respectivamente.

En el GWPCA, las covarianzas son ponderadas según la distancia entre el objeto espacial bajo análisis y sus vecinos cercanos (Fotheringham et al., 2002). Para obtener los componentes principales PG, se descompone la matriz de varianza-covarianza ponderada geográficamente, lo que permite calcular los valores propios y los vectores propios PG. El producto de la i -ésima fila de la matriz de datos con los vectores propios de la ubicación i -ésima genera la i -ésima fila de las puntuaciones de los componentes PG. La matriz de varianza-covarianza ponderada geográficamente es:

$$\Sigma(u, v) = X^T W(u, v) X$$

Donde X es la matriz de datos, con n filas que representan las observaciones y m columnas para las variables, mientras que $W(u, v)$ es una matriz diagonal de pesos geográficos, derivada de una función kernel en la que los pesos dependen de las distancias entre las ubicaciones y el ancho de banda que define el tamaño de la vecindad.

Lu et al. (2014) describen los componentes principales locales en la ubicación (u_i, v_i) de la siguiente forma:

$$L(u_i, v_i)V(u_i, v_i)L(u_i, v_i)^T = \Sigma(u_i, v_i)$$

donde $L(u_i, v_i)$ es una matriz de vectores propios locales, $V(u_i, v_i)$ es una matriz diagonal de autovalores locales, y $\Sigma(u_i, v_i)$ es la matriz de covarianza local. La matriz de puntuaciones de componentes locales, $T(u_i, v_i)$, se obtiene como:

$$T(u_i, v_i) = XL(u_i, v_i)$$

Aquí, el producto de la fila i -ésima de la matriz de datos con los vectores propios locales en la ubicación i -ésima proporciona la i -ésima fila de las puntuaciones de los componentes locales. Al dividir cada autovalor local por $tr(V(u_i, v_i))$, se obtiene una versión localizada de la proporción de varianza total (PVT) explicada por cada componente en los datos originales. Así, para un GWPCA con m variables, en cada ubicación observada se calculan m componentes, m autovalores, m conjuntos de cargas de los componentes (cada uno de dimensión $m \times m$) y m conjuntos de puntuaciones de los componentes (de tamaño $n \times m$). También es posible estimar autovalores y autovectores en ubicaciones no observadas; sin embargo, como no existen datos en estas ubicaciones, no es posible calcular las puntuaciones de los componentes. Las cargas de los componentes reflejan los coeficientes de correlación entre las puntuaciones de los componentes y los datos originales.

En cualquier etapa de calibración de un método PG, es prudente experimentar con diversas funciones de ponderación kernel (Harris et al., 2014). En la Tabla 1 se indica cada uno de estos kernels que se definen de la siguiente manera:

Modelo Global	$w_{ij} = 1$
Gaussiano	$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right)$
Exponencial	$w_{ij} = \exp\left(-\frac{ d_{ij} }{b}\right)$
Box-car	$w_{ij} = \begin{cases} 1 & \text{si } d_{ij} < b \\ 0 & \text{de otra manera} \end{cases}$
Bicuadrado	$w_{ij} = \begin{cases} (1 - (d_{ij}/b)^2)^2 & \text{si } d_{ij} < b \\ 0 & \text{de otra manera} \end{cases}$
Tricubo	$w_{ij} = \begin{cases} (1 - (d_{ij} /b)^3)^3 & \text{si } d_{ij} < b \\ 0 & \text{de otra manera} \end{cases}$

Tabla 1: Funciones kernel

donde el ancho de banda b es la distancia geográfica; d_{ij} es la distancia geográfica entre las ubicaciones espaciales de las filas i y j ; y w_{ij} es el peso geográfico asignado a un punto de observación indexado por j , para un punto de calibración indexado por i . El kernel gaussiano se asocia con una ventana móvil no ponderada, mientras que los demás emplean ventanas ponderadas. Los anchos de banda pueden definirse como una distancia fija o un número fijo de observaciones locales. Comúnmente, un método PG se especifica con un kernel de decaimiento de distancia, como el bi-cuadrado, debido a que esta especificación tiende a proporcionar salidas suavemente variadas en el espacio. No obstante, el kernel box-car puede generar salidas que aparezcan discontinuas en el espacio, útil para detectar relaciones atípicas. Las calibraciones de box-car son útiles porque si se especifica un ancho de banda del 100 %, se obtiene el modelo global correspondiente. El proceso espacial puede ser homogéneo, en cuyo caso un modelo global es suficiente.

3.2. Análisis de Componentes Principales Ponderados Geográfica y Temporalmente (GTWPCA)

El GTWPCA es una extensión del ACP y GWPCA que incorpora los efectos tanto espaciales como temporales en los datos multivariados (Han et al., 2022). El GTWPCA permite explorar la no estacionariedad espacial y temporal en los datos, lo que es particularmente útil en estudios donde las relaciones entre las variables varían tanto en el espacio como en el tiempo.

El GTWPCA considera una ubicación espaciotemporal (u_i, v_i, t_i) para cada observación, y asume que existe un vector de variables observadas locales X_i asociado a dicha ubicación. El análisis parte de la matriz de varianza-covarianza local ponderada espacio-temporalmente, que se define como:

$$\Sigma(u_i, v_i, t_i) = \mathbf{X}^T \mathbf{W}(u_i, v_i, t_i) \mathbf{X} \quad (1)$$

donde X es la matriz de datos espaciotemporales, con n filas (puntos de datos) y m columnas (variables), y $W(u_i, v_i, t_i)$ es la matriz diagonal de pesos geográficos y temporales, que representa la influencia relativa de cada punto alrededor de la ubicación (u_i, v_i, t_i) sobre los componentes locales calculados. Los puntos cercanos a (u_i, v_i, t_i) tienen mayor peso y, por lo tanto, mayor influencia.

Se asume que los datos observados están ubicados en un sistema de coordenadas tridimensional espaciotemporal, donde los efectos de escala en el tiempo y el espacio están relativamente equilibrados. La distancia espaciotemporal d_{ij}^{ST} entre el punto (u_i, v_i, t_i) y su punto circundante (u_j, v_j, t_j) puede expresarse mediante la combinación lineal de la distancia espacial d_{ij}^S y la distancia temporal d_{ij}^T , de la siguiente manera:

$$d_{ij}^{ST} = \alpha d_{ij}^S + \beta d_{ij}^T = \alpha [(u_i - u_j)^2 + (v_i - v_j)^2] + \beta (t_i - t_j)^2$$

Los parámetros α y β se utilizan para ajustar los efectos de las distancias espaciales y temporales. El cálculo de los pesos espacio temporales se basa en una función kernel bicuadrado que considera la distancia espaciotemporal.

$$W_{ij}^{ST} = \begin{cases} \left(1 - \left(\frac{d_{ij}^{ST}}{r^{ST}}\right)^2\right)^2 & \text{si } d_{ij}^{ST} < r^{ST}, \\ 0 & \text{si } d_{ij}^{ST} \geq r^{ST}, \end{cases} \quad (2)$$

donde r^{ST} es el ancho de banda espaciotemporal, d_{ij}^{ST} es la distancia espaciotemporal entre las ubicaciones i y j , y r^{ST} es el radio de influencia espaciotemporal.

La descomposición de la matriz de varianza-covarianza ponderada espaciotemporal se realiza a partir de los autovalores y autovectores locales. La descomposición se formula como:

$$\mathbf{L}(u_i, v_i, t_i) \mathbf{V}(u_i, v_i, t_i) \mathbf{L}^T(u_i, v_i, t_i) = \mathbf{\Sigma}(u_i, v_i, t_i) \quad (3)$$

donde $L(u_i, v_i, t_i)$ es la matriz de autovectores locales y $V(u_i, v_i, t_i)$ es la matriz diagonal de autovalores locales. Los componentes locales en una ubicación espaciotemporal (u_i, v_i, t_i) se calculan multiplicando la i -ésima fila de la matriz de datos X por el autovector ponderado espaciotemporal correspondiente. La expresión para los puntajes de los componentes es:

$$T(u_i, v_i, t_i) = \mathbf{X}_i \mathbf{L}(u_i, v_i, t_i) \quad (4)$$

Cada autovalor local se divide por la traza de la matriz de autovalores locales $tr(V(u_i, v_i, t_i))$ para calcular el PVT explicado por cada componente en la ubicación espaciotemporal (u_i, v_i, t_i) :

$$\text{PVT}(u_i, v_i, t_i) = \frac{\lambda_j(u_i, v_i, t_i)}{\text{tr}(\mathbf{V}(u_i, v_i, t_i))} \quad (5)$$

Este enfoque permite visualizar cómo la varianza multivariante se distribuye a lo largo del espacio y el tiempo, proporcionando una herramienta útil para explorar la heterogeneidad espaciotemporal en los datos.

El GTWPCA permite identificar patrones locales y temporales en los datos, revelando variaciones que podrían no ser capturadas por un ACP convencional. Además, los resultados del GTWPCA pueden ser visualizados en forma de mapas espaciotemporales que muestran la distribución de los puntajes de los componentes, los autovalores y los vectores de carga locales, lo que facilita la interpretación de las variaciones locales en el contexto de fenómenos espaciotemporales complejos.

3.3. Selección de los parámetros para GWPCA y GTWPCA

Un desafío clave en estas metodologías es encontrar la escala en la cual debería operar cada ACP localizado; es decir, elegir el ancho de banda adecuado del kernel. En diversos estudios de GWPCA, este ancho de banda se ha especificado manualmente por el usuario (Lloyd, 2010; Kumar et al., 2012). Los anchos de banda se pueden especificar como una distancia fija o como un número fijo de observaciones locales (es decir, una distancia adaptativa) (Harris et al., 2011). El uso de un ancho de banda fijo resulta apropiado para conjuntos de datos con densidad homogénea, ya que mantiene una escala constante a lo largo de todo el espacio analizado. Por otro lado, un ancho de banda adaptativo es más adecuado para datos con densidad heterogénea, dado que ajusta dinámicamente la vecindad, lo que permite garantizar una representación equilibrada de la información local (Gollini et al., 2015).

En ACP hay m variables en la matriz de datos \mathbf{X} , de modo que cada observación es un vector en un espacio de m dimensiones, y las puntuaciones de componentes correspondientes a los componentes $q+1$ a m representan las distancias euclidianas a lo largo de los ejes de los vectores ortogonales correspondientes a un subespacio lineal de q dimensiones. Aquí, el subespacio de q dimensiones está formado por las primeras q cargas (también vectores de m dimensiones), y es el subespacio que maximiza la varianza de los puntos de datos proyectados en ese subespacio. Para la reducción de dimensiones, se elige q de modo que este subespacio contenga una alta PVT, por lo tanto, los componentes $q+1$ a m representan la desviación de este subespacio. En términos de puntuaciones de componentes, los primeros q componentes se describen por XL_q , y los componentes restantes por $XL_{(-q)}$ (donde L_q denota la matriz de cargas L con todas las columnas eliminadas excepto las primeras q , y $L_{(-q)}$ denota L con las primeras q columnas eliminadas).

Jolliffe y Cadima (2016) muestran que la mejor aproximación de rango q a X es $XL_qL_q^T$, y que la matriz residual S , dada por $S = X - XL_qL_q^T$, también se puede escribir como $S = XL_{(-q)}L_{(-q)}^T$. Así, a través de componentes principales, se desea

encontrar el mínimo de la expresión

$$\sum_{ik} (X_{ik} - S_{ik})^2$$

con respecto a S , donde S es una matriz de rango q . Por lo tanto, los niveles de varianza de los componentes de S miden un criterio de selección denominado bondad de ajuste (GOF, por sus siglas en inglés) de los subplanos proyectados, donde

$$\text{GOF}_i = \sum_{k=q+1}^m s_{ik}^2$$

es la bondad de ajuste para la i -ésima observación, y s_{ik} es la k -ésima puntuación de componente para la observación i ; es decir, el elemento ik de S . La bondad de ajuste total para todo el conjunto de datos es

$$\text{GOF} = \sum_{i=1}^n \text{GOF}_i$$

Para GWPCA y GTWPCA, los componentes principales locales para la i -ésima ubicación representan una proyección similar, pero con las cargas correspondientes definidas localmente. Es decir, se busca minimizar S

$$\sum_{ik} w_i (X_{ik} - S_{ik})^2$$

En el GWPCA, w_i es un peso definido localmente para la ubicación i , mientras que en el GTWPCA se refiere al peso espaciotemporal en la ubicación i . Las estadísticas de bondad de ajuste en el GTWPCA se definen de manera similar a las del GWPCA, con la diferencia de que la matriz de covarianza local S está ponderada por los pesos espaciotemporales en cada ubicación. El parámetro clave en el GTWPCA es determinar la escala adecuada en la que debe operar cada ACP local. Esta escala operativa está determinada por el ancho de banda del kernel r^{ST} , el cual depende de la distancia espaciotemporal y está influenciada por el parámetro de balance τ , que se define como α / β .

Cuando $\alpha = \beta$, el parámetro τ es igual a 1, lo que implica que las distancias espaciales y temporales tienen la misma influencia. Sin embargo, si α y β son distintos, τ refleja el equilibrio entre las distancias espaciales y temporales. En estos casos, para hallar los valores de α y β , se puede asumir que τ fija la relación entre ambos, y luego se optimizan conjuntamente con el ancho de banda usando procedimientos como la validación cruzada, que permiten ajustar los parámetros del modelo de manera que se minimicen los errores de predicción y se evite el sobreajuste, garantizando así un equilibrio adecuado entre las dimensiones espaciales y temporales.

De esta forma, una vez determinado τ , se puede ajustar uno de los parámetros (generalmente β) y calcular el otro ($\alpha = \tau\beta$).

Han et al. (2023) sigue un enfoque automatizado para la selección del ancho de banda, similar al procedimiento utilizado en el GWPCA y el modelo de regresión ponderada geográficamente (GWR), tal como se describe en Zhao et al. (2018). Tanto el GWPCA como el GTWPCA retienen las propiedades del ACP, y el parámetro τ es crucial porque equilibra los efectos de las distancias espaciales y temporales, lo que determina la escala espaciotemporal en la que opera el GTWPCA.

Las estadísticas de bondad de ajuste permiten encontrar el ancho de banda óptimo y el parámetro τ para el GTWPCA. Harris et al. (2011) afirma que el estadístico GOF proporciona los medios para encontrar un ancho de banda óptimo para el GWPCA utilizando un método de “Leave-one-out” o una muestra de retención al calcular los términos del estadístico. El GOF total se puede calcular para diferentes combinaciones de anchos de banda y valores de τ , y el valor mínimo de GOF determina el conjunto óptimo de parámetros. Al igual que en el GWPCA, el número de componentes a retener en el GTWPCA, representado por q , debe decidirse a priori, ya que el ancho de banda y τ no pueden optimizarse si se retienen todos los m componentes.

3.4. Construcción de un indicador

Para ACP, si n es el número de variables correlacionadas: x_1, x_2, \dots, x_n . Entonces, el índice es una combinación lineal de la forma:

$$I = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

donde:

$$c_i = \frac{r_i}{s_i s}$$

siendo s_i la desviación estándar de la variable i ,

$$s = \frac{r_1}{s_1} + \frac{r_2}{s_2} + \dots + \frac{r_n}{s_n}$$

y r_i es la media cuadrática de las correlaciones entre la variable i y el resto de las variables dada por:

$$r_i = \left(\frac{1}{n-1} \sum_{j=1}^n r_{ij}^2 \right)^{\frac{1}{2}}$$

esto es para $i = 1, 2, \dots, n$; con $i \neq j$

Así, el índice constituye un promedio aritmético ponderado, donde la suma de los ponderadores es igual a la unidad y se construye finalmente como:

$$I = \frac{r_1}{s_1 s} x_1 + \frac{r_2}{s_2 s} x_2 + \dots + \frac{r_n}{s_n s} x_n = \sum_{i=1}^n \frac{r_i}{s_i s} x_i$$

El indicador ponderado espaciotemporal en el GTWPCA se construye como una extensión del indicador basado en GWPCA (García, 2020), añadiendo la dimensión temporal para capturar mejor las dinámicas no estacionarias en el espacio y el tiempo. La construcción del indicador se basa en la combinación de los valores estandarizados de las variables seleccionadas, ponderadas por la proporción de varianza explicada en cada componente principal local, adaptada a cada ubicación espaciotemporal.

Para cada localización (u_i, v_i, t_i) , se define el valor del indicador como la suma ponderada de los valores estandarizados de las variables. Las ponderaciones para cada variable en cada punto están dadas por las proporciones de varianza explicada de los componentes principales en esa ubicación y periodo. En términos formales, el indicador espaciotemporal para la ubicación (u_i, v_i, t_i) se expresa como:

$$I_i^{(u_i, v_i, t_i)} = \sum_{j=1}^p \alpha_j^{(u_i, v_i, t_i)} Z_{ij}$$

donde $\alpha_j^{(u_i, v_i, t_i)}$ representa el peso del j -ésimo componente en la ubicación (u_i, v_i, t_i) , que se basa en la varianza explicada por cada componente principal local (espaciotemporal), reflejando la importancia relativa de ese indicador en dicha ubicación. Z_{ij} corresponde al valor estandarizado del j -ésimo indicador parcial para la i -ésima unidad. En un contexto espaciotemporal, Z_{ij} se extiende para incluir tanto las dimensiones espaciales (u_i, v_i) como la temporal t_i . Esto significa que la estandarización no sólo se hace en función de los datos en diferentes ubicaciones espaciales, sino también a lo largo del tiempo. Es decir, los valores de cada variable j se estandarizan considerando tanto la variabilidad espacial como la variabilidad temporal.

El peso $\alpha_j^{(u_i, v_i, t_i)}$ se calcula como:

$$\alpha_j^{(u_i, v_i, t_i)} = \frac{\sum_{k=1}^p \alpha_{jk}^{(u_i, v_i, t_i)} \lambda_k^{(u_i, v_i, t_i)}}{\sum_{j=1}^p \sum_{k=1}^p \alpha_{jk}^{(u_i, v_i, t_i)} \lambda_k^{(u_i, v_i, t_i)}}$$

donde $\lambda_k^{(u_i, v_i, t_i)}$ es el autovalor asociado a la k -ésima componente principal en la ubicación (u_i, v_i, t_i) , y $\alpha_{jk}^{(u_i, v_i, t_i)}$ es el factor de peso correspondiente al j -ésimo

indicador parcial en la k -ésima componente principal para la ubicación (u_i, v_i, t_i) .

El indicador espaciotemporal resultante se normaliza a una escala de 0 a 100 para facilitar su interpretación y comparación entre ubicaciones y períodos temporales. Esta normalización se realiza mediante la fórmula:

$$ICP_i^{(u_i, v_i, t_i)} = \frac{I_i^{(u_i, v_i, t_i)} - \text{mín}(I)}{\text{máx}(I) - \text{mín}(I)} \times 100,$$

donde $I_i^{(u_i, v_i, t_i)}$ es el valor original del indicador, y $\text{mín}(I)$ y $\text{máx}(I)$ son los valores mínimos y máximos del indicador en todo el conjunto de datos. En resumen, este indicador espaciotemporal en el GTWPCA captura la variabilidad de las variables seleccionadas según su importancia relativa, medida a través de la varianza explicada en cada punto del espacio y tiempo. Esta metodología permite captar las dinámicas locales tanto en el espacio como en el tiempo, ofreciendo una visión más detallada y robusta de los patrones no estacionarios en los datos analizados.

3.5. Prueba de Monte Carlo para GWPCA y GTWPCA

La prueba de Monte Carlo es una herramienta clave tanto en GWPCA como en GTWPCA para evaluar si los autovalores locales varían significativamente a lo largo del espacio (en GWPCA) o a lo largo del espacio y el tiempo (en GTWPCA), lo que proporciona evidencia de no estacionariedad en los datos y justifica el uso de estos modelos.

En los casos de GWPCA y GTWPCA, la prueba de Monte Carlo se utiliza para evaluar la significancia de las variaciones en los autovalores locales. Mientras que en el GWPCA el enfoque está dirigido a identificar variaciones espaciales, el GTWPCA extiende este análisis para capturar tanto la variabilidad espacial como temporal. El procedimiento se detalla en los siguientes pasos:

1. **Aleatorización de las ubicaciones espaciales o espaciotemporales:** En el caso del GWPCA, las ubicaciones espaciales del conjunto de datos se aleatorizan secuencialmente, mientras que en el GTWPCA este proceso se extiende para incluir las dimensiones espaciales y temporales, generando permutaciones espaciotemporales.
2. **Aplicación del análisis (GWPCA o GTWPCA) en cada permutación:** Tras cada aleatorización, se realiza el análisis correspondiente: GWPCA para las permutaciones espaciales o GTWPCA para las permutaciones espaciotemporales. En ambos casos, se recalcula un ancho de banda óptimo adaptado a la permutación.
3. **Cálculo de la desviación estándar (DE) del autovalor local:** En cada permutación, se calcula la desviación estándar (DE) de un autovalor local específico, lo que permite recolectar un conjunto de DEs a partir de todas las permutaciones.

4. **Construcción de la distribución aleatorizada y cálculo del valor real:** Las DEs obtenidas en las permutaciones se ordenan para formar una distribución de DEs aleatorizadas. Posteriormente, se calcula la DE real del mismo autovalor local, usando el conjunto de datos original (sin permutar).
5. **Evaluación de significancia:** La DE real se posiciona dentro de la distribución de referencia. En el GWPCA, este análisis determina si existen variaciones espaciales significativas en el autovalor seleccionado, mientras que en el GTWPCA se evalúan variaciones tanto espaciales como temporales.

En ambos métodos, GWPCA y GTWPCA, el primer autovalor local suele ser el más utilizado para evaluar la no estacionariedad. En el caso del GWPCA, esto se debe a que este autovalor domina sobre los demás componentes (Harris et al., 2011; Harris et al., 2014), mientras que en el GTWPCA, su preferencia radica en su mayor influencia sobre la varianza explicada (Han et al., 2022).

3.6. Limitaciones del GWPCA y GTWPCA

El GWPCA y su versión mejorada, el GTWPCA, son adaptaciones localizadas del PCA convencional diseñadas para abordar análisis en contextos espaciales y espacio-temporales. Sin embargo, ambas metodologías presentan limitaciones que restringen su aplicabilidad en ciertos escenarios analíticos. Una de las principales limitaciones del GWPCA es su incapacidad para integrar explícitamente la dimensión temporal, lo que reduce su utilidad en el análisis de fenómenos dinámicos con dependencias temporales significativas (Han et al., 2022).

En términos de interpretación, tanto el GWPCA como el GTWPCA enfrentan desafíos. La interpretación de las cargas y componentes principales locales resulta compleja, ya que los resultados deben analizarse en función de cargas locales extensas o mediante la identificación de “variables ganadoras”, es decir, aquellas más relevantes a nivel local (Harris et al., 2011). En el caso específico del GTWPCA, la metodología genera múltiples resultados en cada posición espacio-temporal: m componentes principales, m valores propios, m conjuntos de cargas y m conjuntos de puntuaciones. Esta abundancia de información plantea desafíos significativos para su visualización y análisis, especialmente en contextos con alta densidad de datos (Han et al., 2023).

Otro aspecto relevante de estos métodos es su sensibilidad a los parámetros de configuración. Aunque los resultados suelen ser robustos frente a la elección de la función kernel utilizada (Wu et al., 2020), el ancho de banda es un parámetro crítico, ya que afecta directamente la ponderación de las observaciones locales y/o temporales. Los cambios en este parámetro pueden introducir variaciones significativas en los resultados obtenidos (Gollini et al., 2015).

Además, los costos computacionales asociadas con el cálculo de matrices de covarianza locales o espacio-temporales pueden ser prohibitivas cuando se trabaja con conjuntos de datos extensos o de alta resolución espacial, lo que limita la apli-

cabilidad práctica de estas técnicas. Otro desafío importante es el sesgo inducido por variables con escalas mayores, que pueden influir desproporcionadamente en el cálculo de las matrices de covarianza. Este efecto es particularmente relevante en metodologías basadas en PCA, dado que su principio fundamental es la maximización de la varianza, lo que puede favorecer a las variables dominantes en detrimento de otras características igualmente significativas (Feijoo, 2023).

4. Datos

Para la construcción del indicador propuesto en este estudio, se utilizaron datos públicos proporcionados por la Fiscalía General de la Nación (FGN, www.fiscalia.gov.co/), el Departamento Administrativo Nacional de Estadística (DANE, www.dane.gov.co/), y la Infraestructura Colombiana de Datos Espaciales (ICDE, <https://datos.icde.gov.co/>). La base de datos final contiene aproximadamente cuatro millones de registros correspondientes a procesos judiciales desarrollados entre los años 2018 y 2023. Dado que las poblaciones de los municipios evaluados presentan diferencias significativas en su tamaño, se calcularon tasas ajustadas por cada 10,000 habitantes. Estas tasas fueron calculadas para cada uno de los grupos delictivos, los cuales se detallan más adelante, utilizando las proyecciones poblacionales del DANE basadas en el Censo de 2018.

Los datos de la FGN se refieren al número total de víctimas, según los reportes de noticias criminales por tipo de delito registrados en el Sistema Penal Oral Acusatorio, conforme a lo establecido en la Ley 906 de 2004 y la Ley 1098 de 2006. Esta información abarca casos procesados en todo el territorio colombiano, los cuales se encuentran en alguna de las siguientes etapas procesales: investigación, indagación o juicio. El año tomado como referencia temporal corresponde al año en que ocurrieron los hechos que constituyen la presunta conducta punible. Para la elaboración del análisis, se seleccionaron 52 delitos, que posteriormente fueron agrupados en 9 grupos (Tabla A1), según su relación con un fenómeno criminal común.

- **Delitos contra el orden social y seguridad pública (TG1):** Este grupo incluye delitos que afectan el orden público, la seguridad del Estado y la salud pública, así como violaciones a las normativas ambientales y sanitarias. Entre estos delitos están los ambientales, del orden económico social, la salud pública, la seguridad del Estado, la seguridad pública y la violación de medidas sanitarias.
- **Delitos contra la administración pública (TG2):** Estos delitos atentan contra la integridad de las instituciones públicas y el correcto funcionamiento del sistema administrativo. Comprenden corrupción administrativa, electoral, judicial, privada y tributaria, así como falsedad en documento, delitos contra la fe pública y la impartición de justicia.
- **Delitos contra la dignidad y los derechos fundamentales (TG3):**

Engloban aquellos crímenes que vulneran la dignidad humana, los derechos fundamentales y la integridad moral. Incluyen actos de discriminación, delitos contra la familia, delitos sexuales, injuria, calumnia, maltrato animal y violencia intrafamiliar.

- **Delitos contra la libertad y la autodeterminación de la persona (TG4):** Delitos que afectan la libertad individual y el derecho de las personas a tomar decisiones sobre su vida. Incluyen amenazas, constreñimiento, desaparición forzada, reclutamiento ilícito, secuestro extorsivo y simple, trata de personas y el uso de menores de edad.
- **Delitos contra la privacidad y la información (TG5):** Estos delitos vulneran la privacidad y los derechos informáticos. Incluyen delitos informáticos y contra los derechos de autor.
- **Delitos contra la propiedad y el patrimonio económico (TG6):** Delitos que afectan el patrimonio económico de las personas y entidades. Abarcan delitos como estafa, extorsión, hurto y otros que comprometen la integridad del patrimonio económico.
- **Delitos contra la vida y la integridad personal (TG7):** Incluyen crímenes que atentan contra la vida y la salud física de las personas. Entre estos delitos están el aborto, feminicidio, homicidio culposo y doloso, lesiones personales, así como el uso de agentes químicos.
- **Delitos contra los derechos colectivos y bienes públicos (TG8):** Delitos que afectan los derechos de la colectividad y los bienes protegidos por el derecho internacional humanitario (DIH). Incluyen concierto para delinquir, desplazamiento, violaciones al régimen constitucional y otros delitos relacionados con bienes públicos.
- **Delitos relacionados con el uso ilegal de armas y drogas (TG9):** Este grupo abarca delitos relacionados con la fabricación, tráfico y porte ilegal de armas, así como los vinculados al tráfico de estupefacientes.

Finalmente, se utilizó la variable de **densidad poblacional (TG10)**, calculada a partir de los registros del Instituto Geográfico Agustín Codazzi (IGAC, www.igac.gov.co/) sobre la superficie en km² de cada municipio. Esta variable proporciona una medida de la cantidad de personas que habitan un territorio en relación con su superficie poblada.

5. Metodología

Para evaluar la presencia de autocorrelación espacial a nivel municipal, primero se calculó el índice de Moran. A continuación, se realizó un ACP sobre el conjunto completo de datos. Con base en los resultados del ACP global, se implementaron

el GWPCA y el GTWPCA, manteniendo el mismo número de componentes en ambos casos. Esto garantiza una comparación equitativa entre los métodos (Han et al., 2023). En el caso del GWPCA, se utilizaron los kernels disponibles en la función `gwpc` del paquete `GWmodel`: gaussiano, exponencial, bicuadrado, tricúbico y box-car. Para este estudio, se utilizaron anchos de banda adaptativos, adecuados para la distribución espacial irregular de los datos de muestra. La selección del ancho de banda óptimo se realizó conforme a los criterios de bondad de ajuste. Para la selección del ancho de banda en el GWPCA (Harris et al., 2011) y el GTWPCA (Han et al., 2023), se utilizó el método de validación cruzada de “Leave-one-out”, basado en las aproximaciones obtenidas a partir de las tres primeras componentes principales. Este método es particularmente adecuado para calibraciones que involucran ponderaciones geográficas, ya que permite evaluar de manera precisa el desempeño del modelo en cada ubicación individual (Farber y Páez, 2007).

El indicador se determinó mediante GTWPCA, identificando los parámetros óptimos para cada combinación de tiempo y espacio, utilizando el kernel bicuadrado con diferentes anchos de banda y evaluando la bondad de ajuste. La combinación de parámetros espaciales y temporales que minimizó el GOF fue seleccionada como la óptima. Para justificar el uso de GWPCA y GTWPCA sobre el ACP, se realizó la prueba de Monte Carlo, evaluando si los autovalores locales varían significativamente a lo largo del espacio y el tiempo. Después de seleccionar el modelo óptimo de GTWPCA, se construyó el indicador y se representó en mapas para mostrar la evolución de la criminalidad procesada de forma clara y accesible. El indicador fue normalizado para ajustar sus valores en un rango de 0 a 100 y, posteriormente, se clasificó en cinco niveles: muy bajo $[0, 20]$, bajo $(20, 40]$, medio $(40, 60]$, alto $(60, 80]$, y muy alto $(80, 100]$.

6. Resultados

En este estudio se analizaron los 1121 municipios de Colombia, observándose una mayor concentración en la región Andina, que abarca departamentos como Cundinamarca, Antioquia, Boyacá y Santander, caracterizados por una elevada densidad poblacional y un alto grado de urbanización. En contraste, las regiones orientales del país, como la Amazonía y la Orinoquía, exhiben una distribución más dispersa de municipios, acorde con la menor densidad poblacional y el predominio de territorios menos urbanizados.

Además de definir las áreas geográficas de estudio, es fundamental evaluar la posible existencia de no homogeneidad en el espacio y en el tiempo. Todas las variables consideradas presentaron un índice de Moran superior a 0.3, lo que indica una autocorrelación espacial positiva significativa en los datos. Los p -valores asociados, cercanos a cero, confirmaron la existencia de un patrón espacial no aleatorio. Estos resultados sugieren que las relaciones espaciales y temporales son suficientemente relevantes como para justificar el uso de metodologías avanzadas, como el GTWPCA. La Tabla 2 presenta los valores del estadístico y los p -valores de las variables

con mayor *I*-Moran.

Variable	Estadístico	P-Valor	I-Moran
TG3 2022	37.2	2.6e-125	0.58
TG3 2020	35.69	5.2e-137	0.56
D 2022	34.97	1.06e-23	0.52
TG3 2017	32.34	6.49e-47	0.51
TG8 2010	29.49	2.42e-155	0.49

Tabla 2: Resultados de autocorrelación espacial

En el análisis multivariado, los datos se estandarizaron previamente a la aplicación del ACP global convencional, asegurando que las diez variables del estudio, independientemente de sus escalas, tuvieran igual importancia al asignarles varianza unitaria. Este procedimiento se replicó en todas las variantes de ACP, permitiendo una comparación equitativa entre ellas. Dado que estas técnicas no son invariantes a la escala, es crucial estandarizar las variables; de lo contrario, los resultados pueden variar significativamente.

Con el propósito de obtener una visión general de los datos relacionados con delitos y densidad poblacional, se realizó inicialmente un ACP. Este análisis permitió identificar patrones subyacentes y tendencias generales en las dimensiones originales de los datos, proporcionando una base sólida antes de aplicar las técnicas de ponderación geográfica y temporal. Los resultados obtenidos del ACP global proporcionaron un punto de partida para comprender el comportamiento general de los datos, tal como se muestra en la Tabla 3.

	CP 1	CP 2	CP 3	CP 4	CP 5	CP 6	CP 7	CP 8	CP 9	CP 10
Autovalores	3.161	1.681	1.100	1.021	0.926	0.727	0.566	0.367	0.329	0.121
PVT	31.615	16.808	11.003	10.206	9.260	7.275	5.665	3.671	3.289	1.209
Cargas										
TG1	0.214		0.500	0.480	0.127	0.665	0.117			
TG2	0.345	-0.325	0.119			-0.251	0.793		-0.242	
TG3	0.333	-0.384	-0.358	0.124	0.172		-0.184	0.718	-0.105	
TG4	0.390	0.284		-0.189	0.104			0.119		0.736
TG5	0.213	-0.265	0.595	0.129		-0.520	-0.476		-0.108	
TG6	0.481			-0.176				-0.109	0.813	-0.238
TG7	0.386	-0.181	-0.437	0.150		0.142	-0.243	-0.659	-0.285	
TG8	0.325	0.572		-0.107			0.109	-0.382	-0.627	
TG9	0.151	0.132	-0.146	0.494	-0.806	-0.157			0.104	
TG10	0.131	-0.254	0.194	-0.621	-0.527	0.406	-0.189		-0.126	

Tabla 3: Resultados del ACP global. *Nota:* Los valores en blanco corresponden a cargas factoriales menores a 0.1. Los valores en negrita representan las dos cargas más altas en valor absoluto para cada variable.

Los resultados indicaron que los cuatro primeros componentes principales presentaron valores propios superiores a 1, explicando en conjunto el 69.6% de la variabilidad total de los datos donde el primer componente capturó el 31.61%. No

obstante, en este estudio se optó por retener únicamente los tres primeros componentes, ya que estos explican aproximadamente el 60 % de la variabilidad, logrando así una reducción efectiva de la dimensionalidad. Para asegurar la comparabilidad entre las metodologías ACP, GWPCA y GTWPCA, se mantuvieron constantes estos tres componentes principales en todas las técnicas aplicadas. De las cargas de ACP, el primer componente (CP1) parecería estar representado fuertemente por TG4, TG6 y TG7, el segundo componente (CP2) por TG3 y TG8; y el tercer componente (CP3) por TG1 y TG5.

Una vez realizado el ACP se llevó a cabo un GWPCA, cuyo objetivo es analizar la estructura espacialmente heterogénea de los datos multivariantes, permitiendo que los componentes principales varíen a lo largo del espacio geográfico. Para la calibración del GWPCA, se seleccionó un ancho de banda óptimo mediante un enfoque de validación cruzada de “Leave-one-out”. En nuestro caso, este proceso arrojó un ancho de banda óptimo de 2988.

El GWPCA fue calibrado utilizando un ancho de banda adaptativo, lo que permite que el número de observaciones dentro de cada vecindario local varíe en función de la densidad espacial de los datos, manteniendo un balance entre sesgo y varianza en áreas de distinta densidad de observaciones. Posteriormente, se realizó un proceso iterativo para la selección del kernel, evaluando el ajuste del modelo GWPCA con diferentes funciones kernel y seleccionando la que minimizara el estadístico de ajuste GOF. El kernel bicuadrado resultó ser el más adecuado en este caso, al proporcionar el menor valor de GOF.

Se llevó a cabo una prueba de Monte Carlo con 999 simulaciones para evaluar si los valores propios locales correspondientes a cada componente variaban significativamente en el espacio. Los resultados, presentados en la Figura A1, muestran un p valor menor a 0.05 ($p = 0.014$) para los valores propios del primer componente dominante. Dado este resultado, se rechazó la hipótesis nula de estacionariedad espacial de los valores propios locales, lo que indica la existencia de variabilidad espacial significativa en el primer componente.

Para visualizar los resultados del GWPCA, es posible generar m mapas, uno para cada componente principal. Sin embargo, en lugar de graficar directamente los valores propios locales, se siguió el enfoque de Brunson et al. (1996), optando por representar el porcentaje de la varianza explicada para los tres primeros componentes en cada ubicación espacial. Estos tres primeros componentes capturan una porción significativa de la variabilidad espacial, con un rango que oscila entre un mínimo de 60.95 % y un máximo de 95.12 % (Figura 1a). Esto sugiere que, en algunas ubicaciones, los tres primeros componentes explican casi la totalidad de la variabilidad en los datos.

Fotheringham et al. (2015) sugieren que es relevante visualizar la “variable ganadora” en cada componente local, entendida como aquella variable que presenta la mayor carga en valor absoluto en el análisis. En la Figura 1b, se muestra esta visualización para el primer componente local del GWPCA. Los resultados revelan que los delitos contra la vida y la integridad personal (TG7) tienen la mayor influencia en el primer componente en gran parte del oriente del país. En el centro,

los delitos contra la privacidad y la información (TG5) resultan ser la variable dominante, mientras que en el suroeste predominan los delitos contra la libertad y la autodeterminación de la persona (TG4). Por su parte, en el norte y noroeste, la densidad poblacional (TG10) es la variable con mayor carga. Esta distribución refleja una significativa heterogeneidad espacial en la importancia de las variables que componen el primer componente local del GWPCA, mostrando cómo distintas variables dominan en diferentes regiones. A diferencia del ACP global, en el que las variables TG4 y TG8 explicaban el primer componente principal para todo el país, el GWPCA permite capturar mejor la variabilidad espacial de los datos. Esto resalta la importancia de tener en cuenta las diferencias locales en lugar de tratar el fenómeno de forma homogénea, ya que la incidencia de los crímenes varía considerablemente según la región.

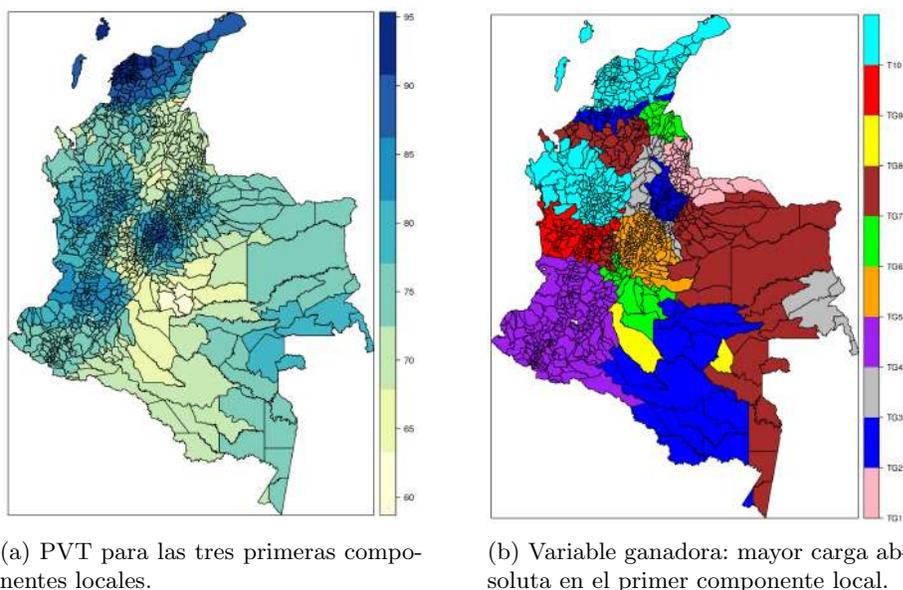


Figura 1: Resultados GWPCA

Después de identificar diferencias geográficas significativas a través del GWPCA, se procedió a incorporar el componente temporal mediante la aplicación del GTWPCA, con el fin de capturar la variabilidad espaciotemporal y proporcionar un análisis más completo y detallado del crimen. Nuevamente, se retuvieron los tres primeros componentes. A diferencia de las metodologías ACP y GWPCA, la técnica GTWPCA no cuenta actualmente con una implementación formal en paquetes de R o Python. Por esta razón, fue necesario desarrollar todos los cálculos y procedimientos analíticos a través de programación personalizada. Se realizó un proceso iterativo para seleccionar el ancho de banda espaciotemporal y el parámetro de balance τ utilizando un esquema de validación cruzada tipo “Leave-One-Out”. Este enfoque basado en la metodología empleada en la función `bw.gwpc` para la

selección de ancho de banda en GWPCA, fue adaptado para incorporar la dimensión temporal. La validación cruzada consistió en excluir una observación en cada iteración y calcular los componentes principales locales para las observaciones restantes, ajustando el modelo de acuerdo con las distancias espaciotemporales entre las ubicaciones y los periodos temporales.

En este procedimiento, se empleó una función kernel bicuadrado y un ancho de banda adaptativo, igual que en el GWPCA, que ponderó las observaciones según su proximidad en el espacio y en el tiempo, considerando que las observaciones más cercanas en ambas dimensiones tienen una mayor influencia. El balance entre las distancias espaciales y temporales, controlado por el parámetro τ , fue ajustado iterativamente para optimizar el modelo, junto con el ancho de banda espaciotemporal. El objetivo de la validación cruzada fue minimizar el error entre las observaciones reales y las proyecciones obtenidas a partir del modelo local en cada punto espaciotemporal. Finalmente, el proceso arrojó un valor óptimo de $\tau = 1$, lo que indica un equilibrio entre la influencia de las distancias espaciales y temporales, y un ancho de banda espaciotemporal de 100, con lo cual se define $\alpha = \beta = 0.5$ (Figura A2).

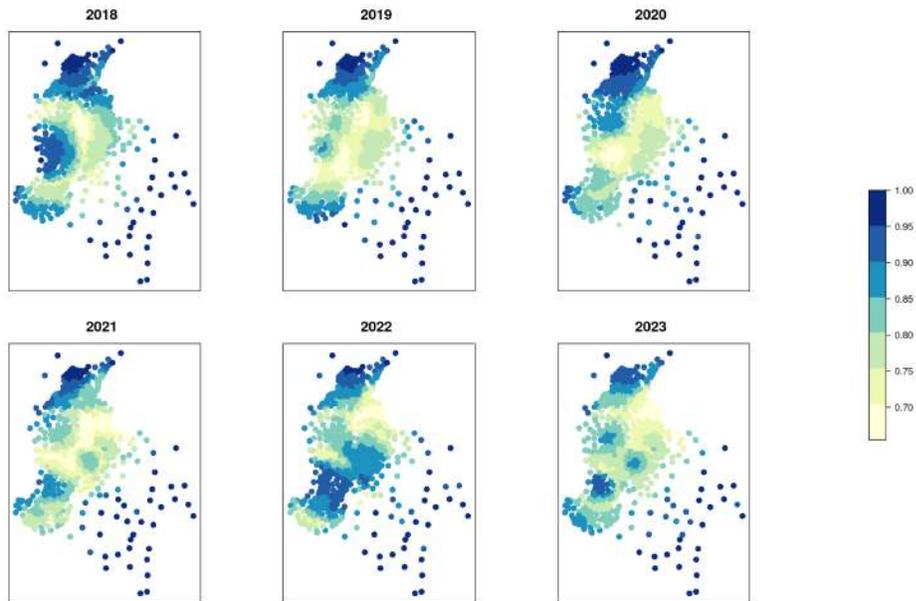


Figura 2: PTV para las tres primeras componentes locales por año (GTWPCA)

Al igual que en el GWPCA se realiza la prueba de Monte Carlo con 999 simulaciones rechazando la hipótesis nula de no estacionariedad en tiempo y espacio con un nivel de significancia del 95 % ($p = 0.038$) (Figura A3). En la Figura 2 se presenta el porcentaje de la varianza explicada por los tres primeros componentes en cada ubicación espacial y para los años analizados. En algunas regiones, estos tres

componentes capturan el 100 % de la variabilidad de los datos, lo que implica una mayor reducción de la dimensionalidad en comparación con el ACP global, donde se requerirían más componentes para alcanzar niveles similares de explicación de la varianza. Este resultado destaca la eficiencia del GTWPCA en la reducción de dimensionalidad, ya que genera estas mediciones de manera localizada para cada punto espacial (u_i, v_i) y cada instante temporal t_i , capturando así variaciones espaciotemporales que el ACP tradicional no puede representar. Se observa una alta estabilidad espacial en la explicación de la varianza, especialmente en el norte y sur oriente del territorio, donde los valores superan consistentemente el 90 %, lo que sugiere que en estas áreas los tres primeros componentes capturan la mayor parte de la variabilidad de los datos durante el periodo analizado.

La Figura 3 muestra la distribución espacial de la “variable ganadora” en el primer componente local a lo largo de los años 2018 a 2023. Se observa una clara heterogeneidad en las variables predominantes según las regiones y el tiempo. En gran parte del territorio, la variable TG1 tiene una influencia significativa, especialmente en el norte y algunas áreas centrales del país. Sin embargo, otras variables como TG2 y TG3 muestran una distribución considerable en el resto del territorio. A medida que avanzan los años, se percibe una mayor diversificación en las variables predominantes, especialmente a partir de 2021, con una mayor presencia de TG5, TG6 y TG7 en zonas específicas. Este patrón resalta la evolución temporal y la variabilidad espacial en la estructura de los datos, indicando que diferentes factores ganan relevancia en distintas ubicaciones y periodos, lo que justifica el uso del GTWPCA para capturar estas dinámicas espaciotemporales.

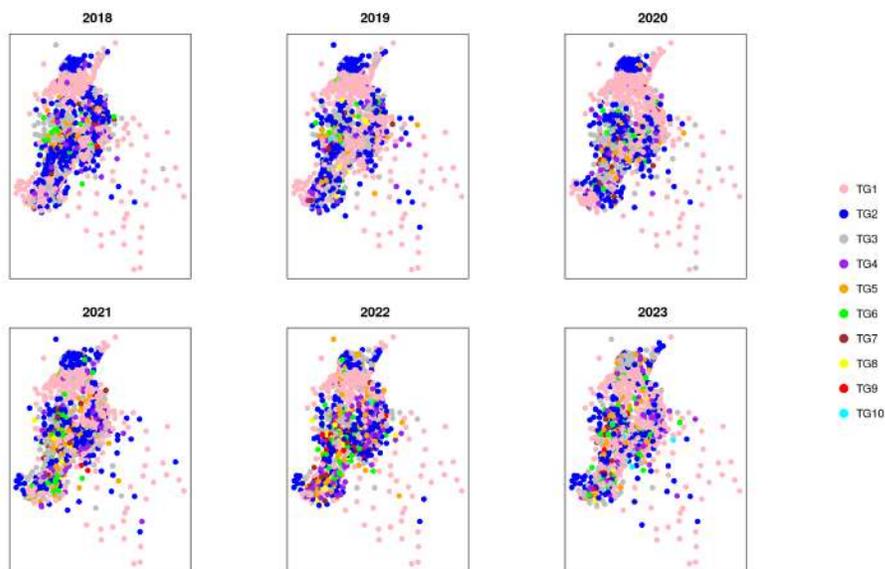


Figura 3: Variable ganadora: mayor carga absoluta en el primer componente local (GTWPCA)

Finalmente se propone un indicador de criminalidad procesada (ICP) a partir de los resultados obtenidos mediante el GTWPCA. Este enfoque permite descomponer la matriz de varianza-covarianza local en cada ubicación espaciotemporal (u_i, v_i, t_i) , capturando la heterogeneidad tanto espacial como temporal de los datos. Mientras que el ACP global asume homogeneidad y estacionariedad, y el GWPCA solo captura la variabilidad espacial, el GTWPCA permite modelar las interacciones entre espacio y tiempo, proporcionando una representación más precisa y adaptada a las dinámicas locales y temporales. Esto asegura que el indicador refleje adecuadamente las fluctuaciones no estacionarias inherentes al fenómeno. La Figura 4 presenta la representación espaciotemporal del indicador directo.

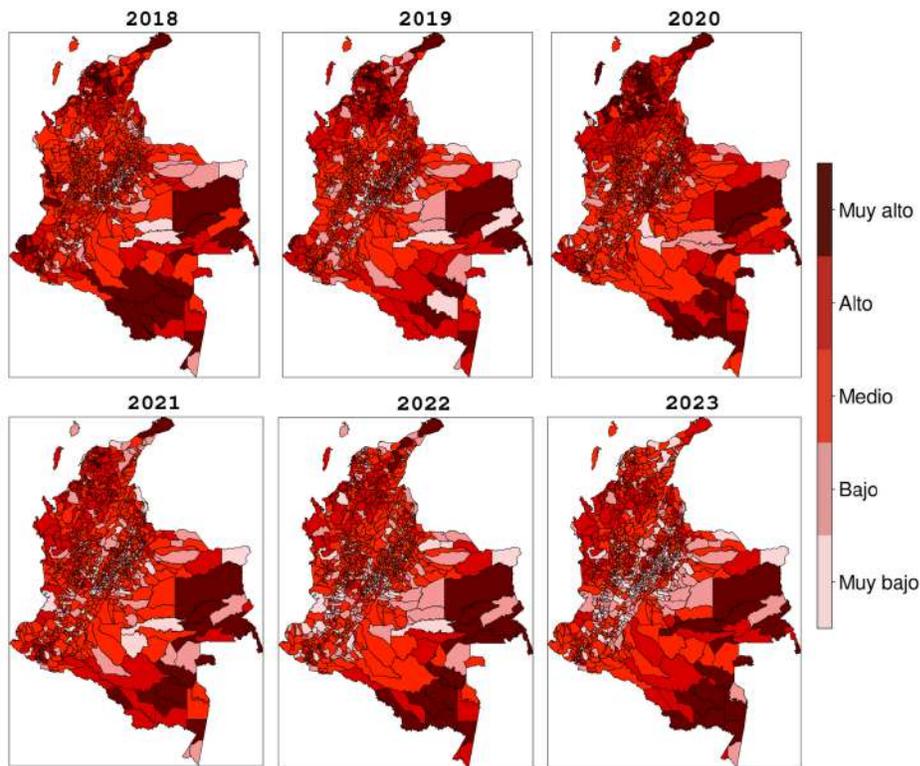


Figura 4: Indicador de criminalidad procesada

El indicador propuesto es reescalado de 0 a 100 y dividido en 5 rangos. Esta reescalación permite una visualización clara de la distribución de la criminalidad procesada en Colombia a lo largo del tiempo. La variabilidad observada en los mapas pone de manifiesto la relevancia de aplicar un enfoque espaciotemporal para analizar patrones delictivos complejos que, de otra manera, no serían detectados mediante metodologías tradicionales globales.

A lo largo del periodo de análisis, se observan cambios significativos en la distri-

bución espacial. Se denota un aumento progresivo de los rangos “altos” y “muy alto” en las regiones del noreste y, en menor medida, en algunas áreas del suroeste. En particular, las zonas centrales del país, que incluyen gran parte de los departamentos de Cundinamarca, Boyacá y Tolima, presentan un comportamiento relativamente estable con valores predominantemente en los rangos “medio” y “alto”, lo que sugiere que estas áreas, aunque afectadas por la criminalidad, no han experimentado fluctuaciones tan drásticas como otras regiones. Por otro lado, la región noroccidental del país, que incluye los departamentos de Antioquia y Córdoba, muestra un patrón mixto, con algunas áreas que han visto una disminución en la criminalidad procesada, moviéndose hacia valores “medio” o incluso “bajos” en ciertos años, lo que indica una posible mejora en las condiciones de seguridad o efectividad de políticas implementadas localmente. En contraste, la región de los Llanos Orientales, que inicialmente mostraba en algunas regiones valores “bajos” para los dos primeros años y se ha visto un incremento sostenido hacia el rango “alto” y “muy alto” en años más recientes, particularmente en áreas cercanas a la frontera con Venezuela, lo que podría estar relacionado con cambios en dinámicas fronterizas y de migración.

Finalmente, las áreas costeras en el norte, especialmente en el Caribe colombiano, también presentan un patrón interesante. Si bien algunas áreas han mantenido valores “medio”, existe un aumento en los rangos “alto” y “muy alto” entre 2021 y 2022, para 2023 se observa una disminución significativa, lo que sugiere posibles mejoras en la seguridad o cambios en las dinámicas delictivas de esta región. En conjunto, estos patrones visuales refuerzan la importancia del enfoque espaciotemporal proporcionado por la metodología GTWPCA, ya que permite observar cómo ciertas regiones experimentan cambios rápidos en su perfil delictivo, mientras que otras muestran una tendencia más estable. Este tipo de análisis detallado facilita la identificación de regiones prioritarias para la implementación de políticas de seguridad ajustadas a sus dinámicas particulares. Este indicador robusto sintetiza la evolución de la criminalidad en distintos municipios a lo largo del tiempo, permitiendo identificar patrones espaciotemporales que no serían detectables mediante enfoques globales, y proporcionando así información más precisa para la toma de decisiones a nivel local.

7. Conclusiones

El ACP es una técnica multivariante ampliamente empleada para la reducción de dimensionalidad y la identificación de patrones en grandes conjuntos de datos. Sin embargo, ignora la dependencia espacial y temporal de los datos, lo que limita su aplicabilidad en contextos donde la variabilidad espacial y temporal es significativa, como en el análisis de datos geográficos y temporales. Para superar esta limitación, en el presente estudio se desarrollaron las metodologías de GWPCA y GTWPCA. Estas técnicas permiten capturar la variación local tanto en el espacio como en el tiempo, proporcionando una comprensión más detallada de los fenómenos estudiados.

El estudio realizado permitió identificar la heterogeneidad espaciotemporal subyacente en los patrones delictivos de Colombia durante el periodo 2018-2023. En este proceso, la incorporación de ponderaciones geográficas y temporales fue fundamental para capturar la no estacionariedad inherente en el fenómeno delictivo, permitiendo la construcción de un indicador robusto y dinámico. Los resultados revelan estructuras no aleatorias en la distribución de los delitos, evidenciando que su incidencia no es homogénea, ni espacial ni temporalmente, destacándose áreas de alta concentración geográfica en núcleos urbanos. Estos hallazgos son consistentes con las hipótesis de autocorrelación espacial, confirmadas mediante el índice de Moran moderado, lo que valida la pertinencia de aplicar modelos multivariantes ponderados espacial y temporalmente. El análisis reveló patrones espaciotemporales diferenciados, con incrementos destacados en los Llanos Orientales, estabilidad relativa en las zonas centrales y disminuciones en el noroccidente. Además, se identificaron fluctuaciones en el Caribe colombiano, con un descenso en los niveles “muy alto” para 2023.

La implementación del GTWPCA permitió descomponer la varianza multivariante de manera localizada, ajustando el análisis a las características espaciales y temporales de cada municipio. Esta descomposición permite una mayor precisión en la captación de las dinámicas espaciotemporales delictivas, las cuales no son accesibles a través de enfoques globales convencionales. El ajuste óptimo de los anchos de banda y kernels seleccionados mediante validación cruzada “Leave-one-out”, maximizó la eficiencia del modelo, proporcionando proyecciones de componentes principales que reflejan fielmente las dinámicas delictivas observadas. La metodología propuesta establece una base matemática sólida para el diseño de políticas públicas focalizadas, optimizando la asignación de recursos mediante la identificación de territorios de mayor riesgo.

La creación del indicador espaciotemporal de criminalidad procesada mediante GTWPCA permitió capturar variaciones significativas en la estructura espacial de los delitos tipificados, revelando zonas de alta criminalidad a lo largo del tiempo. Al integrar variabilidad espacial y temporal, el indicador representa una contribución novedosa al análisis multivariado aplicado a fenómenos socioeconómicos. Es importante destacar que el GTWPCA es una metodología relativamente reciente y aún en desarrollo. Actualmente, no existen librerías formalmente desarrolladas en lenguajes de programación como R o Python para su automatización, lo que limita su aplicación generalizada y conlleva elevados costos computacionales. Este trabajo contribuye al desarrollo de esta técnica al proponer una implementación manual, lo que abre nuevas oportunidades para su expansión y refinamiento, subrayando la necesidad de desarrollar herramientas automatizadas que faciliten su uso en diversos contextos analíticos. Finalmente, se propone un indicador basado en GTWPCA, que debería considerarse como una aproximación innovadora para la medición de la evolución del crimen en Colombia.

Este trabajo representa un avance metodológico significativo al integrar dimensiones espaciales y temporales mediante el GTWPCA, lo que supera las limitaciones observadas en estudios previos que utilizaron enfoques unidimensionales o modelos

espaciales simples (Bello et al., 2014; Ayoola et al., 2015; Summers y Caballero (2017); Aguirre-Rodríguez et al., 2021). En el contexto colombiano, investigaciones como las de Quintero Cuello et al. (2008) y Cetina et al. (2022) desarrollaron índices multivariados para analizar la criminalidad, pero sin incorporar dinámicas espaciotemporales, limitando su capacidad para identificar patrones complejos. En contraste, este estudio genera un indicador robusto y comparable en tiempo y espacio, captando patrones heterogéneos que no eran detectables mediante metodologías anteriores y proporcionando un marco analítico práctico para comprender la criminalidad procesada en Colombia y una herramienta efectiva para orientar estrategias de seguridad adaptadas a las necesidades locales.

Finalmente, trabajos futuros podrían explorar la extensión del GTWPCA en otros contextos y áreas de aplicación, como la epidemiología, la economía regional y la planificación territorial, donde las dinámicas espaciotemporales sean relevantes. Asimismo, sería valioso comparar el desempeño del indicador con otras metodologías no lineales o híbridas que integren técnicas de aprendizaje automático, con el fin de evaluar su robustez en diferentes contextos. Por último, futuros estudios teóricos podrían enfocarse en la optimización del ancho de banda en GTWPCA, ya que este parámetro tiene un impacto significativo en los resultados locales, y su calibración adecuada podría mejorar aún más la precisión y aplicabilidad del indicador en diferentes escalas espaciales.

Recibido: Octubre 21 de 2024
Aceptado: Febrero 10 de 2025

Referencias

- E. Y. Aguirre-Rodríguez, J. P. Santos-Fernández, E. C. Aguirre-Rodríguez, and F. A. Silva-Marins. Análisis espacial de la criminalidad e identificación de clústeres en la provincia de trujillo. *Revista Ciencia y Tecnología*, 17(2):53–61, 2021. URL <https://revistas.unitru.edu.pe/index.php/PGM/article/view/3581>.
- H. A. Andersen and K. Mueller-Johnson. The Danish Crime Harm Index: How It Works and Why It Matters. *Cambridge Journal of Evidence-Based Policing*, 2(1-2):52–69, 5 2018. URL <https://doi.org/10.1007/s41887-018-0021-7>.
- F. J. Ayoola, M. A. Adeyemi, and Jabaru. On the Estimation of Crime Rate in the Southwest of Nigeria: Principal Component Analysis Approach. *Global Journal of Human Social Science*, 15(2), 3 2015.
- Y. Bello, Y. U. Batsari, and A. S. Charanchi. Principal component analysis of crime victimizations in katsina senatorial zone. *International journal of science and Technology*, 3(4):192–202, 2014.
- C. Brunsdon, A. S. Fotheringham, and M. E. Charlton. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographi-*

- cal Analysis*, 28(4):281–298, 10 1996. URL <https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>.
- P. Cadena-Urzúa, A. Iftimi, and F. Montes-Suay. Seguridad ciudadana como elemento fundamental para el desarrollo humano: análisis de delitos en Chile mediante un modelo logístico espaciotemporal. *Revista iberoamericana de estudios de desarrollo = Iberoamerican journal of development studies*, 12(1):30–49, 5 2023. URL https://doi.org/10.26754/ojs_ried/ijds.699.
- Y. A. M. Cetina, D. B. Rubio, and H. Y. R. Pinzón. Índice multivariado de medición del conflicto interno colombiano. *Perspectivas en Inteligencia*, 14(23): 171–200, 12 2022. URL <https://doi.org/10.47961/2145194x.343>.
- M. A. Cohen. The Effect of Crime on Life Satisfaction. *The Journal of Legal Studies*, 37(S2):S325–S353, 6 2008. URL <https://doi.org/10.1086/588220>.
- G. I. A. T. O. Crime. Índice global de crimen organizado 2023, 2023. URL <https://globalinitiative.net/wp-content/uploads/2023/09/1%CC%81ndice-global-de-crimen-organizado-2023.pdf>.
- J. R. B. Cubides and E. N. Céspedes. Registros de la criminalidad en Colombia y actividad operativa de la Policía Nacional durante el año 2015. *DOAJ (DOAJ: Directory of Open Access Journals)*, 8 2016. URL <https://dialnet.unirioja.es/servlet/articulo?codigo=5737185>.
- S. Curtis-Ham and D. Walton. Mapping crime harm and priority locations in New Zealand: A comparison of spatial analysis methods. *Applied Geography*, 86: 245–254, 6 2017a. URL <https://doi.org/10.1016/j.apgeog.2017.06.008>.
- S. Curtis-Ham and D. Walton. The New Zealand Crime Harm Index: Quantifying Harm Using Sentencing data. *Policing A Journal of Policy and Practice*, 12(4): 455–467, 7 2017b. URL <https://doi.org/10.1093/police/pax050>.
- M. de Justicia y del Derecho. Criminalidad y política pública criminal: Propuesta de lineamientos de política criminal para la formulación del plan integral de política criminal 2016-2020, 2017. URL <https://www.politicacriminal.gov.co/Portals/0/documento/CriyPolPubCri.pdf?ver=2017-06-06-094951-850>.
- U. Demšar, P. Harris, C. Brunson, A. S. Fotheringham, and S. McLoone. Principal Component Analysis on Spatial Data: An Overview. *Annals of the Association of American Geographers*, 103(1):106–128, 7 2012. URL <https://doi.org/10.1080/00045608.2012.689236>.
- W. M. Erick, S. W. Wanyonyi, and C. Muchwanju. Application of Principal Component Analysis to Crime Data, Case Study: Mathare Slums, Nairobi County in Kenya. *American Journal of Theoretical and Applied Statistics*, 8(1):7, 1 2019. URL <https://doi.org/10.11648/j.ajtas.20190801.12>.

- S. Farber and A. Páez. A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. *Journal of Geographical Systems*, 9(4):371–396, 8 2007. URL <https://doi.org/10.1007/s10109-007-0051-3>.
- L. E. J. Feijoo. BIPLLOT Logístico Ponderado Espacio Temporal (TSWLB): una aplicación a datos de mortalidad por cáncer de mama en el Ecuador, 1 2023. URL <https://doi.org/10.14201/gredos.153242>.
- A. S. Fotheringham, C. Brunson, and M. Charlton. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley, 10 2002. URL http://www.dlfiles24.com/pdf60/168092213-geographically_weighted_regression_the_analysis_of_spatially_varying_relationships_chris_brunson.pdf.
- A. S. Fotheringham, R. Crespo, and J. Yao. Geographical and Temporal Weighted Regression (GTWR). *Geographical Analysis*, 47(4):431–452, 3 2015. URL <https://doi.org/10.1111/gean.12071>.
- F. García. Construcción de un índice espacial de bancarización : un estudio para la región centro de Argentina. *Panorama Económico*, 28(4):232–241, 10 2020. URL <https://doi.org/10.32997/pe-2020-3327>.
- G. S. García. La construcción social de la realidad. Las ficciones del discurso sobre la impunidad y sus funciones sociales. *Via Inveniendi Et Iudicandi*, 17(1), 1 2022. URL <https://doi.org/10.15332/19090528.7743>.
- I. Gollini, B. Lu, M. Charlton, C. Brunson, and P. Harris. GWmodel: AnRPackage for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software*, 63(17), 1 2015. URL <https://doi.org/10.18637/jss.v063.i17>.
- O. O. Gómez Pinto. *Comprensión dinámica del fenómeno de la criminalidad en Colombia para la evaluación de alternativas de prevención*. Phd thesis, Universidad de Bogotá Jorge Tadeo Lozano, 2022.
- J. Han, X. Kang, Y. Yang, and Y. Zhang. Geographically and temporally weighted principal component analysis: a new approach for exploring air pollution non-stationarity in China, 2015–2019. *Journal of Spatial Science*, 68(3):451–468, 1 2022. URL <https://doi.org/10.1080/14498596.2022.2028270>.
- J. Han, Y. Yang, X. Yang, D. Wang, X. Wang, and P. Sun. Exploring air pollution characteristics from spatio-temporal perspective: A case study of the top 10 urban agglomerations in China. *Environmental Research*, 224:115512, 2 2023. URL <https://doi.org/10.1016/j.envres.2023.115512>.
- P. Harris, C. Brunson, and M. Charlton. Geographically weighted principal components analysis. *International Journal of Geographical Information Science*, 25(10):1717–1736, 8 2011. URL <https://doi.org/10.1080/13658816.2011.554838>.

- P. Harris, A. Clarke, S. Juggins, C. Brunson, and M. Charlton. Enhancements to a Geographically Weighted Principal Component Analysis in the Context of an Application to an Environmental Data Set. *Geographical Analysis*, 47(2): 146–172, 9 2014. URL <https://doi.org/10.1111/gean.12048>.
- I. T. Jolliffe and J. Cadima. Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.
- Y.-A. Kim and J. R. Hipp. Density, diversity, and design: Three measures of the built environment and the spatial patterns of crime in street segments. *Journal of Criminal Justice*, 77:101864, 11 2021. URL <https://doi.org/10.1016/j.jcrimjus.2021.101864>.
- S. Kumar, R. Lal, and C. D. Lloyd. Assessing spatial variability in soil characteristics with geographically weighted principal components analysis. *Computational Geosciences*, 16(3):827–835, 3 2012. URL <https://doi.org/10.1007/s10596-012-9290-6>.
- F. Kärholm, P. Neyroud, and J. Smaalund. Designing the Swedish Crime Harm Index: an Evidence-Based Strategy. *Cambridge Journal of Evidence-Based Policing*, 4(1-2):15–33, 3 2020. URL <https://doi.org/10.1007/s41887-020-00041-4>.
- C. D. Lloyd. Analysing population characteristics using geographically weighted principal components analysis: A case study of Northern Ireland in 2001. *Computers Environment and Urban Systems*, 34(5):389–399, 3 2010. URL <https://doi.org/10.1016/j.compenvurbsys.2010.02.005>.
- B. Lu, P. Harris, M. Charlton, and C. Brunson. The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models. *Geo-spatial Information Science*, 17(2):85–101, 4 2014. URL <https://doi.org/10.1080/10095020.2014.917453>.
- V. J. Mojica, A. Choi, R. N. Leong, and F. Co. Spatial analysis of violent crimes in Metro Manila, Philippines. *International Journal of Comparative and Applied Criminal Justice*, 43(1):29–47, 11 2017. URL <https://doi.org/10.1080/01924036.2017.1398669>.
- E. Norza-Céspedes and G. P. Espino-Duque. Criminalidad y análisis espacial de los delitos en Colombia, 2010. *Revista Criminalidad*, 53(1):17–43, 2011.
- D. Ortega, D. Mejía, and K. Ortiz. Un análisis de la criminalidad urbana en Colombia. *CAF Development Bank Of Latinamerica*, (810), 2015. URL <https://ideas.repec.org/p/dbl/dblwp/810.html>.
- C. A. Pardo. *Criminalidad, conflicto y justicia penal en Colombia*. Universidad Católica de Colombia, 2018. URL <https://repository.ucatolica.edu.co/entities/publication/b8c4b4c6-774a-4122-8c75-6bfaebdea734>.

- M. S. Posada. Violencia Política, Conflicto Social y su Impacto en la Violencia Urbana. *DOAJ (DOAJ: Directory of Open Access Journals)*, 1 1999. URL <https://doaj.org/article/9aa8a37e4f7948fc9f5a9c3f80954ebe>.
- D. M. Quintero Cuello, Y. Lahuerta Percipiano, and J. M. Moreno Acosta. Un índice de criminalidad para colombia. *Revista Criminalidad*, 50(1):37–58, 2008.
- L. Summers and M. Caballero. Spatial conjunctive analysis of (crime) case configurations: Using monte carlo methods for significance testing. *Applied geography*, 84:55–63, 2017. URL <https://doi.org/10.1016/j.apgeog.2017.05.002>.
- J. L. V. Tornero and R. M. B. Andrade. Eficacia de las medidas políticas e índices internacionales de criminalidad: el caso de la república del ecuador. *Gestión y Análisis de Políticas Públicas*, (23):152–166, 2020.
- C. Wu, N. Peng, X. Ma, S. Li, and J. Rao. Assessing multiscale visual appearance characteristics of neighbourhoods using geographically weighted principal component analysis in Shenzhen, China. *Computers Environment and Urban Systems*, 84:101547, 9 2020. URL <https://doi.org/10.1016/j.compenvurbsys.2020.101547>.
- H. Zhao, S. Guo, and H. Zhao. Impacts of GDP, Fossil Fuel Energy Consumption, Energy Consumption Intensity, and Economic Structure on SO2 Emissions: A Multi-Variate Panel Data Model Analysis on Selected Chinese Provinces. *Sustainability*, 10(3):657, 3 2018. URL <https://doi.org/10.3390/su10030657>.

A. Apéndice

GRUPOS	CÓDIGO	DELITOS
Delitos contra el orden social y seguridad pública	TG1	Delitos ambientales Orden económico social Salud pública Seguridad del Estado Seguridad pública Violación medidas sanitarias
Delitos contra la administración pública	TG2	Administración pública Corrupción administrativa Corrupción electoral Corrupción judicial Corrupción privada Corrupción tributaria Falsedad en documento Fe pública Impartición de justicia
Delitos contra la dignidad y los derechos fundamentales	TG3	Actos de discriminación Contra la familia Delitos sexuales Injuria y calumnia Maltrato animal Violencia intrafamiliar
Delitos contra la libertad y la autodeterminación de la persona	TG4	Amenazas Constreñimiento Desaparición forzada Libertad individual y otras garantías Reclutamiento ilícito Secuestro extorsivo Secuestro simple Trata de personas Uso de menores de edad
Delitos contra la privacidad y la información	TG5	Delitos informáticos Derechos de autor
Delitos contra la propiedad y patrimonio económico	TG6	Estafa Extorsión Hurto Patrimonio económico
Delitos contra la vida y la integridad personal	TG7	Aborto Feminicidio Homicidio culposo Homicidio doloso Lesiones por agentes químicos Lesiones personales Lesiones personales culposas Vida e integridad personal
Delitos contra los derechos colectivos y bienes públicos	TG8	Concierto para delinquir Derechos de reunión y asociación Desplazamiento Otros delitos Personas y bienes protegidos por el DIH Régimen constitucional y legal
Delitos relacionados con el uso ilegal de armas y drogas	TG9	Estupefacientes Fabricación, tráfico y porte de armas

Tabla A1: Clasificación de delitos

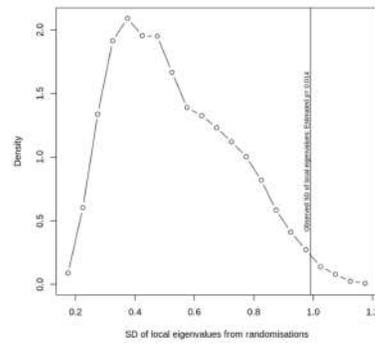


Figura A1: Estadístico de prueba para valores propios no estacionarios GWPCA

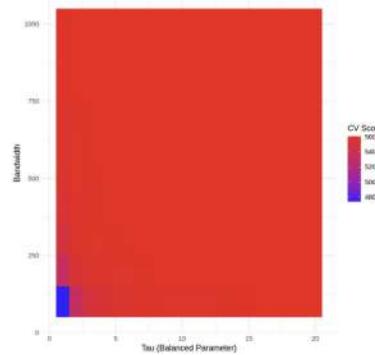


Figura A2: Superficie de validación cruzada

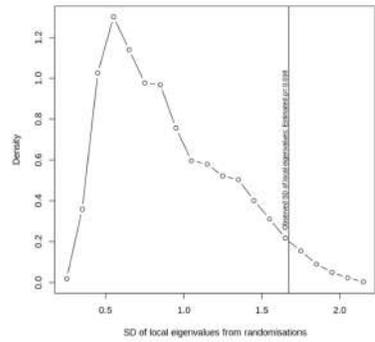


Figura A3: Estadístico de prueba para valores propios no estacionarios GTWPCA