

Análisis geoespacial de transmisión del SARS-CoV-2 haciendo uso de componentes principales para la reducción de dimensiones

Juliana Damelines Pareja¹

Sofía Ríos Oliveros¹

Javier Rubio Sáenz²

Juan David Serrano Medina¹

Kimberly Piñeros Chía³

Diane Moyano Romero⁴

Agradecimientos: Leonardo Salas

Resumen

Las tecnologías de información geográfica y el análisis geoestadístico permiten focalizar la observación analítica en los entornos donde se desarrolla la vida de las personas por medio de los datos cuantificables y localizados que hablan del territorio, orientando las acciones de salud y la gestión de los recursos de manera costo-efectiva. Frente a la emergencia sanitaria, los sistemas de información geográfica visibilizaron la configuración del evento de interés COVID-19 y su progresiva expansión. Se utilizó una metodología de parcelación espacial en Bogotá, por la cual se delimitaron cuadrantes con el fin de observar constantes y medir variables de interés epidemiológico y espacial para el virus SARS-CoV-2. Asimismo, a través de un análisis factorial de componentes principales fue posible agrupar dos grupos de variables que se

definieron como “Transmisibilidad” y “Severidad”. Entre los resultados se obtuvo una escala de “gravedad” y su respectiva medición para cada cuadrante (parcela) que fue calculada semanalmente desde mayo de 2020, y usada hasta la fecha de edición del presente artículo con miras a la búsqueda activa de personas potencialmente transmisoras o que puedan tener complicaciones en su estado de salud debido al COVID-19. Los resultados que se presentan han permitido mejorar la gestión y planeación en salud pública en lo referente a la configuración del COVID-19, así como evidenciar la necesidad de realizar este tipo de análisis como herramienta dirigida a fortalecer la gobernanza en la toma de decisiones.

Palabras clave: análisis factorial, análisis geoespacial, componentes principales, pandemia, SARS-CoV-2, reducción de dimensionalidad, transmisión, severidad.

1. Introducción

Este artículo expone la metodología empleada y los resultados obtenidos del diseño y la construcción del Índice de Transmisibilidad y Severidad (ITS) de la Secretaría Distrital de Salud (SDS) para el seguimiento de la configuración geográfica del SARS-CoV-2 en Bogotá, y con miras al despliegue institucional de acciones de respuesta en los diferentes lugares y sectores de la ciudad.

En el campo de salud pública, el análisis geográfico y espacial ha tenido una connotación de gran relevancia para la observación de la distribución de eventos de interés en un espacio o unidad administrativa. Sin embargo, algunos planteamientos desde corrientes no hegemónicas han expuesto diversas posturas alrededor de la forma en la que se geolocalizan las condiciones de salud en las urbes y cómo se analiza la producción de daños en salud y

1 Profesionales especializados. Dirección de Epidemiología, Análisis y Gestión de Políticas de Salud Colectiva. Secretaría Distrital de Salud.

2 Profesional especializado. Dirección de Planeación Sectorial. Secretaría Distrital de Salud.

3 Profesional especializada. Subdirección de Determinantes en Salud. Secretaría Distrital de Salud.

4 Directora de Epidemiología, Análisis y Gestión de Políticas de Salud Colectiva. Secretaría Distrital de Salud.

las condiciones de bienestar en un territorio a través del tiempo, a fin de develar las condiciones y determinar el proceso de salud-enfermedad sin reducirla a la localización de estos (1, 2). De igual forma, se reconocen las bondades y ventajas que en el transcurso de los años ha conllevado el análisis geoespacial interdisciplinar que involucra la estadística, la salud pública, la epidemiología y la ciencia de datos en torno a los eventos de salud pública, dentro de los cuales se destaca la posibilidad de identificar fenómenos no observados en el espacio geográfico, los cuales determinan la progresividad de una situación particular en el propósito de planear y evaluar intervenciones en salud pública (1–3).

Algunos estudios frente a la pandemia actual generada por el COVID-19 visibilizaron la utilidad de los análisis espaciales que permiten identificar factores dentro de las características urbanas que pueden configurar patrones que favorecen la propagación de la infección (4), develando así el involucramiento tecnológico y la posibilidad de recolección de fuentes de información espaciales para el análisis.

En Colombia, el avance en lo que se refiere a fuentes de información tiene desarrollos limitados en cuanto a la captura, el registro y la estructuración. Aún se cuenta con las problemáticas develadas alrededor de los años sesenta en diferentes estudios, en los cuales la falta de fuentes de información en línea e interoperables entre las acciones públicas y privadas, así como la dificultad de la notificación sistemática y la carencia del acceso de la población a los servicios de manera ágil, son factores determinantes para que el análisis geográfico-espacial dé cuenta de la "realidad" y no de un conteo o una distribución atemporal de la situación (1). Sin embargo, es válido mencionar que el país ha realizado un gran esfuerzo por mejorar estas fuentes de información y, aunque aún nuestras posibilidades de geolocalización no superan

el 80 % de concordancia en el dato, se cuenta con avances que permiten realizar algunas aproximaciones en salud. Sin embargo, se debe mencionar también que los estudios publicados en el país en los que se reflejen estos detalles de la geolocalización son escasos (5–7).

En el presente artículo se expone la metodología empleada para el seguimiento del comportamiento en la configuración espacial del virus SARS-CoV-2 en la ciudad, y cómo dicha configuración permitió converger la respuesta institucional y ciudadana (cuarentena social, los protocolos de bioseguridad, comunicación e información sobre prevención y tratamiento de la pandemia y las políticas de asistencia social a los grupos más vulnerables) en espacios geográficos determinados.

De igual forma, este escrito da cuenta de una necesidad sentida por los equipos de salud pública frente a la planeación de las acciones territoriales necesarias para mitigar el riesgo mostrado en el ámbito internacional a causa de la alta transmisión presentada por SARS-CoV-2. Es importante mencionar que el primer caso de COVID-19 en la ciudad y el país se identificó por el Sistema de Vigilancia en Salud Pública en marzo del 2020; rápidamente, el incremento de casos se presentaba, mientras la incertidumbre y el panorama internacional mostraban afectaciones en diferentes escalas bajo diferentes sistemas de salud. Esto conllevó a la toma de decisiones gubernamentales que tenían un énfasis en el principio de precaución, por lo que se iniciaron de manera temprana cierres y cuarentenas en la ciudad, así como el uso de tapabocas y las medidas denominadas de bioseguridad (lavado de manos, distanciamiento físico, adecuada ventilación), las cuales en la actualidad son aún efectivas (8).

Sin embargo, el análisis geográfico habitual no permitía observar o dimensionar la afectación real. Se conocía que, acorde a la división

administrativa de la ciudad, se contaba con una serie de datos frente a la densidad poblacional, a lo que se sumaba “el escaso” número de casos captados en el sistema (máximo 47 en una Unidad de Planeación Zonal - UPZ) y sus posibles áreas de contactos identificados. Por tanto, era necesario avanzar con una propuesta que permitiera analizar el comportamiento espacial de la pandemia en Bogotá. Teniendo en cuenta esto, el objetivo de este escrito es mostrar los resultados obtenidos luego de un análisis espacial en el que se hace uso de componentes principales con el fin de reducir las dimensiones para el análisis, y conducir a la ciudad y a los tomadores de decisión frente al comportamiento de la transmisión y la severidad.

El análisis factorial de componentes principales ha sido ampliamente utilizado en salud pública. En este caso, se complementa con el desarrollo de una metodología de análisis geoespacial. Algunos autores definen el análisis de componentes principales como “una técnica de representación (descripción) de los datos que se utiliza sin referencia a hipótesis ni a un modelo estadístico particular”, en el que el principal objetivo es reducir el número de variables en un número más pequeño denominado “factores”, los cuales son una combinación lineal de las variables originales que dan máxima información sin ser redundantes (9).

2. Materiales y métodos

El desarrollo metodológico consta de tres momentos para identificar algunos patrones configurativos del evento. Corresponde el primero a definir la línea de partida para identificar y analizar los patrones de concentración espacial de casos; en el segundo momento se genera un procedimiento geoestadístico de análisis factorial por componentes principales con el propósito de resumir la dimensionalidad de

la información y facilitar el estudio y la toma de decisiones; finalmente, el ejercicio analítico de observar la trazabilidad comparada del índice de transmisibilidad y su uso. El periodo de análisis que se muestra en este documento corresponde a los datos desde el inicio de captación de casos de la pandemia generada por el SARS-CoV-2 en Bogotá, de marzo 2020 hasta la semana epidemiológica 53 del 2021. Para la comparación del índice de transmisibilidad-severidad se realiza con los datos obtenidos de la semana 1-53 del 2021, teniendo en cuenta la consolidación de los datos. A continuación, se explica cada uno de los tres momentos.

Momento I: identificar y analizar patrones de concentración espacial de casos

Basados en estudios como los de Olmedo Lucerón (2018) (10) y Mutheneni(2018) (11) se tomaron los casos localizados de COVID-19 para realizar un análisis de correlación por medio del índice de Moran Global, a fin de identificar el grado de interrelación espacial en regiones de la ciudad, basada en ubicaciones y valores de entidades de forma simultánea. A partir de un conjunto de entidades espaciales —como los casos positivos⁵— y un atributo asociado —como la dirección de residencia, la localidad y la UPZ a la que pertenece la dirección— el índice evalúa si el patrón expresado está agrupado, disperso o es aleatorio.

Posteriormente al cálculo de concentraciones, se suscitaba una pregunta de planeación en los equipos de salud pública, el cual era identificar áreas caminables por estos en una unidad geográfica, con el fin de realizar las acciones de educación, información y comunicación en salud pertinentes. Para esto, se utilizó la medida de 500 m, que es una medida

5 Las personas que posterior a una prueba PCR o de antígenos tienen un resultado positivo y son registrados por el Sistema de Vigilancia en Salud Pública (Sivigila).

óptima y una distancia caminable real, puesto que representa aproximadamente diez minutos de caminata, mientras que una distancia de 1.000 metros representa una caminata de veinte minutos a una velocidad urbana promedio de tres kilómetros por hora (12).

El análisis fue realizado con la herramienta ArcGis desktop y su aplicación de puntos calientes G_i^* de Getis-Ord para identificar si existe o no concentración de casos en diversos sectores de la ciudad. En el momento de hacer el análisis de clúster de los casos positivos también se consideró la distancia de 500 m en el parámetro *distance band or threshold distance*. Esta herramienta funciona mediante la búsqueda de cada entidad dentro del contexto de entidades vecinas, indicando a partir de las "puntuaciones z" y los "valores p" resultantes del procesamiento dónde se agrupan espacialmente las entidades con valores altos o bajos. Las "puntuaciones z" y los "valores p" permiten identificar si los registros analizados presentan una concentración o dispersión estadísticamente significativa. Específicamente, las "puntuaciones z" son desviaciones estándar que permiten rechazar una hipótesis nula dependiendo de su puntuación; por su parte, los "valores p" son una probabilidad para observar que el patrón espacial se haya creado mediante algún proceso aleatorio.

A partir del resultado obtenido en el análisis de clúster se procedió a hacer una unión espacial para identificar en la cuadrícula cuántas personas según el censo del DANE podrían habitar, los posibles contactos según el análisis de proximidad en tiempo y distancia⁶, la cantidad de establecimientos comerciales de alimentos y las zonas de aglomeración o de concentración de personas. De esta manera, a cada cuadrante se le identificaron posibles

⁶ Se recopilan a través de proximidad de tiempo y distancia en un momento del tiempo determinado, entre celulares de personas que han sido diagnosticadas como positivas. Esta información se consolida de manera anónima a partir del análisis de redes establecido por Servinformación.

atributos espaciales que tuviesen relación con el comportamiento del evento. A partir de esta información se definió la metodología para calcular el índice de transmisibilidad y severidad, la cual fue un análisis factorial por componentes principales.

Momento II: análisis factorial por componentes principales

Con el fin de determinar uno o varios indicadores que logran resumir la situación dentro de cada cuadrante definido, se exploraron las posibles fuentes de información y las variables que se podían recopilar para cada uno de estos, entre las que se destacan las que se enlistan a continuación.

- » Índice de vulnerabilidad por manzana con el uso de variables demográficas y comorbilidades - DANE, IETS, DNP: este índice permite identificar manzanas de mayor vulnerabilidad; esto implica que son zonas en las que habitan personas que pueden complicarse en caso tal que la transmisión sea alta.
- » Índice de pobreza multidimensional del DANE: permite identificar familias censadas que están en pobreza y pobreza extrema.
- » Información de diferentes sectores de la administración distrital: permite identificar en el territorio zonas de posible aglomeración; la fuente principal es la Infraestructura de Datos Espaciales para el Distrito Capital (Ideca).
- » Servinformación para ubicación de contactos⁷
- » Registros Individuales de Prestación de

⁷ Se recopilan a través de proximidad de tiempo y distancia en un momento del tiempo determinado, entre celulares de personas que han sido diagnosticadas como positivas. Esta información se consolida de manera anónima a partir del análisis de redes establecido por Servinformación.

Servicios (RIPS) para identificación de comorbilidades.

» Con los datos mencionados se organizó la in-

formación conforme a las siguientes categorías que se presentan en la Tabla 1.

Tabla 1. Categorías y tipo de información incluida en el análisis geoespacial

Categoría	Tipo	Fuente
Establecimientos	Constante	Secretaría de Planeación Distrital
Demográficas	Constante	Población Censal DANE 2018
Instituciones educativas	Constante	Ideca-Secretaría de Educación Distrital
Personas positivas de COVID-19, desagregado por su estado	Variables	Sistema de Vigilancia en Salud Pública
Contactos por distancia y tiempo	Variable	Servinformación

Fuente: elaboración propia grupo ACCVSyE.

Un problema central en el análisis de datos multivariantes es la reducción de la dimensionalidad, por lo que, con el objetivo de identificar una medida o indicador que categorizara los cuadrantes en una escala de gravedad en el contexto de la pandemia, se determinó hacer uso de componentes principales (13,14). Esta metodología permite describir con precisión los “valores de p ” variables por un pequeño subconjunto ($r < p$) de ellas, lo que redujo la dimensionalidad a costa de una pequeña pérdida de información. El análisis de componentes principales tiene el objetivo de analizar —si es posible— y representar adecuadamente, dadas n observaciones de p variables, con un número menor de variables construidas como combinaciones lineales de las originales. Por ejemplo, con variables con alta dependencia es frecuente que un pequeño número de nuevas variables (menos del 20 % de las originales) expliquen la mayor parte (más del 80 %) de la variabilidad original (15).

Se planteó hacer un análisis de componentes principales buscando que la rotación

de los ejes garantizara la varianza máxima y se lograra representar la mayor variabilidad. Los valores propios de los componentes principales representan la varianza representada por cada componente, de manera que se puede estudiar la variabilidad absorbida por cada uno y determinar el número de componentes a utilizar.

En efecto, la metodología propuesta es relevante para solucionar el problema, ya que permite reducir cinco variables, las cuales, se sabe, están altamente correlacionadas en pares, a una o dos variables adimensionales que absorben un gran porcentaje de la variabilidad y facilitarán el análisis del estado de los cuadrantes.

Los intervalos para la clasificación de los cuadrantes acorde a los índices obtenidos se construyen a partir de los pasos que se enumeran a continuación.

1. Número óptimo de intervalos: a fin de establecer el número de intervalos a construir se sigue la regla de Sturges, la cual señala el número de clases a considerar a partir de la

siguiente fórmula: $c=1+Log(M)$, donde M es el tamaño de la muestra y el logaritmo con base 2.

2. Puntos de corte: una vez definidos el número de intervalos o clases se establecen los puntos de corte mediante método de cortes naturales (Jenks) que minimiza la desviación de los casos en cada clase y la maximiza entre grupos.

Momento III. Análisis de la trazabilidad comparada del índice de transmisibilidad y uso

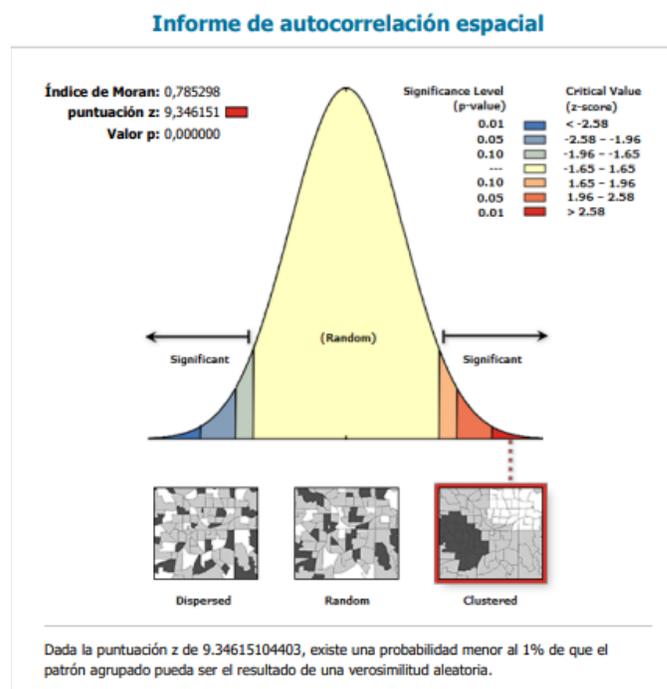
Una vez obtenido el índice, se realiza un seguimiento semanal teniendo en cuenta la disponibilidad de las variables, y se hace el recuento que permite unificar las clases obtenidas. Estas permiten comparar cada cuadrante en diferentes momentos de la pandemia, bajo los mismos parámetros del índice. En lo anterior fue importante conocer el análisis de cada uno de los momentos pandémicos, para lo que se identificaron los casos por semana epidemiológica de los años 2020 y 2021, registrados en la base de positivos por fecha de inicio de síntomas o consulta. Fue necesario procesar la información por semana epidemiológica a fin de que fuera comparable con el seguimiento semanal del índice. De igual forma, teniendo en cuenta el histórico de puntuaciones obtenidos en el índice de transmisibilidad, se realizó una recategorización de las clases según la regla de Sturges y el método de cortes naturales (Jenks). Además, fue necesario en la comparación determinar lo que se determinó como "pico" y "valle", haciendo referencia a las semanas en las que hubo un mayor número de casos por inicio de síntomas o consultas identificado por medio de la representación de la curva y el valle cuando este número de casos descendía.

3. Resultados

Momento I

Se realizó el cálculo del índice de Moran, el cual identificó que existían patrones de agrupación de casos. La puntuación z fue de 9.34615104403, y el valor p fue menor al 1%. Esto indicó que el patrón agrupado de casos COVID localizados pudiera ser el resultado de una verosimilitud aleatoria (Figura 1).

Figura 1. Índice de Moran para COVID-19, Bogotá 2020



Resumen del índice global I de Moran

Índice de Moran:	0,785298
Índice esperado:	-0,001748
Varianza:	0,007091
puntuación z:	9,346151
Valor p:	0,000000

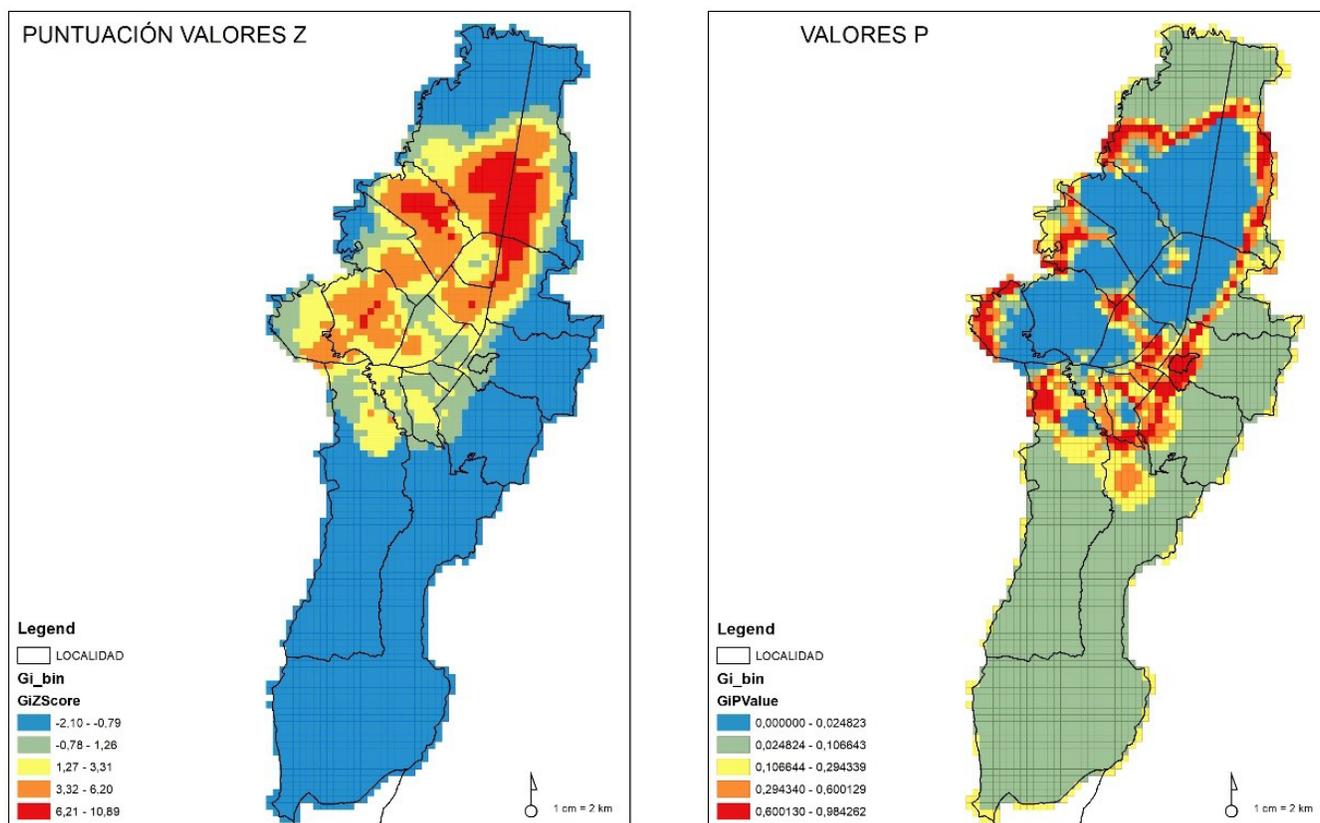
Fuente: elaboración propia a partir de la base positivos COVID-19 SDS.

Posteriormente, se realizó el análisis de densidad de puntos calientes optimizados. En este aná-

lisis se integró las puntuaciones z y los valores p obtenidos a partir del índice de Moran, el cual indicó que es posible rechazar la hipótesis nula, puesto que las puntuaciones z mostraron valores altos en zonas muy específicas de la

ciudad; por su parte, los valores p arrojaron que no existe la probabilidad de observar que el patrón espacial se haya creado mediante algún proceso aleatorio.

Figura 2. Análisis puntos calientes Bogotá, COVID-19, año 2020



Fuente: elaboración propia a partir de los datos obtenidos de base positivos COVID-19, Sivigila 2020, datos preliminares. Corte mayo de 2020.

Los mapas de la Figura 2 exponen los resultados obtenidos de los valores más altos de z y más bajos de p . Estos resultados evidencian que el patrón de concentración de casos ubica a las localidades de Usaquén, Chapinero, Suba, Engativá, Kennedy, Bosa, Barrios Unidos y Teusaquillo con un patrón agrupado, difícilmente producto de una verosimilitud aleatoria y la unión espacial de elementos geográficos que pudiesen tener relación en la transmisibilidad y severidad del COVID-19.

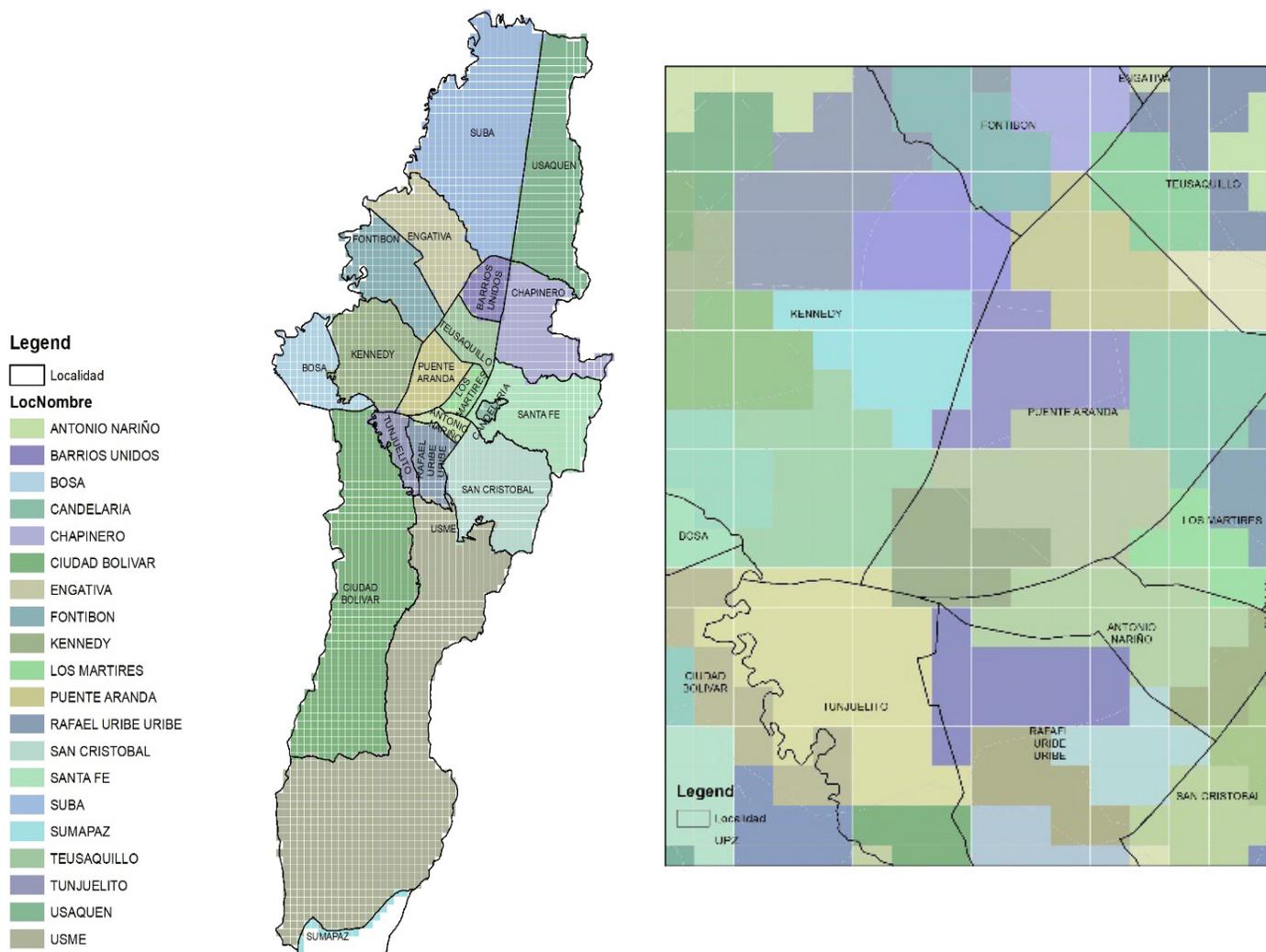
A partir del resultado obtenido en el análisis de clúster, el cual arrojó una cuadrícula que parceló en partes iguales la ciudad (un cuarto de kilómetro cuadrado) en su parte urbana⁸, se procedió a realizar una unión espacial con el fin de identificar cuántos elementos podrían existir de casos diarios, los contactos, el porcentaje de personas mayores a sesenta años, el porcentaje de personas con comorbilidad y el

⁸ Teniendo en cuenta que para el momento del análisis no se habían identificado casos positivos en las zonas rurales, excepto en la UPR Río Tunjuelo.

número de establecimientos en cada uno de los cuadrantes. De esta manera, a cada cuadrante se le identificaron posibles atributos espaciales

que tuviesen relación con el comportamiento del evento (Figura 3).

Figura 3. División de cuadrantes COVID-19 Bogotá-análisis espacial



Fuente: elaboración propia a partir de los datos obtenidos de base positivos COVID-19, Sivigila 2020, datos preliminares. Corte mayo de 2020.

Momento II

A partir de la caracterización de objetos geográficos presente en cada uno de los cuadrantes, se realizó el análisis de las constantes y variables que podrían ser objeto de inclusión en el análisis factorial. En la Tabla 2 se resume la información obtenida para cada cuadrante.

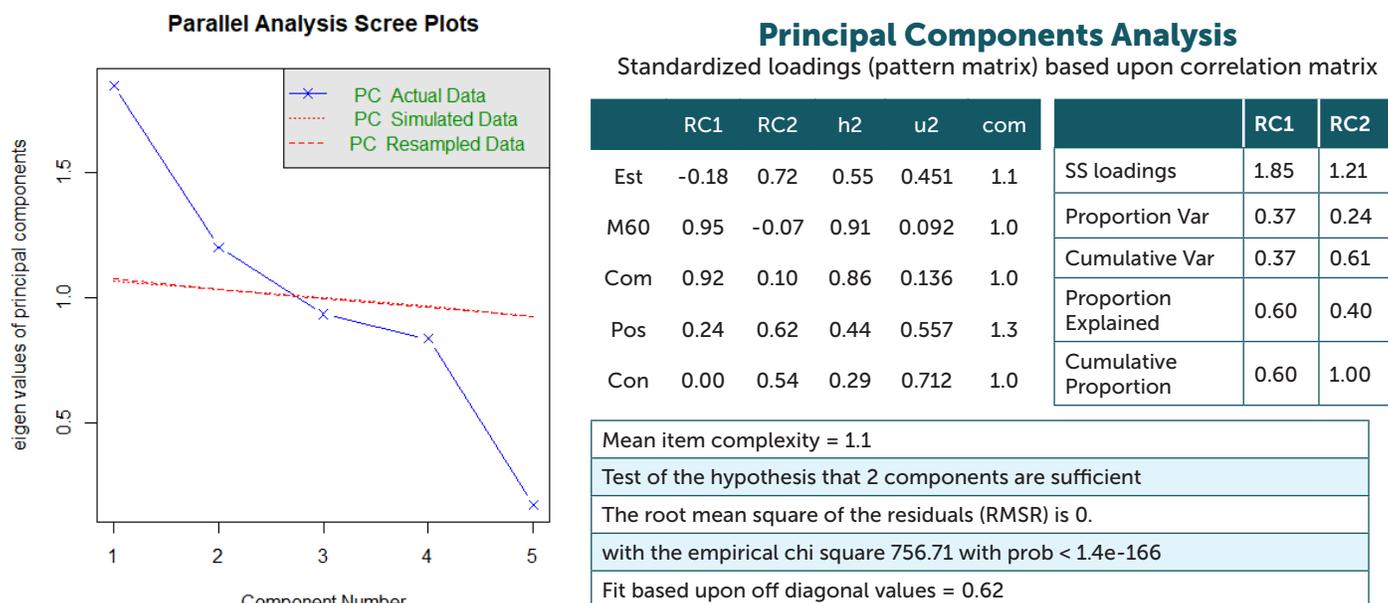
El resultado determinó que las variables que mejor presentaban correlación y dependencia son: casos positivos diarios, contactos, porcentaje de personas mayores a seis años, porcentaje de personas con comorbilidad y número de establecimientos (Figura 4).

Tabla 2. Información obtenida en cada cuadrante según su clasificación

Constantes	Variables-recopilación semanal
<p>Establecimientos: Centros geriátricos, unidades de salud, hipermercados, minimercados, panaderías, supermercados</p> <p>Demográficas: Número de personas por grupos de edad, número de personas por sexo, número de personas por condición de comorbilidad según el índice de vulnerabilidad por manzana, número de personas por condición laboral</p> <p>Instituciones educativas: Número de colegios, número de maestros, número de directivos y número de personal administrativo</p>	<ol style="list-style-type: none"> 1. Casos COVID-19 acumulados 2. Casos activos 3. Número de casos fallecidos 4. Número de casos recuperados 5. Número de contactos (Servinformación) 6. Número de personas con comorbilidad contagiadas 7. Número de personas mayores de sesenta años contagiadas

Fuente: elaboración propia.

Figura 4. Determinación de número de variables a usar

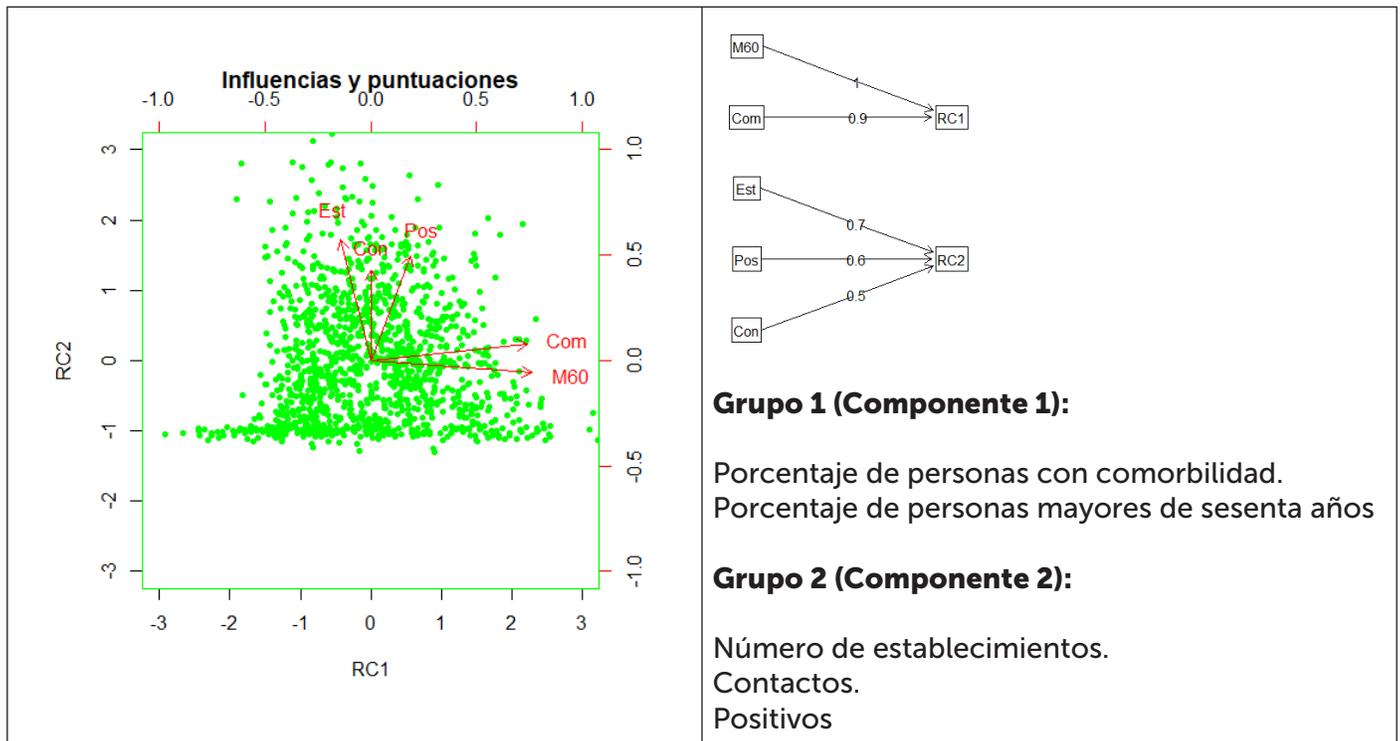


Fuente: elaboración propia.

En el gráfico de la Figura 4 se presentan los valores propios de las componentes. Se puede determinar que dos componentes principales son suficientes para explicar la variabilidad de las cinco variables que mejor presentaron correlación y dependencia. La componente 1 absorbe el 37 %, y la componente 2 el 24 % de la

variabilidad; en su conjunto acumulan el 61 %. Adicionalmente, en el gráfico de influencias y puntuaciones de la Figura 5 podemos observar la posición relativa del centro de gravedad de cada una de las variables en los nuevos ejes rotados RC1 y RC2. Se observa una clara correlación entre dos grupos de variables.

Figura 5. Influencias y puntuaciones, análisis factorial



Fuente: elaboración propia.

El primer grupo altamente correlacionado con el eje RC1 está conformado por las influencias del porcentaje de personas con comorbilidades y mayores de sesenta años en el cuadrante; dado que el porcentaje de personas con comorbilidades y mayores de sesenta años inducen a la severidad con la que puede afectar el COVID-19, se nombró a ese grupo "Severidad" (RC1). El segundo grupo altamente correlacionado con el eje RC2 está conformado por las influencias del número de establecimientos, los contactos con infectados y el número de casos

activos por cuadrante; en razón a que el número de casos, los contactos y los establecimientos inducen a la transmisión del virus, se llamó a ese grupo "Trasmisibilidad" (RC2)⁹.

El resultado es entonces un indicador para la trasmisibilidad (Grupo 2) y un indicador para la severidad (Grupo 1). El indicador es una proyección perpendicular de las variables originales sobre los nuevos ejes definidos. Estas

⁹ La denominación de los factores o componentes del índice surgieron por consenso en el grupo de trabajo, no hacen referencia a una categoría teórica de los mismos.

proyecciones o valores son adimensionales y conforman un indicador para cada grupo de variables (Severidad y Transmisibilidad). La denominación que se le da a los grupos se hace acorde al significado que tienen las variables dentro de cada grupo. La comorbilidad y la población mayor a sesenta años hace referencia a lo gravedad de la enfermedad, mientras que el número de establecimientos, los contactos y los casos positivos hacen referencia a la propagación o transmisión del virus. Entre más pequeño sea el valor de la componente, menor será la medida de “gravedad” para cada grupo. En caso contrario, entre más alto sea el valor de la componente, mayor será la medida de “gravedad” para cada grupo.

Posterior a esto, por medio de la regla de Sturges (16) se determinó el número de intervalos de clasificación del índice, los cuales se representaron a partir de los puntos de corte definidos por la metodología Jenks (17). Esto permitió intensificar los cuadrantes que se encuentran en una peor situación en el momento de la medición respecto a los demás cuadrantes (Figura 6). (Figura 6 ver página 16).

Fase III

Teniendo la trazabilidad del índice de transmisibilidad y severidad desde mayo de 2020, se logró establecer para diferentes momentos de la pandemia el análisis en dos momentos del tiempo: se identificó la necesidad de estandarizar las clases en la medición de las categorías. Lo anterior, en razón a que, si se hace énfasis en la última categoría de cada uno de los mapas para la medición del 30 de octubre, la categoría más “transmisible” cuenta con un rango de 3,8 a 14,10. Sin embargo, para el 9 de mayo se obtiene en esta última categoría 5, 7 a 8, 7 como los cuadrantes más transmisibles. Esto se debe a que el método de Jenks no permite comparar el estado de los cuadrantes

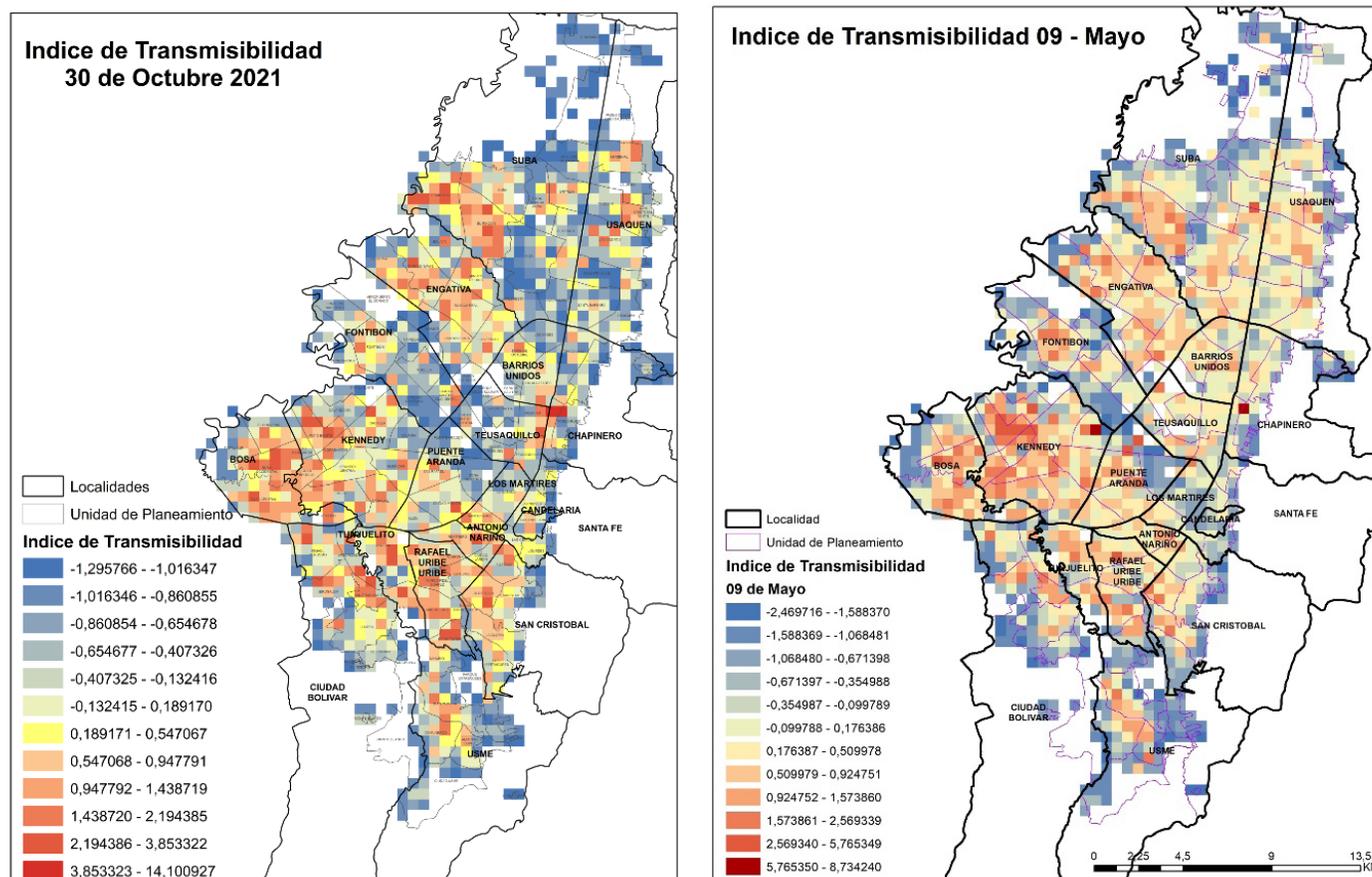
a través del tiempo, ya que los límites de cada clase no son constantes: cambian cada semana de acuerdo con la variabilidad del indicador entre cuadrantes. Por tal motivo, a fin de comparar la evolución de los cuadrantes en Bogotá D. C. haciendo uso de este indicador se construyen las clases de manera convencional.

Momento III

A partir de los resultados del cálculo histórico de los indicadores se definen los límites de clase, de acuerdo con el método convencional en el que todas las clases tienen el mismo ancho a excepción de la primera y última clase. En la Tabla 3 (ver página 16) se muestran los resultados para el índice de transmisibilidad.

Para la comparación del análisis del índice se identificó, a partir de la curva epidémica por semana epidemiológica, cuatro momentos del tiempo. Se escogieron las semanas epidemiológicas que representaron un pico y un valle en cuatro momentos pandémicos y en cada uno de ellos se identificó el comportamiento de los cuadrantes. Esta comparación entre picos y valles se presenta en el mapa de la Figura 7 (ver página 18). La selección de estos picos y valles se realiza teniendo en cuenta la curva construida por semana epidemiológica; es válido mencionar que durante el 2021 se definió como pico las SE 2 y 22 y los valles SE 5 y 44. La definición de estos últimos se realizó según el análisis de olas definido en la ciudad, sin embargo, para el inicio del 2021 este no es epidemiológicamente muy claro, por lo que no hubo un descenso tan claro en el número de casos como al final del año.

Figura 6. Índice de transmisibilidad en dos momentos de tiempo



Fuente: elaboración propia.

Tabla 3. Recategorización de clases para el índice de transmisibilidad

Clase	(]	Marca de clase
C1	-∞	-2,29	-2,51
C2	-2,29	-1,84	-2,06
C3	-1,84	-1,39	-1,61
C4	-1,39	-0,94	-1,16
C5	-0,94	-0,49	-0,71
C6	-0,49	-0,04	-0,26
C7	-0,04	0,41	0,19
C8	0,41	0,86	0,64
C9	0,86	1,31	1,09
C10	1,31	1,76	1,54
C11	1,76	2,21	1,99
C12	2,21	∞	2,44

Fuente: elaboración propia a partir de la trazabilidad del índice de transmisibilidad.

A partir de los datos presentados se puede indicar que existen zonas de la ciudad que, por su mayor número de establecimientos, de casos activos, de contactos, de personas de la tercera edad y de personas con comorbilidades, son constantes y mantendrán una transmisibilidad alta. Esto permite identificar puntos de referencia dentro de la ciudad claves en el análisis, como lo son las zonas de circulación conocida, o los lugares de tipo centralidad ciudadana considerados a la hora de identificar algunos de los puntos, así como dónde llevar diferentes estrategias alusivas a lugares transitados. Con esto se identificaron zonas de severidad y de mayor transmisibilidad para el análisis diferencial, al igual que las estrategias de restricción de la movilidad, acciones intersectoriales y la necesidad de focalizar acciones en salud de manera semanal.

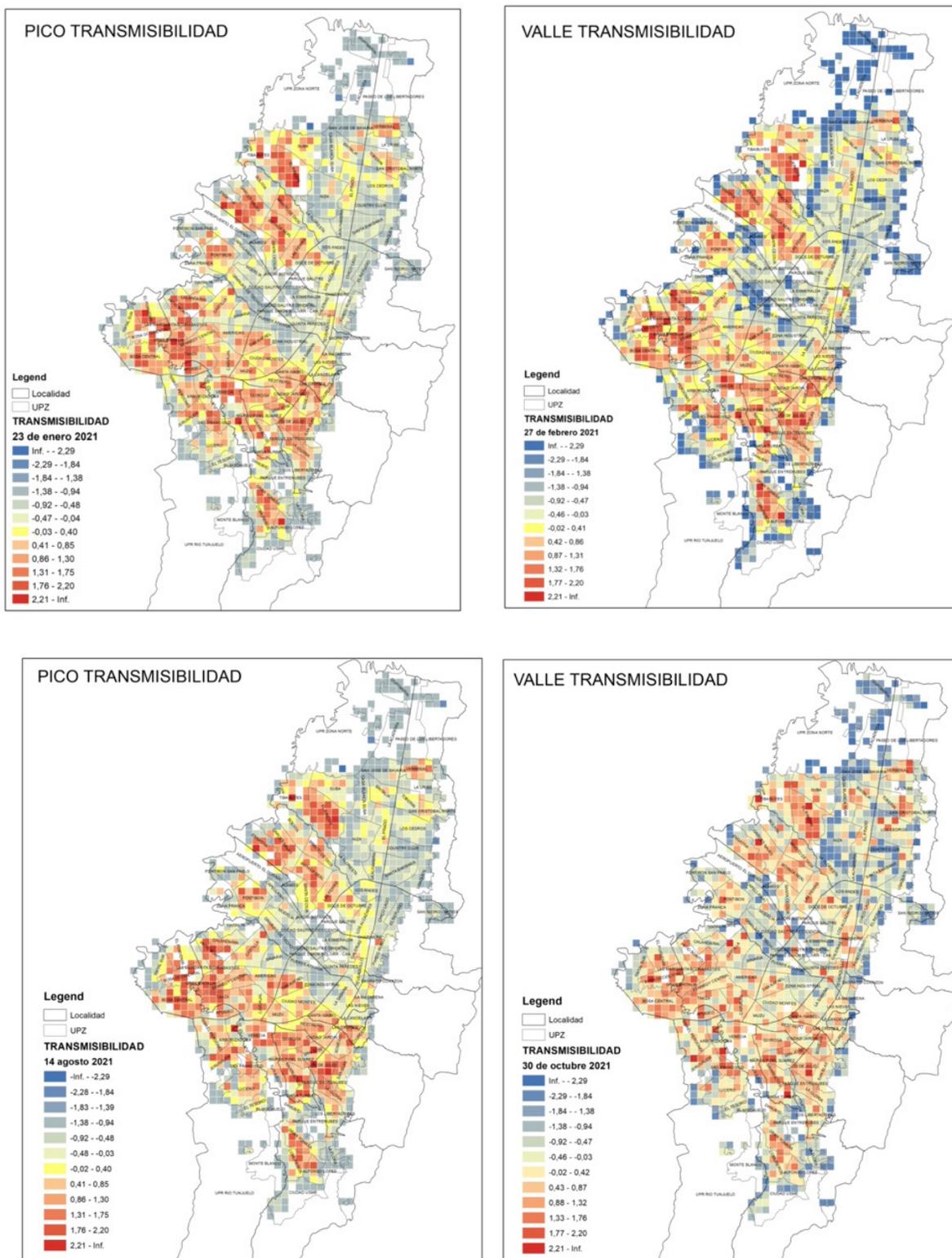
Los puntos de referencia con valores altos y que son constantes permiten observar cómo cambia la transmisibilidad alrededor de estas áreas, puesto que la ciudad tiende a tornarse entre amarillo y rojo, mientras que cuando la transmisibilidad esta baja la ciudad tiende a tornarse de azul y rojo. Este tipo de interpretaciones se ha utilizado como guía para ejercer la acción territorial de la vigilancia en salud pública, por medio de la búsqueda activa de grupos específicos y vulnerables, así como el direccionamiento de acciones de educación, comunicación e información en salud frente al COVID-19.

A partir de la distribución y la concentración espacial identificada fue posible llevar a

cabo acciones de planeación en salud pública. Esta fue diferencial, de acuerdo con el momento epidemiológico. En el inicio de la pandemia en la ciudad fue necesario con miras a identificar las zonas donde se realizaban diferentes acciones territoriales y, de acuerdo con la normativa vigente, se realizaban cuarentenas focalizadas que permitían en una ciudad con la magnitud poblacional direccionar de manera más efectiva los equipos y los recursos humanos definidos en el nivel sectorial e intersectorial para el desarrollo de las acciones de promoción identificadas en el COVID-19.

Posteriormente, con la reactivación de estas concentraciones permitieron que los equipos de Vigilancia en Salud Pública, complementariamente con los entornos de vida cotidiana en la ciudad, realizaran acciones de búsqueda activa con su posterior identificación o tamización para el SARS-Cov-2 debido a la alta externalidad presentada, permitiendo, además, la ubicación de zonas con una configuración histórica, determinada por el número de casos en la ciudad que permitía que los equipos realizaran acciones de educación, comunicación e información en salud de manera más dirigida. Para el caso de la severidad se lograron realizar acciones en los cuadrantes desde la ruta de promoción y mantenimiento de la salud-RIAS, caracterizando a la población, lo que permitió canalizar a población identificada como vulnerable a las RIAS de atención específica acorde a la necesidad.

Figura 7. Diferencias entre picos y valles, índice de transmisibilidad



Fuente: elaboración propia, a partir de datos índice de transmisibilidad.

4. Discusión

Los resultados aquí presentados han permitido mejorar la gestión y la planeación en salud pública en lo referente a la configuración de COVID-19, así como evidenciar la necesidad de realizar y mejorar este tipo de análisis como herramienta para fortalecer la gobernanza en la toma de decisiones (18,19). Una alternativa estadística era plantear un modelo de regresión que explicara el número de casos en cada cuadrante y establecer como variables explicativas las demás anteriormente expuestas que contribuyeran y mostraran dependencia. Sin embargo, esta opción fue poco viable ya que el comportamiento de la pandemia responde a un modelo de ecuaciones diferenciales el cual es más adecuado para proyectar escenarios bajo supuestos robustos, determinados por epidemiólogos y expertos en el tema. No obstante, el ejercicio permitió concluir que había 5 variables que presentaban alta correlación entre sí: el número de establecimientos, el número de casos activos, el número de contactos, el número de personas de la tercera edad y el número de personas con comorbilidades por cada cuadrante.

En cuanto a las variables incluidas en el análisis de componentes principales, se evidencia que presentan correlación entre ellas. Aunque exista correlación con el número de casos diarios y las demás variables contempladas, además de la gravedad de la enfermedad en casos de alta vulnerabilidad no es adecuado plantear un modelo de dependencia estadístico, ya que el comportamiento endémico responde a un modelo de ecuaciones diferenciales, tal como se menciona en (20). Además, debido a la incertidumbre de la información o a rezagos en cortos periodos de tiempo, la incertidumbre no permite hacer predicciones con altas certezas ni mucho menos a mediano o largo plazo.

Existen cuadrantes cuyo valor de transmisibilidad generalmente ha estado por encima de 1,76, los cuales en los momentos en los que ha habido valles en la transmisibilidad tienden a concentrarse sobre ellos mismos. En los picos hay una tendencia a dispersarse y aumentar el valor en los cuadrantes vecinos a ellos. Los cuadrantes con los valores superiores a 1,76 y que son reiterativos en las diversas mediciones pueden interpretarse como puntos de referencia asociados a un espacio fijo en el que se pueden observar movimientos en las mediciones y, a su vez, identificar los contextos en los que se proyectan o no determinadas acciones, considerando diversos factores dentro de las características urbanas que pueden configurar patrones de transmisibilidad. Estos resultados pueden estar relacionados con patrones de transmisibilidad estudiados en zonas urbanas. Un estudio exploratorio sobre la relación espacial entre las enfermedades infecciosas emergentes y las características urbanas muestra que la planeación urbana, la densidad y otros factores ambientales que son considerados estáticos pueden estar relacionadas con la aparición de enfermedades infecciosas (20).

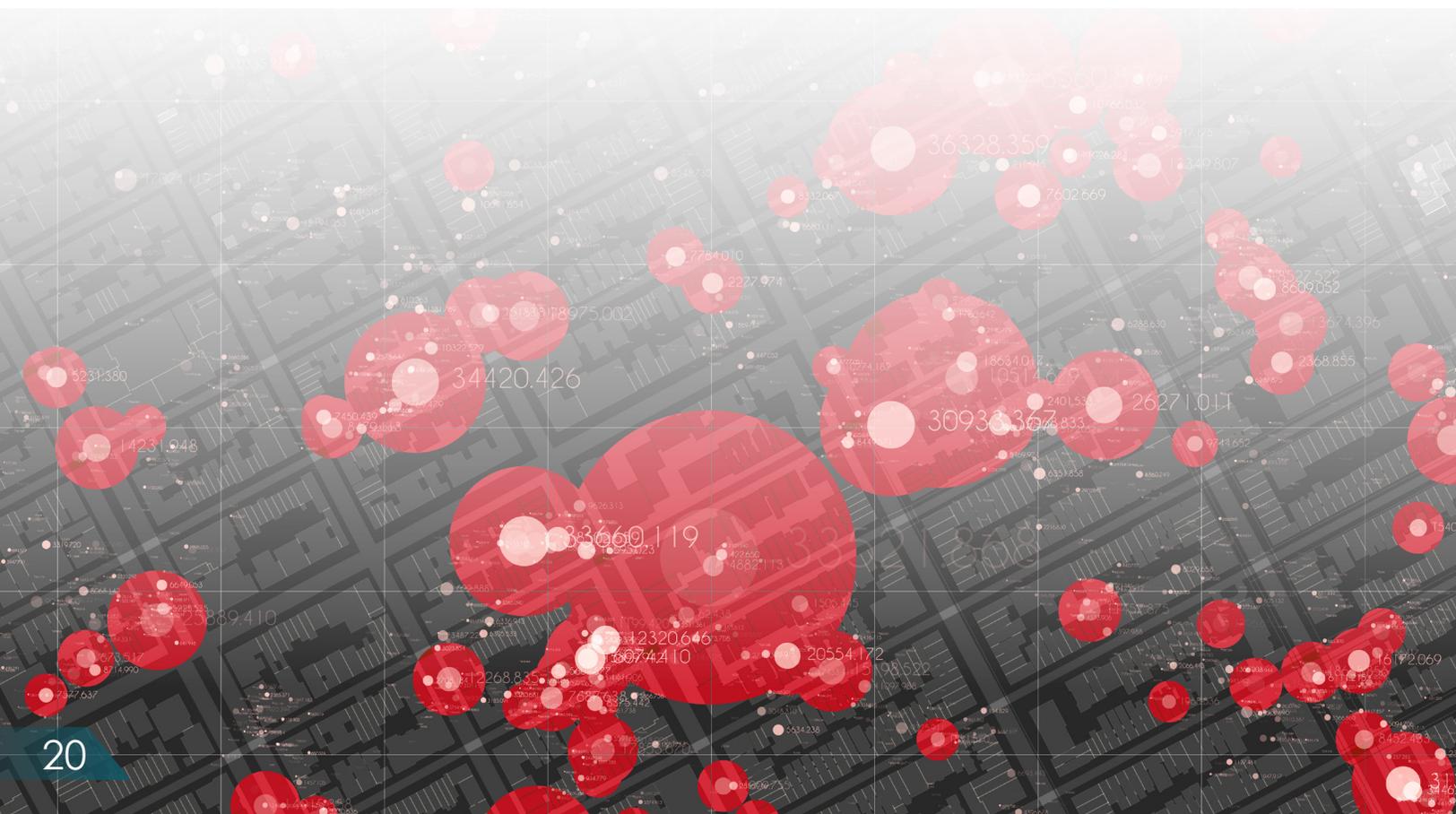
Además, un estudio en particular analizó la geometría urbana y las características socio-demográficas de Hong Kong, en esta ciudad se observan características geográficas en la cual además se presenta una alta densidad poblacional concentrada en partes específicas del territorio dado que tiene una estructura montañosa; en este estudio la densidad y la altura del edificio tiene un factor relevante en la propagación del virus: se analizan las cercanías humanas y lo que mencionan como el "calor humano", en el cual un factor determinante es la falta de ventilación, dado que al ser edificios de veinte o cincuenta pisos no se puede contar con un flujo de aire constante, lo cual supone una connotación de riesgos que fue posible medir y explicar, en parte, la toma de decisiones diferenciales como, por ejemplo,

los cierres o cuarentenas epidemiológicas (21). Si bien en este estudio no se realiza un análisis de las condiciones geográfica y geológica del territorio, permite identificar que tanto en los momentos valles como en los pico existen sectores cuyos valores marcan cifras superiores a 1,76 de transmisibilidad; estos valores pueden interpretarse en donde las categorías dependientes se ven remarcadas por los contactos y los casos registrados en el cuadrante, de manera que permiten observar posibles zonas de circulación conocida. Es importante para futuros análisis determinar características físicas del territorio y conocer si en Bogotá tienen una relación con la propagación de enfermedades infecciosas.

Estar en capacidad de generar una información territorial en salud sobre la configuración del evento COVID-19 en la ciudad

por medio de representaciones ideográficas —como, por ejemplo, la cartografía— facilitaron la comprensión local de la transmisión. Adicionalmente, se logró por medio del índice establecer puntos de intervención en Salud Pública, generando en los tomadores de decisión, principalmente en las alcaldías locales y las secretarías del Distrito, confianza y delimitaciones que les permitieron de manera complementaria apoyar las medidas de mitigación establecidas en el Distrito.

Es importante fortalecer el análisis espacial de los eventos propagados en zonas urbanas como lo es Bogotá, pues esto le permitirá mantenerse a la vanguardia y realizar acciones de planeación urbana; además, permite direccionar las estrategias de salud pública de manera más precisa y diferencial.



Anexos

Anexo 1. Relación de la densidad poblacional y configuración del evento por COVID-19 en Bogotá

Figura A1. Densidad poblacional por UPZ y casos COVID-19 identificados Corte a abril de 2020.

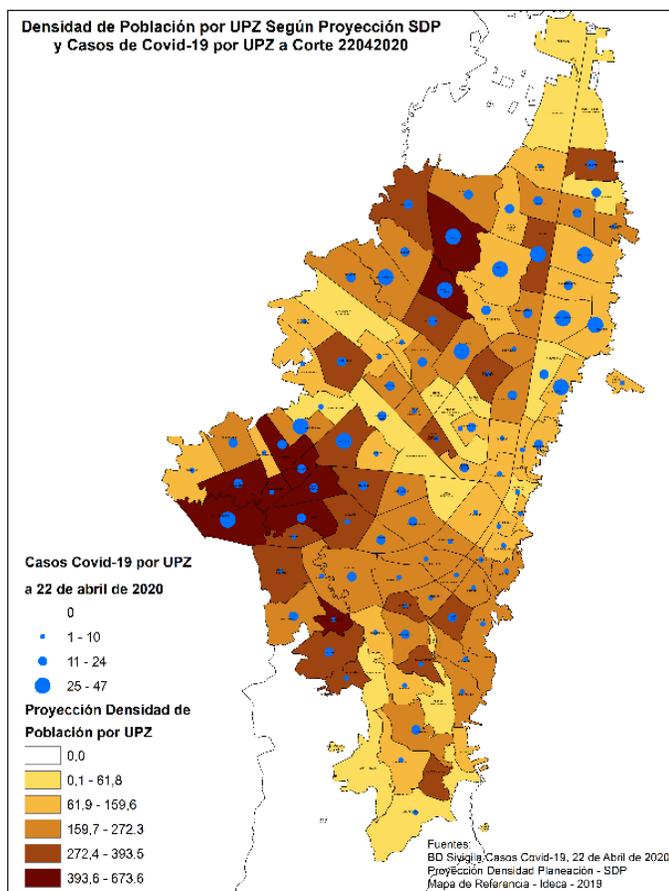
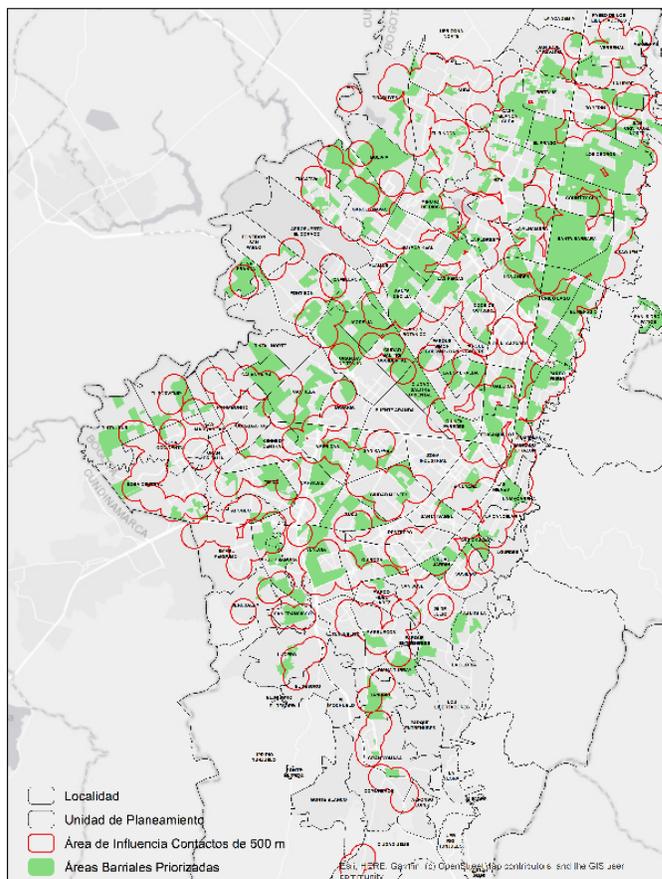


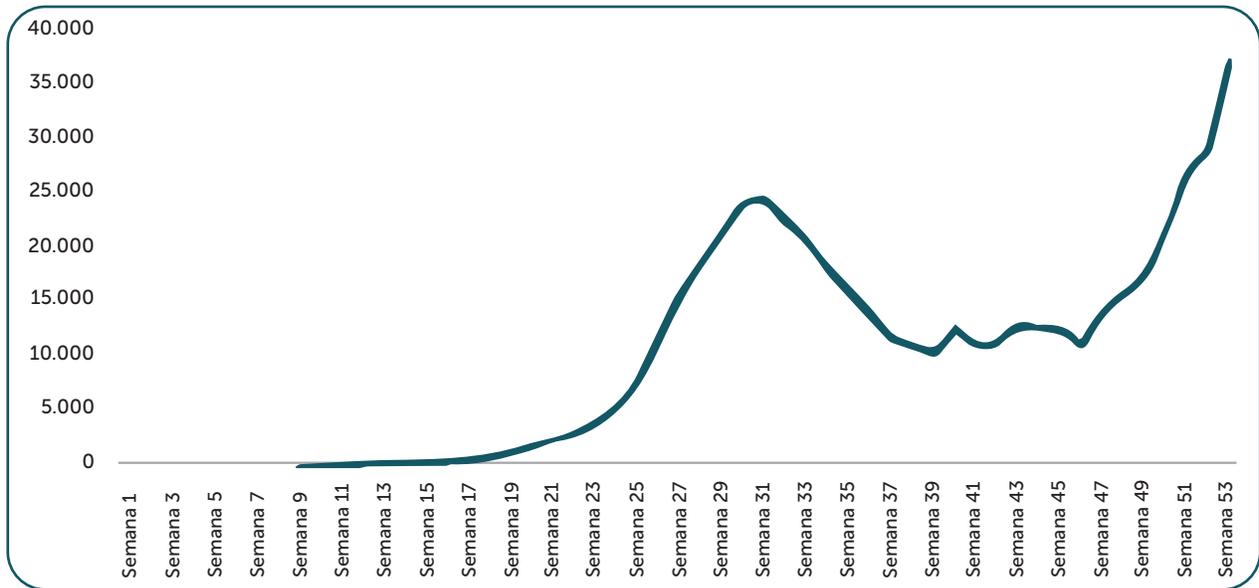
Figura A2. Área de influencia de contacto de casos reportados. Corte abril 2020



Fuente: elaboración propia grupo ACCVSyE a partir de datos base positivos COVID-19, Sivigila 2020, datos preliminares. Censo DANE 2018 proyección población 2020.

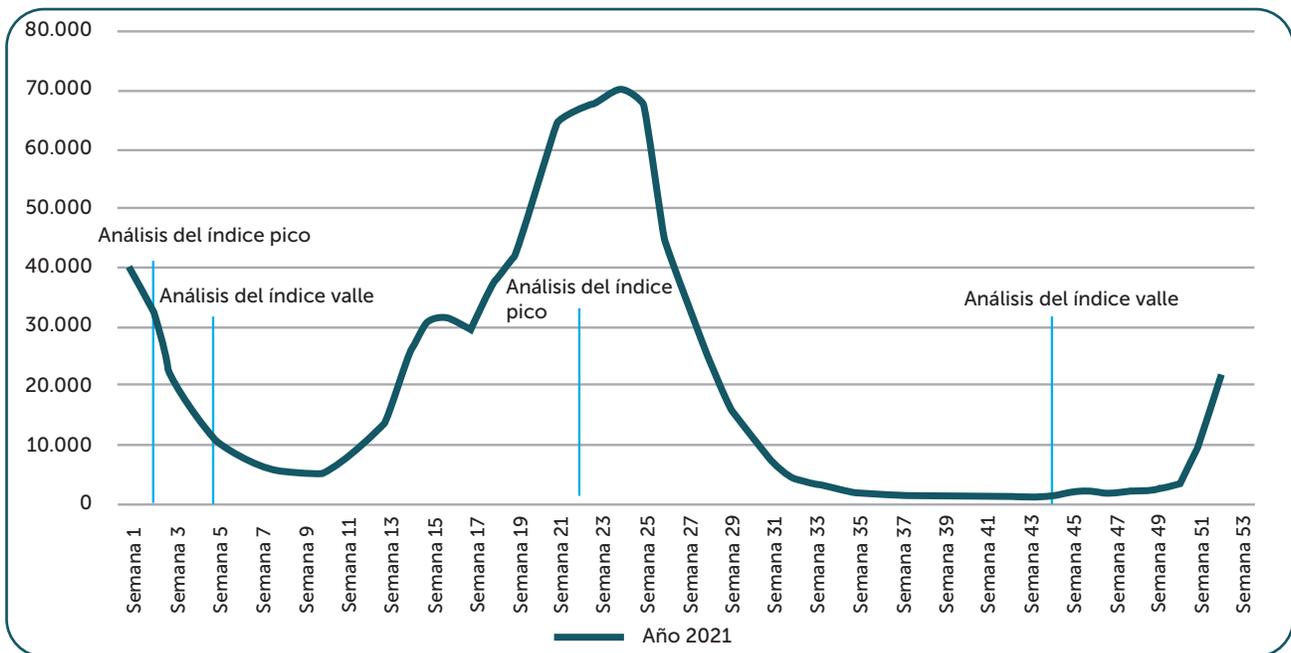
Anexo 2. Curvas epidémicas por año para la identificación de picos y valles

Figura A3. Casos por semana epidemiológica según fecha de inicio de síntomas o consulta 2020 Bogotá SARS-CoV-2



Fuente: elaboración propia grupo ACCVSyE a partir de datos base positivos COVID-19, Sivigila 2020, datos preliminares.

Figura A4. Casos por semana epidemiológica según fecha de inicio de síntomas o consulta 2021 Bogotá SARS-CoV-2



Fuente: elaboración propia grupo ACCVSyE a partir de datos base positivos COVID-19, SIVIGILA 2021, datos preliminares.

Referencias

1. Íñiguez Rojas L, Barcellos C. Geografía y salud en América Latina: evolución y tendencias. *Rev Cuba Salud Pública*. 2003;29(4):330–43.
2. Breilh J. La epidemiología crítica: Una nueva forma de mirar la salud en el espacio urbano. *Salud Colect*. 2010;6(1):83–101.
3. Shrestha S, Bauer CXC, Hendricks B, Stopka TJ. Spatial epidemiology: An empirical framework for syndemics research. *Soc Sci Med* [Internet]. 2020;(August):113352. Available from: <https://doi.org/10.1016/j.socscimed.2020.113352>
4. Rezaeian M, Dunn G, St. Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: A multidisciplinary glossary. *J Epidemiol Community Health*. 2007;61(2):98–102.
5. Tamayo Barrera JA. Estimación de la gestión geoespacial en Colombia. *Análisis geográficos* [Internet]. 2019 [cited 2022 Jul 25];1(56):45–69. Available from: <https://bit.ly/3PUbjaw>
6. Junta extremadura. Consejería de Economía de infraestructuras., Fondo Europeo de Desarrollo Regional. Tecnologías de la información Geográfica en el análisis espacial. Aplicaciones en los sectores público, empresarial y universitario. Nieto Masot, A. Grupo de investigación en Desarrollo Sostenible y planificación territorial de la Universidad de Extremadura., Grupo de Investigación Geo-Ambiental de la Universidad de Extremadura., editors. Barcelona, España ; 2016. 1–283 p.
7. Castellanos-Barrero DF. Aplicación de los sistemas de información geográfica en el ordenamiento territorial. *Vent informática*. 2010 Jun;22:39–53.
8. Rios Oliveros S, Salas Zapata L, Moyano Romero D. Intervenciones no farmacológicas para enfrentar la pandemia de COVID-19 en Bogotá D.C. [Internet]. Vol. 17, *Boletín Epidemiológico Distrital*. Bogotá ; 2020 [cited 2022 Jan 20]. p. 4–18. Available from: [http://www.saludcapital.gov.co/DSP/Boletines epidemiologicos/2019/BED_VOL_17_No_12.pdf](http://www.saludcapital.gov.co/DSP/Boletines%20epidemiologicos/2019/BED_VOL_17_No_12.pdf)
9. Buzai GD, Baxendale C. *Análisis espacial con Sistemas de Información Geográfica. Aportes de la Geografía para la elaboración del diagnóstico en el Ordenamiento Territorial*. Servicio de Publicaciones. Universidad de Alcalá; 2009. 20 p.
10. Olmedo Lucerón C, Pérez Meixeira AM, Iriaso Calle A, Aránguez Ruiz E, Abad Sanz I. Evolución y análisis espacial de los casos de leishmaniasis en una zona rural colindante al brote comunitario de Madrid: 2001-2017. *Rev Esp Salud Publica* [Internet]. 2018;92:e201811084. Available from: <https://pubmed.ncbi.nlm.nih.gov/30458458>
11. Mutheneni SR, Mopuri R, Naish S, Gunti D, Upadhyayula SM. Spatial distribution and cluster analysis of dengue using self organizing maps in Andhra Pradesh, India, 2011–2013. *Parasite Epidemiol Control*. 2018 Feb 1;3(1):52–61.
12. Institute for transportation and development policy. *Desarrollo orientado al transporte DOT -Estándar* [Internet]. 2017 [cited 2022 Jun 1]. Available from: <https://coreciudades.com/wp-content/uploads/2019/07/Desarrollo-Orientado-al-Transporte-DOT-Estandar.pdf>
13. Escofier B, Pagés J. *Análisis factoriales simples y múltiples. Objetivos, métodos e interpretación*. 1st ed. Universidad del País Vasco, editor. 1992. 1–286 p.
14. Cuadras CM. *Nuevos métodos de análisis multivariante*. [Internet]. Segunda edición. Barcelona ; 2007 [cited 2022 Jan 20]. 1–249 p. Available from: http://www.est.uc3m.es/esp/nueva_docencia/getafe/estadistica/

