



**UNIVERSIDAD
DE ANTIOQUIA**

**MODELOS DE REGRESIÓN DE USOS DEL SUELO PARA LA
CARACTERIZACIÓN ESPACIAL DE LA CONTAMINACIÓN
DEL AIRE POR PM_{2.5} EN LA CIUDAD DE MEDELLÍN-
COLOMBIA, 2018**

Sara Catalina Grisales Vargas

Universidad de Antioquia
Facultad de Ingeniería, Escuela Ambiental
Medellín, Colombia
2020



Modelos de regresión de usos del suelo para la caracterización espacial de la contaminación del
aire por PM_{2.5} en la ciudad de Medellín-Colombia, 2018

Sara Catalina Grisales Vargas

Trabajo de investigación como requisito para optar al título de:
Magíster en Ingeniería Ambiental.

Directora:

Ruth Marina Agudelo Cadavid, PhD en Ingeniería.

Asesor:

Libardo Antonio Londoño Ciro, PhD en Ingeniería.

Universidad de Antioquia
Facultad de Ingeniería, Escuela Ambiental
Medellín, Colombia

2020

A mi familia

Agradecimientos

Durante mi proceso académico, he encontrado muchas personas a quién agradecer por su constante apoyo. De manera muy especial quiero agradecer a mi directora Ruth Marina Agudelo Cadavid, por su confianza y paciencia. También agradezco a mi asesor metodológico Libardo Londoño Ciro, por su gran apoyo en la construcción de este trabajo.

Agradezco a la Universidad de Antioquia, la Facultad de Ingeniería, la Facultad Nacional de Salud Pública, al Grupo Salud y Ambiente y a Colciencias, por financiar mi posgrado y por el crecimiento académico, laboral y personal que llevaron a la consolidación de este trabajo. A los coinvestigadores del proyecto de Colciencias, agradezco el empujón y la idea para investigar la aplicación de esta metodología en nuestro entorno.

Con mis compañeros de trabajo y estudio, me disculpo y agradezco los momentos difíciles que sobrellevaron conmigo. A mis amigos de siempre, mis padres, hermanos y mi perro Gauss, gracias por el apoyo.

Resumen

Actualmente, la contaminación atmosférica es un problema ambiental y de salud pública, de especial interés por su consecuente afectación a la salud humana. En los últimos años, se han presentado alertas ambientales debido a los niveles de $PM_{2.5}$ en la ciudad de Medellín, Colombia, y aunque se tiene un sistema de vigilancia de calidad del aire con estaciones continuamente activas, se establece una limitación relacionada con la distribución espacial de la misma. Con el fin de caracterizar espacialmente la contaminación del aire por $PM_{2.5}$ en la ciudad de Medellín, para el año 2018, se construyeron modelos de regresión de usos del suelo (*Land Use Regression - LUR*), a partir de variables meteorológicas, demográficas, de movilidad y de usos del suelo urbano. Se seleccionaron doce modelos mensuales que cumplieran los criterios de especificación del método de Mínimos Cuadrados Ordinarios (OLS), donde, un buffer de 200 m de radio, la altura a nivel de piso de la estación de monitoreo, el flujo vehicular, el gradiente de temperatura y la velocidad del viento promedio mensual, fueron las variables más comunes. Los modelos seleccionados explicaron entre 26% y 79% de la variabilidad del $PM_{2.5}$, siendo el mejor modelo LUR, el obtenido para el mes de agosto. Según los mapas de predicción, las áreas más contaminadas se encontraron al sur y las áreas menos contaminadas se encontraron en el noroeste del área de estudio. La inclusión de variables que representaban fuentes de emisión y dispersión de contaminantes, permitió visualizar la influencia del flujo vehicular y la meteorología en los niveles de $PM_{2.5}$. La metodología LUR es una alternativa simple y replicable para estimar la exposición a partículas; sin embargo, este método es susceptible a sitios de monitoreo limitados.

Palabras clave: Modelo de regresión de usos del suelo (LUR), caracterización espacial multivariada, contaminación del aire.

Abstract

Nowadays, air pollution is an environmental and public health problem of special interest due to its impacts on human health. In recent years, $PM_{2.5}$ levels in the city of Medellin have exceeded Colombian threshold, and although there is an active air quality monitoring system, an important limitation related to spatial distribution is associated. In order to characterize the spatial distribution of fine particles in the city of Medellin for 2018, monthly Land Use Regression Models (LUR) were developed based on meteorological, demographic, traffic and urban land uses data. Twelve monthly models were selected based on the specification criteria of the Ordinary Minimum Squares (OLS) method, where, a buffer of 200 m radius, the sampler height, the monthly average traffic intensity, the monthly average temperature gradient and the monthly average wind speed, were the most common variables. The selected models explained between 26% and 79% of the variability of $PM_{2.5}$, being the best LUR model, the one obtained for the month of August. Based on the prediction maps, the most contaminated areas were found in the southern region of the study area and the least contaminated areas were found in the northwest region. The selected models included variables that represented emission sources and pollutant dispersion, showing the influence of vehicle fleet and meteorology on $PM_{2.5}$ levels. The LUR methodology is a simple and replicable alternative to estimate exposure to particulate matter; however, this method is susceptible to limited measurements sites.

Key words: Land Use Regression (LUR), spatial characterization, air pollution.

Tabla de contenido

1.	Introducción.....	15
2.	Planteamiento del problema.....	16
2.1.	Antecedentes.....	16
2.2.	Descripción del problema.....	20
3.	Objetivos.....	22
3.1.	Objetivo general.....	22
3.2.	Objetivos específicos.....	22
4.	Marco de referencia.....	23
4.1.	Marco teórico.....	23
4.1.1.	Contaminantes atmosféricos.....	23
4.1.2.	Usos del suelo urbano.....	28
4.1.3.	Caracterización espacial.....	29
4.1.4.	Modelos de regresión de usos del suelo – LUR.....	37
4.2.	Marco legal.....	39
4.2.1.	Marco jurídico.....	39
4.2.2.	Marco normativo.....	40
5.	Metodología.....	41
5.1.	Generalidades del estudio.....	41
5.2.	Diseño experimental y metodológico.....	43
5.3.	Momento metodológico uno: Análisis exploratorio de las variables.....	45
5.3.1.	Recolección de información.....	45
5.3.2.	Análisis exploratorio de variables ambientales y aforos vehiculares.....	56
5.3.3.	Criterios de selección de estaciones.....	56
5.3.4.	Excedencias a la norma de calidad del aire.....	57
5.4.	Momento metodológico dos: Preprocesamiento de las variables.....	57
5.4.1.	Interpolaciones espaciales.....	57
5.4.2.	Población.....	59
5.4.3.	Zonas de influencia o buffers.....	59
5.4.4.	Variables predictoras SIG.....	59

5.5.	Momento metodológico tres: Construcción de modelos LUR mensuales	60
5.6.	Momento metodológico cuatro: Selección del mejor modelo LUR mensual.....	62
6.	Resultados	69
6.1.	Momento uno: Análisis exploratorio de las variables	69
6.1.1.	Variable dependiente: PM _{2.5} (µg/m ³).....	69
6.1.2.	Variables ambientales.....	74
6.1.3.	Aforos vehiculares.....	86
6.2.	Momento dos: Preprocesamiento de las variables	91
6.2.1.	Interpolaciones y gradientes espaciales de las variables ambientales.....	91
6.2.2.	Interpolación espacial de aforos vehiculares	95
6.2.3.	Densidad poblacional	96
6.2.4.	Buffers e intersecciones espaciales	97
6.3.	Momento tres: Construcción de modelos LUR mensuales.....	98
6.4.	Momento cuatro: Selección de los mejores modelos LUR mensuales.....	100
6.4.1.	Modelos LUR mensuales	100
6.4.2.	Mapas de predicción y validación de modelos LUR.....	111
7.	Discusión.....	119
8.	Conclusiones	124
9.	Recomendaciones.....	126
10.	Referencias	128
11.	Anexos.....	139

Lista de tablas

Tabla 1. Tipos de SVCA, según número de habitantes	26
Tabla 2. Clasificación de estaciones de monitoreo de calidad del aire	27
Tabla 3. Jerarquía vial de vías urbanas en el municipio de Medellín	29
Tabla 4. Características de los métodos de interpolación espacial	30
Tabla 5. Estadísticos de error calculados para la validación de los datos estimados	36
Tabla 6. Decretos y resoluciones relacionados a la calidad del aire extradomiciliario, Colombia, 2019.	39
Tabla 7. Documentos vigentes relacionados a la calidad del aire extradomiciliario, 2019	40
Tabla 8. Grupos de clasificación de las posibles variables explicativas	43
Tabla 9. Número de modelos construidos teniendo en cuenta los seis buffers explorados	44
Tabla 10. Información necesaria para la construcción de modelos LUR y fuente de información	45
Tabla 11. Estaciones de la Red de Monitoreo de Calidad del Aire del Valle de Aburrá exploradas	47
Tabla 12. Estaciones con registros de temperatura (T), velocidad del viento (Vv) y/o precipitación (P), para el año 2018.....	48
Tabla 13. Cámaras de fotodetección insumo para estimación del flujo vehicular	53
Tabla 14. Niveles diarios máximos permisibles para PM _{2.5}	57
Tabla 15. Variables explicativas evaluadas en la construcción de los modelos LUR	59
Tabla 16. Criterios de especificación del método OLS.....	62
Tabla 17. Tabla resumen con los métodos espaciales empleados en la construcción de modelos LUR mensuales.....	63
Tabla 18. Tabla resumen con los estadísticos empleados en la construcción de modelos LUR mensuales.....	64
Tabla 19. Estadísticos descriptivos para las series de PM _{2.5} diarias, en todo el periodo de análisis	69
Tabla 20. Concentraciones de PM _{2.5} promedio mensual para las estaciones de monitoreo, insumo para la construcción de modelos LUR.....	73
Tabla 21. Estadísticos descriptivos para las series de temperatura promedio diaria, en todo el periodo de análisis	74
Tabla 22. Estadísticos descriptivos para las series de velocidad del viento promedio diaria, en todo el periodo de análisis	78
Tabla 23. Estadísticos descriptivos para las series de precipitación diaria, en todo el periodo de análisis	82
Tabla 24. Estadísticos descriptivos para las series de aforos vehiculares diarios en todo el periodo de análisis.....	86
Tabla 25. RMSE de las interpolaciones espaciales para las variables ambientales	92
Tabla 26. Datos de temperatura promedio mensual en cada estación de monitoreo de PM _{2.5}	92

Tabla 27. Datos de velocidad del viento promedio mensual en cada estación de monitoreo de $PM_{2.5}$	93
Tabla 28. Datos de precipitación acumulada mensual en cada estación de monitoreo de $PM_{2.5}$	93
Tabla 29. Datos de gradiente de temperatura promedio mensual en cada estación de monitoreo de $PM_{2.5}$	94
Tabla 30. Datos de gradiente de velocidad del viento promedio mensual en cada estación de monitoreo de $PM_{2.5}$	94
Tabla 31. Datos de gradiente de precipitación acumulada mensual en cada estación de monitoreo de $PM_{2.5}$	94
Tabla 32. Modelos de semivariograma teóricos ajustados para los aforos vehiculares	95
Tabla 33. Variables explicativas excluidas de la construcción de los modelos LUR	98
Tabla 34. Frecuencias de las variables explicativas en los modelos LUR	99
Tabla 35. Criterios de especificación OLS para los mejores modelos LUR mensuales	100
Tabla 36. RMSE de las interpolaciones espaciales de los valores estimados de $PM_{2.5}$ con los modelos LUR.....	112
Tabla 37. Estadísticos de error estimados de $PM_{2.5}$ para cada modelo mensual.....	117
Tabla 38. Estadísticos de error para cada sitio de monitoreo	117

Lista de figuras

Figura 1. Semivariograma teórico.....	33
Figura 2. Ubicación y mapa de elevaciones de la zona de estudio	43
Figura 3. Momentos metodológicos principales.....	45
Figura 4. Usos industriales, residenciales y zonas verdes en la zona de estudio.....	50
Figura 5. Malla vial de la zona de estudio, insumo para la construcción de los modelos LUR.....	52
Figura 6. Número de habitantes en la zona de estudio	55
Figura 7. Variables fijas para cada mes y para cada buffer	61
Figura 8. Flujograma metodológico para el Momento 1.....	65
Figura 9. Flujograma metodológico para el Momento 2.....	66
Figura 10. Flujograma metodológico para el Momento 3	67
Figura 11. Flujograma metodológico para el Momento 4.....	68
Figura 12. Gráfica de registros de PM _{2.5} en todo el periodo de análisis.....	70
Figura 13. Distribución de los cuartiles para las series de PM _{2.5} diarias seleccionadas.....	71
Figura 14. Número de días con excedencias a la norma nacional de los niveles de PM _{2.5}	71
Figura 15. Promedios mensuales de PM _{2.5} para las estaciones seleccionadas	72
Figura 16. Ubicación estaciones de monitoreo de PM _{2.5} seleccionadas	73
Figura 17. Gráfica de registros de temperatura promedio diaria en todo el periodo de análisis	75
Figura 18. Distribución de los cuartiles para las series de temperatura promedio diaria seleccionadas.....	76
Figura 19. Temperatura promedio mensual para las estaciones seleccionadas.....	76
Figura 20. Ubicación estaciones de temperatura seleccionadas.....	77
Figura 21. Gráfica de registros de velocidad del viento promedio diaria en todo el periodo de análisis	79
Figura 22. Distribución de los cuartiles para las series de velocidad del viento promedio diaria seleccionadas.....	80
Figura 23. Velocidad del viento promedio mensual para las estaciones seleccionadas	80
Figura 24. Ubicación estaciones de velocidad del viento seleccionadas	81
Figura 25. Gráfica de registros de precipitación diaria en todo el periodo de análisis.....	83
Figura 26. Distribución de los cuartiles para las series de precipitación diaria seleccionadas	84
Figura 27. Precipitación acumulada mensual para las estaciones seleccionadas	85
Figura 28. Ubicación estaciones de precipitación seleccionadas.....	86
Figura 29. Gráfica de registros de aforos vehiculares diarios en todo el periodo de análisis	88
Figura 30. Distribución de los cuartiles para las series de aforos vehiculares diarios seleccionados.....	89
Figura 31. Aforos vehiculares mensuales para las cámaras seleccionadas	90
Figura 32. Ubicación de cámaras de fotodetección seleccionadas.....	91
Figura 33. Flujos vehiculares estimados por el método Kriging para enero, junio y diciembre	96
Figura 34. Densidad poblacional para la zona de estudio	97
Figura 35. Intersecciones espaciales para los seis buffers evaluados, en tres estaciones de PM _{2.5}	98

Figura 36. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de febrero	102
Figura 37. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de marzo	103
Figura 38. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de abril	104
Figura 39. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de mayo	105
Figura 40. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de junio	106
Figura 41. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de julio	107
Figura 42. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de agosto	108
Figura 43. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de septiembre	109
Figura 44. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de octubre	110
Figura 45. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de diciembre	111
Figura 46. Mapas de regresión de $PM_{2.5}$ promedio mensual, para el primer semestre del año	113
Figura 47. Mapas de regresión de $PM_{2.5}$ promedio mensual, para el segundo semestre del año ..	114
Figura 48. Concentraciones de $PM_{2.5}$ estimadas de los mapas de regresión y valores promedio mensuales observados	116

Lista de anexos

Anexo A. Porcentaje de valores perdidos mensuales para para las estaciones de PM _{2.5}	139
Anexo B. Porcentaje de valores perdidos mensuales para las estaciones de temperatura	139
Anexo C. Porcentaje de valores perdidos mensuales para las estaciones de velocidad del viento	140
Anexo D. Porcentaje de valores perdidos mensuales para las estaciones de precipitación	140
Anexo E. Porcentaje de valores perdidos mensuales para los aforos vehiculares	141
Anexo F. Criterios de especificación de los 72 modelos LUR construidos	142
Anexo G. Estadístico de error ARE para cada sitio de monitoreo de PM _{2.5}	147

Sigla	Significado
LUR	<i>Land Use Regression</i> – Regresión de Usos del Suelo
PM _{2.5}	Material particulado con diámetro aerodinámico menor a 2.5 micrómetros
SCVA	Sistema de Vigilancia de Calidad del Aire
IDW	<i>Inverse Distance Weighted</i> – Distancia Inversa Ponderada
OLS	<i>Ordinary Least Squares</i> – Mínimos Cuadrados Ordinarios
GWR	<i>Geographically Weighted Regression</i> – Regresión Geográficamente Ponderada
µm	Micrómetros
SIATA	Sistema de Alertas Tempranas del Valle de Aburrá
AMVA	Área Metropolitana del Valle de Aburrá
RMSE	<i>Root Mean Square Error</i> - Error Cuadrático Medio
ASE	<i>Average Standard Error</i> - Error Estándar Promedio
RMSS	<i>Root Mean Square Error Standarized</i> - Error Cuadrático Medio
ARE	<i>Absolute Relative Error</i> – Error Absoluto Relativo
MARE	<i>Mean Absolute Relative Error</i> – Error Absoluto Relativo Promedio

1. Introducción

En este trabajo se realizó una caracterización espacial multivariada de la contaminación atmosférica por $PM_{2.5}$, para el año 2018, por medio de la construcción de modelos de regresión de usos del suelo (LUR). Se definió como variable dependiente las concentraciones promedio mensuales de $PM_{2.5}$ en diferentes estaciones de monitoreo. Además, se establecieron seis zonas de influencia o buffers circulares de diferentes radios, con centro en las estaciones de monitoreo de $PM_{2.5}$; para intersecar con las posibles variables explicativas, las cuales se dividieron en cinco grupos: dos variables de posición (Grupo 1), seis variables ambientales (Grupo 2), seis variables de usos del suelo (Grupo 3), cinco variables de movilidad (Grupo 4) y una variable demográfica (Grupo 5). Las variables ambientales variaban para cada mes del año 2018, las otras variables variaban según el buffer definido y el flujo vehicular variaba para cada mes y cada buffer, dando en total 210 posibles variables explicativas.

La metodología fue dividida en cuatro momentos y se presenta esquemáticamente por medio de cuatro flujogramas, cada uno representando uno de los momentos metodológicos, listados a continuación:

- Momento 1: Análisis exploratorio de variables (Figura 8)
- Momento 2: Preprocesamiento de variables (Figura 9)
- Momento 3: Construcción de modelos LUR mensuales (Figura 10)
- Momento 4: Selección de los mejores modelos LUR, para cada mes (Figura 11)

2. Planteamiento del problema

2.1. Antecedentes

Diferentes estudios a nivel mundial han asociado la exposición a contaminantes atmosféricos con un posible deterioro de la salud del ser humano. Generalmente, para realizar este tipo de asociaciones, se parte de información sobre admisiones hospitalarias y se relacionan a los niveles del contaminante a evaluar, a lo largo del tiempo, en donde autores han encontrado que incrementos en los niveles de material particulado pueden asociarse a aumentos en el riesgo de ingresos hospitalarios diarios por enfermedades respiratorias y cardiovasculares (Camilo Blanco-Becerra et al., 2014; Linares & Díaz, 2010; Luong, Phung, Sly, Morawska, & Thai, 2017; Rodopoulou et al., 2014; Tam et al., 2014) e incrementos en las tasas de suicidios (Kim et al., 2015). Por otro lado, se han desarrollado estudios de cohorte para evaluar relaciones de causalidad, para individuos seleccionados, entre los contaminantes atmosféricos y determinadas enfermedades, que han indicado que la exposición a partículas finas puede causar daño en el ADN (Huang et al., 2012) y defectos cardíacos congénitos (B. Zhang et al., 2016). Sin embargo, los estudios mencionados no van más allá de las medidas agregadas de exposición, es decir que no permiten obtener información sobre la exposición a una escala individual de los habitantes en una zona de estudio definida, necesaria para mejorar el alcance en estudios epidemiológicos (Briggs, 2005; Setton, Hystad, & Keller, 2005).

Para manejar las limitaciones relacionadas a la estimación espacial de la exposición ambiental a contaminantes atmosféricos, se han utilizado diferentes métodos, como: algoritmos de interpolación de datos de estaciones de monitoreo en sitios fijos (Jerrett et al., 2005; J. Li & Heap, 2011), métodos geoestadísticos (Dabass et al., 2016; Jerrett et al., 2005), modelos de redes neuronales (Blanes-Vidal, Cantuaria, & Nadimi, 2017), modelos de transporte (Shekarzifard et al., 2015), modelos de dispersión (de Hoogh et al., 2014; Marshall, Nethery, & Brauer, 2008) y regresiones espaciales (Briggs et al., 1997; Eeftens et al., 2012). En la revisión realizada por Li & Heap (2011), se encontró que el algoritmo de interpolación espacial Inverso de la Distancia (IDW), el Kriging ordinario (OK) y el Cokriging ordinario (OCK), eran los métodos más utilizados en las ciencias ambientales, en donde, la variación en los datos es un factor de impacto dominante y tiene efectos significativos en el desempeño de los métodos. Dabass et al (2016), utilizó el algoritmo geoestadístico Kriging para estimar la exposición diaria a material particulado menor a $2.5 \mu\text{m}$ ($\text{PM}_{2.5}$) y ozono (O_3) en algunas residencias, permitiendo estudiar los efectos agudos de los contaminantes sobre la mortalidad por enfermedades cardiovasculares en una gran área metropolitana.

En el estudio realizado por Shekarzifard et al. (2015), se implementó un modelo de transporte con una capacidad detallada para modelar las emisiones del tráfico, el cual simula los flujos y emisiones de tráfico en una región metropolitana, y estima las emisiones de óxidos de nitrógeno (NO_x) en el tráfico por carretera. Se encontró que, al compararlo con un método de regresión espacial implementado previamente, la correlación entre los resultados fue alta para las áreas afectadas por un alto nivel de contaminación relacionada con el tráfico (Shekarzifard et al., 2015).

En el estudio desarrollado por Blanes-Vidal et al. (2017), se comparó un enfoque para la evaluación de la exposición utilizando un sistema de inferencia difuso neuro-adaptativo (NFIS), donde las entradas del modelo eran medidas de proximidad a fuentes y la salida era la exposición residencial a un contaminante específico, con un modelo de dispersión y un modelo de proximidad lineal, para amoníaco (NH_3). Los resultados del modelo dispersión presentaron un comportamiento similar al NFIS, a diferencia del modelo de proximidad, indicando que la implementación de esta metodología podría ser una herramienta útil para estimar exposiciones individuales cuando no se disponga de datos de dispersión de emisiones (Blanes-Vidal et al., 2017).

Los modelos de dispersión se han utilizado en el estudio de la variación espacial de los contaminantes atmosféricos. Estos representan una alternativa robusta de estimación, sin embargo, la información requerida para su implementación es abundante y costosa (Beelen, Voogt, Duyzer, Zandveld, & Hoek, 2010; de Hoogh et al., 2014; Marshall et al., 2008). Diferentes estudios han comparado la implementación de modelos de dispersión y modelos de regresión espacial, específicamente, modelos de regresión de usos del suelo (LUR). En el estudio realizado por Marshall et al. (2008), se compararon tres diferentes enfoques para estimar la variabilidad espaciotemporal intraurbana en las concentraciones de cuatro contaminantes en Vancouver (óxido de nitrógeno (NO), dióxido de nitrógeno (NO_2), monóxido de carbono (CO) y ozono (O_3), por medio de un modelo empírico-estadístico basado en análisis geográficos (LUR); y un modelo de dispersión Euleriano (*Community Modeling and Analysis System, CMAQ*). Se concluyó que ambos modelos tienen un buen desempeño en la predicción de las concentraciones en los sitios de monitoreo, generando el modelo LUR una mayor resolución temporal y espacial en los resultados (Marshall et al., 2008).

Por otro lado, Hoogh et al. (2014) y Beelen, Voogt, Duyzer, Zandveld, & Hoek (2010) observaron que, tanto los modelos de dispersión, como los de regresión de usos del suelo, pueden ser útiles para estimar las variaciones a pequeña escala de la contaminación atmosférica relacionada con la combustión al aire libre, típicamente relacionada con el tráfico de la carretera; sin embargo, es importante mencionar que el modelo LUR no puede manejar fácilmente diferentes periodos de tiempo (horas, días, semanas, años) ni datos de concentraciones o meteorología en tiempo real, en comparación con los modelos de dispersión (de Hoogh et al., 2014).

A pesar de que existen diferentes metodologías para estimar los niveles de los contaminantes atmosféricos en una zona determinada, en estudios epidemiológicos es común el desarrollo de modelos de regresión de usos del suelo para estimar la exposición a los contaminantes atmosféricos relacionados con el tráfico vehicular (Cattani et al., 2017; Chen, Shi, Li, Bai, & Wang, 2017; Khreis & Nieuwenhuijsen, 2017; Mølter, Lindley, de Vocht, Simpson, & Agius, 2010a; Shekarrizfard et al., 2015). Este método busca predecir la concentración de un contaminante atmosférico en un lugar determinado, en función de las características del terreno circundante a las estaciones de monitoreo del contaminante o *buffers*, como: los usos del suelo, la intensidad del tráfico, la proximidad a fuentes de emisión y la meteorología, basándose en una regresión espacial

multivariada (Briggs et al., 1997; Buteau et al., 2017; Cordioli et al., 2017; Eeftens et al., 2012; Jerrett et al., 2005).

La metodología fue propuesta por Briggs et al (1997), como respuesta a la necesidad de mapear la contaminación del aire relacionada con el tráfico, con el fin de identificar puntos de altas de concentraciones en ambientes urbanos, definir grupos poblaciones en riesgo, mostrar cambios en patrones espaciales de los contaminantes y mejorar las estimaciones de la exposición en estudios epidemiológicos. En este estudio, se construyeron ecuaciones de regresión para el NO₂ medido en 80 puntos de monitoreo para cuatro ciudades europeas: Huddersfield (Reino Unido), Ámsterdam (Países Bajos), Praga (República Checa) y Poznan (Polonia); siguiendo la siguiente metodología (Briggs et al., 1997):

1. Selección de sitios de muestreo
2. Definición de predictores SIG
3. Desarrollo de modelos de regresión multivariados
4. Evaluación de la calidad y el rendimiento del modelo

Donde para cada estación de monitoreo seleccionada, se construyeron áreas circulares con centro en las estaciones y de radios variantes entre 0 y 300 m (buffers circulares), los cuales fueron intersectados con las variables espaciales: malla vial, flujo vehicular y usos del suelo. Se encontró que los modelos desarrollados podían explicar entre el 61 y 92% de la variabilidad del NO₂, y que todos diferían entre sí, debido a que, dependiendo de las características de la zona de estudio, la ecuación para este contaminante estaba en función de diferentes variables, siendo el flujo vehicular y el uso residencial e industrial a 300 m, las variables más comunes (Briggs et al., 1997).

En diferentes estudios realizados a nivel mundial, se puede observar que el contaminante comúnmente evaluado con esta metodología es el NO₂, por estar directamente relacionado a emisiones vehiculares (Allen, Amram, Wheeler, & Brauer, 2011; Beelen et al., 2013; Bertazzon, Johnson, Eccles, & Kaplan, 2015; Cordioli et al., 2017; Cowie et al., 2019; Crouse, Goldberg, & Ross, 2009; Dijkema et al., 2011; M. Wang et al., 2012). Sin embargo, existen diferentes estudios que han probado el rendimiento de la implementación de modelos LUR para material particulado (Cattani et al., 2017; Chen et al., 2018, 2017; Clougherty et al., 2013; de Hoogh et al., 2018, 2016; Dirgawati et al., 2016), partículas ultrafinas (diámetro menor a 0.1 µm - UFP) (Cattani et al., 2017; Montagne et al., 2015; Van Nunen et al., 2017) y *black carbon* (BC) (Clougherty et al., 2013; de Hoogh et al., 2018; Dons et al., 2014; Montagne et al., 2015).

Las áreas de estudio en donde se ha aplicado la metodología LUR a nivel mundial, varían en tamaño, características topográficas y fuentes de emisión de los contaminantes. Se han desarrollado modelos en diferentes centros urbanos de Europa, con una varianza explicada entre el 55 y 92 % para NO₂, el 31 y 94% para PM_{2.5}, el 35 y 59% para BC y el 36 y 71% para UFP (Beelen et al., 2013; Cattani et al., 2017; de Hoogh et al., 2018; Eeftens et al., 2012; Montagne et al., 2015). En Asia, se desarrollaron modelos con una varianza explicada entre 84 y 91% para NO₂, y el 42 y 78% para PM_{2.5} (Chen et al., 2017; Shi, Ho, Xu, & Ng, 2018; Shi, Lau, & Ng, 2017). En Perth,

Australia, se desarrolló un modelo que explicaba hasta el 67% de la variabilidad del PM_{2.5} (Dirgawati et al., 2016). En Norteamérica se han desarrollado modelos que explicaban la variabilidad del NO₂ entre un 55 y 84% y del NO entre un 55 y 56% (Allen et al., 2011; Bertazzon et al., 2015; Crouse et al., 2009).

Para Latinoamérica, se encontró un modelo desarrollado en Ciudad de México que explicaba el 82 y 76% de la variabilidad para NO₂ y PM_{2.5}, respectivamente (Son et al., 2018). Hasta la fecha, en Colombia sólo se ha desarrollado un modelo de regresión de usos del suelo, correspondiente al estudio realizado por Londoño & Cañón (2015), en el cual se describe la metodología para desarrollar un modelo LUR para PM₁₀ con información del mes de enero de 2007, para la ciudad de Medellín, con una variabilidad explicada del 79%. Por otro lado, se ha reenfocado la metodología y se ha explorado el desarrollo de modelos LUR en regiones más grandes, como Irlanda con 78% de NO₂ explicado (Naughton et al., 2018), China con 82% para PM_{2.5} (Chen et al., 2018), y se ha desarrollado un modelo global para NO₂, con una variabilidad explicada del 52% (Larkin et al., 2017).

El rendimiento de un modelo LUR, está altamente relacionado con la cantidad de estaciones de monitoreo del contaminante de interés. No existe un umbral mínimo requerido de puntos de muestreo para aplicar esta metodología, sin embargo, se ha observado que modelos desarrollados con una pequeña cantidad de puntos tienden a seleccionar un conjunto diferente de variables cada vez, a ser muy sensibles a la cantidad de predictores potenciales ofrecidos y a ser propensos a un posible sobreajuste (Basagaña et al., 2012; Eeftens et al., 2012). La cantidad de estaciones difiere en gran medida para cada zona de estudio, desde 10 hasta 161 estaciones para modelos en centros urbanos (Cordioli et al., 2017; Eeftens et al., 2012; Londoño & Cañón, 2015); de 15 estaciones para modelar el NO₂ en Irlanda (Naughton et al., 2018), hasta 1452 estaciones para un modelo LUR de PM_{2.5} en China (Chen et al., 2018) y 5220 estaciones para un modelo LUR global de NO₂ (Larkin et al., 2017).

En los modelos LUR, las variables predictoras o explicativas más comunes son: tipos de uso del suelo urbano, variables de tráfico como longitud y tipo de vía y flujo o densidad vehicular, obtenidas a partir de la intersección con buffers entre 25 y 5000 m, con tamaños de paso de 50, 100 y 1000 m (Bertazzon et al., 2015; Cordioli et al., 2017; de Hoogh et al., 2016, 2014; Eeftens et al., 2012; Londoño & Cañón, 2015; Marshall et al., 2008; Shi et al., 2017). En zonas donde existen calentadores a base de carbón, es común que la densidad poblacional sea significativa, especialmente en los modelos estacionales de invierno. Por otro lado, algunos estudios han incluido la disponibilidad del viento representada por variables geomorfológicas (índice de área frontal, índice de barlovento/sotavento, rugosidad superficial) (Cordioli et al., 2017; Shi et al., 2017), velocidades del viento interpoladas (Arain et al., 2007; Bertazzon et al., 2015; Londoño & Cañón, 2015), estimaciones de las concentraciones de contaminantes a partir de modelos químicos de transporte para aumentar la densidad de monitoreo disponible en la zona (Marshall et al., 2008; Michanowicz, Shmool, Cambal, et al., 2016; Michanowicz, Shmool, Tunno, et al., 2016; Mølter et al., 2010a) y de imágenes satelitales (Cowie et al., 2019; de Hoogh et al., 2016; R. Li, Ma, Xu, &

Song, 2018; Shi et al., 2018), niveles de ruido, precipitación y gradientes de temperatura y velocidad del viento (Londoño & Cañón, 2015).

Desde la estructuración de la metodología LUR, propuesta por Briggs et al (1997), en la que se desarrolló una caracterización espacial multivariada con variables derivadas de una evaluación de proximidad desde puntos de monitoreo, a las posibles fuentes de emisión y de remoción de contaminantes atmosféricos, muchos estudios han aplicado esta metodología, hasta llegar al punto de ser un estándar en la epidemiología ambiental para estimar la exposición de la población a diferentes contaminantes relacionados con el tráfico vehicular, por su relativa facilidad de aplicación, costos y por el resultado asociado, el cual permite obtener una distribución espacial de los contaminantes en escalas más finas.

2.2. Descripción del problema

Actualmente, la contaminación atmosférica es un problema ambiental de especial interés por su consecuente afectación a la salud humana. Estudios previos han sugerido que las exposiciones a diferentes contaminantes, a corto y largo plazo, están asociados a incrementos en la mortalidad y a enfermedades respiratorias y cardiovasculares (Camilo Blanco-Becerra et al., 2014; Janssen, Fischer, Marra, Ameling, & Cassee, 2013; Yorifuji et al., 2013). De acuerdo con la Organización Mundial de la Salud (OMS), alrededor de 3 millones de muertes a nivel mundial son atribuibles únicamente a la contaminación del aire, afectando todas las regiones, entornos, grupos socioeconómicos y grupos de edad; implicando costos en salud, ocasionados por la degradación de la calidad del aire y el ambiente. Sin embargo, la magnitud de la exposición de las personas que viven en un área determinada, varía en función de las condiciones geográficas heterogéneas, presentes en la zona (World Health Organization, 2016).

En el Valle de Aburrá, se estiman los costos en salud, asociados a una calidad del aire deteriorada, en \$2,8 billones de pesos, equivalentes al 5% del Producto Interno Bruto (PIB) del área (Golub, Klychnikova, Sanchez-Martinez, Belausteguigoitia, & Molina, 2014); donde, en los últimos años, el contaminante atmosférico de interés ha sido el material particulado menor a 2.5 micrómetros ($PM_{2.5}$), el cual ha presentado niveles alarmantes con superaciones frecuentes de los umbrales establecidos en la norma colombiana de $50 \mu\text{g}/\text{m}^3$ (Área Metropolitana del Valle de Aburrá, 2016c, 2017b; Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010b); algo que se ha atribuido, además de a las fuentes de emisión, a la orografía y a las condiciones meteorológicas del Valle de Aburrá, por su influencia en procesos de estabilidad atmosférica, propiciando un bajo transporte de las partículas (Área Metropolitana del Valle de Aburrá, 2013). Los inventarios de emisiones realizados en el año 2013 y 2015 en el Valle de Aburrá han sugerido que cerca del 80% de $PM_{2.5}$ emitido directamente a la atmósfera proviene de las fuentes móviles, es decir, del parque automotor; mientras que el porcentaje restante es atribuido al sector industrial (Área Metropolitana del Valle de Aburrá, 2016b, 2017a). Por otro lado, las emisiones de gases precursores de $PM_{2.5}$ como los óxidos de azufre, de nitrógeno y los compuestos orgánicos volátiles,

indican que, para el primer contaminante, cerca del 90% es emitido por el sector industrial, mientras que para los otros, el 83% y el 92%, respectivamente, se atribuyen a las fuentes móviles (Área Metropolitana del Valle de Aburrá, 2016b).

Con la implementación de la Red de Monitoreo de Calidad del Aire del Valle de Aburrá, administrada por el Sistema de Alerta Temprana del Valle de Aburrá (SIATA), se han evidenciado incrementos estacionales en las series de material particulado, donde, para los años 2015 y 2016, se han identificado episodios de contaminación por $PM_{2.5}$ durante el mes de marzo, en dos municipios ubicados en el Valle de Aburrá (Área Metropolitana del Valle de Aburrá, 2018). Como respuesta al problema que propone estos incrementos en los niveles de $PM_{2.5}$, se construyó el Plan Operacional para Enfrentar Episodios Críticos de Contaminación Atmosférica – POECA, en donde se establece el $PM_{2.5}$ como contaminante crítico, debido al deterioro de la calidad del aire que este ocasiona, generando niveles “Dañinos para grupos sensibles” y “Dañinos para la salud” (Área Metropolitana del Valle de Aburrá, 2016a). Como consecuencia, se ha ahondado en la medición y el análisis del $PM_{2.5}$ en el Valle de Aburrá, implicando un aumento en la cantidad de estaciones de monitoreo de este contaminante. Por estas razones, en conjunto a la existencia de información suficiente para la construcción de modelos LUR, el desarrollo del presente estudio se ha enfocado al comportamiento del $PM_{2.5}$.

En Medellín y los otros 9 municipios que conforman el Valle de Aburrá, durante el año 2018 han funcionado 24 estaciones de $PM_{2.5}$ en la red de monitoreo. El diseño de dicha red no permite conocer la concentración del contaminante en lugares donde no se encuentre una estación ubicada, ni saber el comportamiento del mismo a medida que se aleja de la estación. En otras palabras, existe una distribución espacio-temporal limitada en el sistema de vigilancia de calidad del aire, generando la necesidad de densificar la red de monitoreo o de emplear estrategias para estimar el valor del contaminante en otras zonas, con el fin de disminuir dicha limitante y de evaluar la variación espacial de $PM_{2.5}$ en una escala intraurbana, tanto por sus impactos ambientales, como por sus efectos a la salud.

Según lo expuesto anteriormente, los altos niveles de $PM_{2.5}$ en el Valle de Aburrá y, por ende, en la ciudad de Medellín, se ven influenciados por las variables meteorológicas, sin embargo, estos no son los únicos factores que condicionan el comportamiento del contaminante en la ciudad. De acuerdo a diferentes estudios, se podría asociar la variabilidad espacial de los niveles de este contaminante en una zona urbana con variables no meteorológicas como la distancia a uso industrial y la distancia a zonas verdes desde la estación de monitoreo (Eeftens et al., 2012; Wu et al., 2017), la densidad vehicular (Eeftens et al., 2012), además de las variables meteorológicas como la precipitación, el gradiente de temperatura y velocidad del viento (Londoño, 2018a; Londoño & Cañón, 2015). Lo anterior, conlleva a establecer la siguiente pregunta de investigación: *¿existen variables no meteorológicas que incidan en la variabilidad espacial de las concentraciones de este contaminante en una escala intraurbana?*

3. Objetivos

3.1. Objetivo general

Proponer modelos de regresión de usos del suelo mensuales para caracterizar espacialmente la contaminación del aire por $PM_{2.5}$ en el municipio de Medellín en el año 2018, que permitan estimar el alcance geográfico de posibles fuentes cercanas y sumideros, con el fin de mejorar las redes de vigilancia de la calidad del aire existentes, y el diseño de futuras redes de monitoreo.

3.2. Objetivos específicos

1. Explorar las series del contaminante atmosférico $PM_{2.5}$ registrado en cada estación de monitoreo de Medellín, y las posibles variables explicativas, para el año 2018.
2. Construir modelos de regresión con datos espaciales para la concentración mensual de $PM_{2.5}$, en función de distintos buffers y variables explicativas.
3. Seleccionar el mejor modelo de regresión de usos del suelo mensual para $PM_{2.5}$, aplicando los criterios de especificación del método de mínimos cuadrados ordinarios (OLS).

4. Marco de referencia

4.1. Marco teórico

4.1.1. Contaminantes atmosféricos

La contaminación atmosférica es un problema local, regional y global, definido como la presencia de sustancias en el aire que interfieren el bienestar y la salud humana, y pueden producir otros efectos dañinos para el medio ambiente. En general, el análisis de los contaminantes se ha concentrado en seis contaminantes criterio, debido a sus posibles afectaciones a la salud, al medio ambiente y a la infraestructura; dichos contaminantes, establecidos en las normas de calidad del aire son: material particulado (PM_{10} y $PM_{2.5}$), ozono (O_3), monóxido de carbono (CO), dióxido de azufre (SO_2) y dióxido de nitrógeno (NO_2) (Ministerio de Ambiente y Desarrollo Sostenible. República de Colombia, 2017; Vallero, 2008). Estos contaminantes tienen diversas fuentes de emisión, propiedades fisicoquímicas e impactos; entre los cuales, se destacan los efectos sobre la transformación, la depositación y el transporte de especies químicas, el forzamiento radiativo y la salud humana (Squizzato et al., 2017).

Cuando estos contaminantes son emitidos directamente desde la fuente de emisión, se consideran contaminantes primarios, mientras que, cuando son producto de reacciones químicas en la atmósfera, son secundarios; es decir que los contaminantes pueden ser emitidos directamente como resultado de actividades antropogénicas o por causas naturales, o pueden ser un subproducto de las reacciones de otros contaminantes en el ambiente (Donahue, 2017; Squizzato et al., 2017). Por otro lado, los contaminantes también se clasifican por su composición química, en donde la mayoría de los contaminantes son moléculas individuales o grupos estrechamente unidos de moléculas; sin embargo, cuando se habla del material particulado (PM), las partículas son clasificadas según sus propiedades físicas. A diferencia de los otros contaminantes criterio, el PM no es una entidad química específica, por el contrario, es una mezcla de partículas de diferentes tamaños, composiciones y propiedades; características que dependen de la fuente de emisión de la sustancia y que residen en una fase condensada (no acuosa) (Donahue, 2017; Vallero, 2008).

La composición química de las partículas en el aire incluye iones inorgánicos, compuestos metálicos, carbono elemental, compuestos orgánicos y sustancias de la corteza terrestre. Una clasificación importante para el material particulado es su tamaño (diámetro esférico equivalente), donde el tamaño de las partículas está determinado por la fuente de emisión, debido a que procesos mecánicos pueden generar partículas gruesas, mientras que procesos de combustión son relacionados a partículas más pequeñas. Entre las partículas más grandes se encuentra el PM_{10} , que corresponde a material particulado con diámetro aerodinámico menor a 10 micrómetros (μm); por otro lado, el $PM_{2.5}$ corresponde a partículas con diámetro aerodinámico menor a 2.5 μm , llamadas partículas finas o fracción respirable, son de especial interés debido a que, por su tamaño, se transportan por largas distancias y pueden llegar profundamente en el sistema respiratorio (Donahue, 2017; Vallero, 2008).

Entre las fuentes antropogénicas que generan mayor impacto en la calidad del aire, se encuentran las emisiones originadas por el tráfico vehicular y las industrias (Colville, Hutchinson, Mindell, & Warren, 2001; Squizzato et al., 2017). En las zonas urbanas, la combustión y las emisiones relacionadas con el tráfico, a menudo representan las principales fuentes primarias de material particulado; sin embargo, según las características del área de estudio, otras fuentes pueden afectar la calidad del aire, como las actividades industriales y la resuspensión del polvo en la carretera. Por otro lado, el material particulado puede presentarse como subproducto de procesos químicos que involucran compuestos orgánicos o inorgánicos emitidos por diferentes fuentes (Squizzato et al., 2017).

Las partículas finas ($PM_{2.5}$), que generalmente provienen de procesos de combustión industrial o vehicular, pueden penetrar fácilmente en el tracto respiratorio y llegar al sistema circulatorio, y ser más tóxicas que las partículas gruesas. Cambios en los niveles del contaminante se han asociado significativamente con un incremento en la mortalidad y la morbilidad por enfermedades respiratorias y cardiovasculares (Rodopoulou et al., 2014). Además, existe una amplia evidencia epidemiológica sobre posibles asociaciones entre la exposición a corto y largo plazo a $PM_{2.5}$ y el riesgo de resultados adversos en el embarazo, insuficiencias cardíacas, EPOC (Enfermedad Pulmonar Obstructiva Crónica) y asma (Huang et al., 2012; Luong et al., 2017; Rodopoulou et al., 2014; Sava & Carlsten, 2012; B. Zhang et al., 2016).

Para estimar las concentraciones del material particulado, generalmente se mide la masa que se encuentra en las dos categorías de tamaño mencionadas previamente (diámetro equivalente de 2.5 y 10 μm). Estas mediciones se toman por instrumentos que utilizan mecanismos de exclusión de tamaño para segregar la masa de cada fracción, en donde dicha masa puede determinarse para una partícula predominantemente esférica, por: microscopía (óptica o electrónica), dispersión de la luz, movilidad eléctrica de la partícula o su comportamiento aerodinámico (Vallero, 2008). Las técnicas de medición de contaminantes atmosféricos pueden ser divididas en cuatro grupos: sensores remotos, muestreadores pasivos, muestreadores activos (manuales o semiautomáticos) y analizadores automáticos, siendo los dos últimos utilizados específicamente para material particulado en Colombia (Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010a).

Los muestreadores activos manuales recolectan las muestras de contaminantes por métodos físicos, debido a que toman un volumen conocido de aire y lo bombean a través de un filtro, por un periodo de tiempo determinado; después, el colector es retirado y llevado al laboratorio para su análisis. Estos equipos pueden ser: *Hi Vol (High Volumen Sampler)* o *Low Vol (Low Volumen Sampler)*. Por otro lado, los muestreadores activos semiautomáticos son equipos activos con un sistema electrónico incorporado con el fin de mejorar la calidad de los datos y de reducir la incertidumbre, sin embargo, también requiere de un análisis posterior la muestra. Una de las principales ventajas de estos equipos radica en que permite muestrear varios días cambiando automáticamente los filtros y, al intercambiar el cabezal, se pueden medir diferentes fracciones de material particulado

(PST, PM₁₀, PM_{2.5} y PM₁) (Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010a).

Finalmente, los analizadores automáticos permiten realizar mediciones puntuales con alta resolución temporal (promedios horarios o cada 10 minutos). Las muestras son analizadas por medio de métodos electro-ópticos (absorción UV, infrarrojo no dispersivo, fluorescencia o quimioluminiscencia) y los datos pueden ser transmitidos en tiempo real. Se recomiendan en gran medida si se desean constituir mediciones continuas de un contaminante, debido a su amplio cubrimiento temporal y poco esfuerzo operacional. Para el material particulado se recomiendan los equipos tipo *Beta Gauge* (atenuación de radiación beta) o *Tapered Element Oscillating Microbalance* (micro-balanza oscilante de elemento cónico) (Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010a).

– Variables meteorológicas y material particulado

Los contaminantes del aire alcanzan los receptores al ser transportados y, en algunos casos, transformados en la atmósfera, siendo la dispersión de estos condicionada por procesos de advección y difusión. El proceso de transporte de contaminantes más directo es la advección, donde la especie química se mueve junto con el fluido transportador, es decir que la velocidad y dirección de transporte están determinadas por la velocidad y la dirección de flujo del fluido. Sin embargo, desde un punto de vista físico, varios factores son determinantes en la dispersión de contaminantes atmosféricos: la estabilidad atmosférica; el viento y la turbulencia inducidos por fenómenos meteorológicos o por la rugosidad de la superficie de la tierra y la combinación de los factores térmicos y mecánicos, la orografía y los cambios de temperatura (Vallero, 2019).

Las concentraciones de material particulado varían fuertemente con las condiciones meteorológicas, esto puede favorecer la ventilación de las capas inferiores de la tropósfera y, por lo tanto, la dispersión de los contaminantes, o llevar a una acumulación cerca de las fuentes de emisión (Kim Oanh, Chutimon, Ekbordin, & Supat, 2005; Llargeron & Staquet, 2016; Vallero, 2019). Por ejemplo, en condiciones de estabilidad atmosférica, las concentraciones de material particulado pueden acumularse, generando una posible disminución la radiación solar entrante y la visibilidad, ocasionando a su vez un incremento de las concentraciones de CO₂ y otros gases. Esta situación aumenta las temperaturas de la superficie y la magnitud del efecto invernadero local, induciendo episodios de altas concentraciones de otros contaminantes en fase gaseosa y del mismo material particulado (Vallero, 2019).

Las condiciones meteorológicas influyen el transporte de largo alcance y la formación de contaminantes secundarios, modificando la concentración y composición del material particulado; esto implica que los cambios en las condiciones ambientales pueden reflejarse en los patrones de distribución de estos aerosoles. La precipitación actúa como un mecanismo de remoción húmeda de material particulado en la atmósfera, teniendo una relación negativa con el contaminante, al igual que la humedad relativa (Llargeron & Staquet, 2016; Pateraki, Asimakopoulos, Flocas, Maggos, & Vasilakos, 2012).

La temperatura del aire favorece la presencia de partículas de diferentes diámetros en la atmósfera debido a las reacciones fotoquímicas que se presentan y que permiten la formación de partículas secundarias. Además, cambios en la temperatura están asociados a la radiación solar incidente, incrementando la energía cinética turbulenta, afectando la altura de la capa de mezcla, asociada a la dispersión de contaminantes en la tropósfera. Debido a esto, la temperatura puede considerarse como uno de los predictores más fuertes de los niveles de contaminación (Kim Oanh et al., 2005; Papanastasiou, Melas, & Kioutsioukis, 2007; Pateraki et al., 2012).

Por otro lado, la velocidad del viento influencia de manera significativa las concentraciones de contaminantes en el aire, debido a que es determinante en el transporte horizontal y la dispersión de los mismos. Elevados niveles de contaminación se encuentran asociados a velocidades del viento bajas o calmas, mientras que con un aumento en la magnitud de esta variable, se esperan condiciones propicias para la dispersión de contaminantes en el aire (Papanastasiou et al., 2007).

– Sistemas de vigilancia de calidad del aire (SVCA)

La definición de una red de monitoreo de calidad del aire depende del conocimiento previo de las variables que afectan la dispersión de contaminantes atmosféricos, la norma vigente, la distancia a fuentes de emisión, los objetivos de la red, el área de influencia y el comportamiento espacial de los contaminantes (Londoño, 2018a). A partir de los lineamientos contemplados en el Protocolo para el Monitoreo y Seguimiento de la Calidad del Aire, en Colombia se han implementado diferentes estaciones de monitoreo que hacen parte de Sistemas de Vigilancia de Calidad del Aire (SVCA), con el fin de controlar y mitigar los efectos producidos por la contaminación atmosférica sobre la salud humana y el ambiente (IDEAM, 2018). Los SVCA tienen diferentes objetivos, dependiendo de las necesidades de la zona donde se planea implementar, ya sea desde determinar el cumplimiento de las normas de calidad del aire, observar tendencias a mediano y largo plazo, estudiar fuentes de emisión, validar modelos de dispersión del aire o diseñar procesos de control en episodios críticos de contaminación (Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010a).

Para dar cumplimiento a los objetivos establecidos en un SVCA, se implementan redes de monitoreo de calidad del aire en una región determinada, con el fin de muestrear, analizar y procesar las concentraciones de uno o más contaminantes atmosféricos (Londoño, 2018a). Dependiendo de la población de la zona a evaluar o del posible deterioro de la calidad del aire asociado a actividades económicas específicas, como la minería, se establecen diferentes tipos de SVCA (Tabla 1) (Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010a).

Tabla 1. Tipos de SVCA, según número de habitantes

Tipo de SVCA	Condiciones
Tipo I - Indicativo	Población entre 50,000 y 150,000 habitantes
Tipo II - Básico	Población entre 150,000 y 500,000 habitantes
Tipo III - Intermedio	Población entre 500,000 y 1,500,000 habitantes

Tipo IV - Avanzado	Población mayor a 1,500,000 habitantes
SEVCA - Sistema Especial de Vigilancia de Calidad del Aire	Población con problemáticas específicas en calidad del aire
SVCAI - Sistema de Vigilancia de Calidad del Aire Industrial	La autoridad ambiental establece las actividades con obligación de implementar un SVCA

Fuente: Ministerio de Ambiente, Vivienda y Desarrollo Territorial. República de Colombia. (2010a). Manual de diseño de sistemas de vigilancia de la calidad del aire. Protocolo para el Monitoreo y Seguimiento de la Calidad del Aire. Bogotá D.C. Recuperado de http://www.minambiente.gov.co/images/AsuntosambientalesySectorialUrbana/pdf/contaminacion_atmosferica/Protocolo_Calidad_del_Aire_-_Manual_Diseño.pdf.

Para evaluar la calidad del aire, las estaciones de monitoreo son clasificadas en diferentes niveles, con el fin de conocer explícitamente las características y representatividad de los datos registrados por cada estación. Las estaciones se clasifican según el tipo de área (urbana, suburbana o rural), el tiempo de muestreo (fija o indicativa) y las emisiones dominantes (de tráfico, punto crítico, industrial y de fondo), como se muestra en la Tabla 2. Según la clasificación de la estación, se han establecido los contaminantes que comúnmente se producen en los procesos relacionados a la fuente de emisión y ubicación del monitor, donde el PM_{2.5} es un compuesto de interés en todas las estaciones ubicadas en áreas urbanas y suburbanas, sin importar el tipo de fuente predominante (Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010a).

Tabla 2. Clasificación de estaciones de monitoreo de calidad del aire

Clasificación según	Tipo	Descripción
Tiempo de muestreo	Fija	Permanece activa en un punto por un periodo de tiempo superior a un año
	Indicativa	Permanece en un punto por periodos inferiores a un año
Tipo de área	Urbana	Área totalmente urbanizada
	Suburbana	Área mayormente urbanizada, con mezclas de áreas no urbanizadas (áreas agrícolas, lagos, bosques, grandes zonas verdes, entre otras)
	Rural	Todo lo que no satisfaga requisitos de área urbana o suburbana
Emisiones dominantes	De tráfico	Nivel de contaminación determinado principalmente por emisiones de fuentes móviles
	Punto crítico	Ubicadas a nivel de piso y usadas como apoyo a estudios epidemiológicos
	Industrial	Nivel de contaminación determinado principalmente por emisiones de fuentes fijas
	De fondo	Nivel de contaminación no está influenciado por fuentes directas, sino por la dinámica atmosférica

Fuente: Ministerio de Ambiente, Vivienda y Desarrollo Territorial. República de Colombia. (2010a). Manual de diseño de sistemas de vigilancia de la calidad del aire. Protocolo para el Monitoreo y Seguimiento de la Calidad del Aire. Bogotá D.C. Recuperado de http://www.minambiente.gov.co/images/AsuntosambientalesySectorialUrbana/pdf/contaminacion_atmosferica/Protocolo_Calidad_del_Aire_-_Manual_Diseño.pdf.

Por otro lado, los SVCA también incluyen una componente encargada de describir la meteorología de la zona de estudio. Se utilizan estaciones meteorológicas que midan velocidad (m/s) y dirección del viento (°), temperatura (°C), radiación solar (W/m²), humedad relativa (%), pluviometría (mm) y presión barométrica (mm Hg). La ubicación de estas estaciones están sujetas a distintos factores, como: la rugosidad de la superficie terrestre, los usos del suelo, la infraestructura urbana, las condiciones micro-climáticas de la capa límite planetaria y la capa de mezcla (Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia, 2010a).

Durante el año 2017, en Colombia operaron 26 SVCA, con 204 estaciones de monitoreo, donde 166 fueron estaciones fijas y 38 indicativas. Los sistemas de vigilancia estuvieron distribuidos en las regiones Andina, Caribe, Pacífico y Orinoquía, en un total de 22 departamentos y 91 municipios; donde 89 de las 204 estaciones evaluaron PM_{2.5}. La autoridad ambiental con mayor número de estaciones en funcionamiento fue el Área Metropolitana del Valle de Aburrá (AMVA), con 40 estaciones. Por otro lado, en el mismo año se encontraron 73 estaciones que registraron información las variables meteorológicas mencionadas, distribuidas en 14 sistemas de vigilancia, entre ellos el AMVA (IDEAM, 2018).

4.1.2. Usos del suelo urbano

En el Plan de Ordenamiento Territorial (POT) vigente del municipio de Medellín (Acuerdo 48 de 2014), se definen los usos del suelo urbano como las normas generales para la distribución espacial de las actividades urbanas en el municipio, las cuales establecen las características, utilización, grado de intensidad y ocupación del suelo, con el fin de optimizar la infraestructura urbana e impulsar relaciones socio-económicas más eficientes, mitigando los impactos generados en su entorno y atendiendo a las características del territorio (Alcaldía de Medellín, 2014).

En el municipio de Medellín, la clasificación de Los Usos Generales del Suelo Urbano está en función de la mezcla de actividades urbanas presentes en la zona, definiendo límites tanto para el área residencial, como para las otras actividades. Los usos se clasifican en baja, media y alta mixtura, donde las categorías generales de uso aplicables son: residencial, comercial, servicios, industrial, dotacional, espacios públicos existentes y espacios públicos proyectados (Alcaldía de Medellín, 2014).

La asignación de usos del suelo promueve el cumplimiento de los objetivos del sistema de ordenamiento territorial, asociados a una ciudad compacta con crecimiento hacia adentro, la responsable mezcla de usos, el control de los impactos que algunas actividades pueden generar en el entorno y las políticas de movilidad sostenible. Parte importante en el POT se centra en el desarrollo de este último ítem, con el cual se busca contribuir a la reducción de la contaminación atmosférica y sus efectos sobre la salud de la población, el consumo excesivo de energía y la saturación de las vías de circulación (Alcaldía de Medellín, 2014). A continuación, se hace una descripción de la Red de Movilidad Motorizada, eje central de las políticas de movilidad sostenible.

– Red de movilidad motorizada

La promoción y aplicación de prácticas de movilidad responsable y el desarrollo de tecnologías que amplíen las opciones de movilidad sostenible son cruciales; por ende, existe un enfoque en privilegiar los modos no motorizados, la proximidad de servicios en nuevos procesos de urbanización y la regulación responsable del uso de vehículos en la ciudad (Alcaldía de Medellín, 2014). Sin embargo, como primer paso hacia la movilidad sostenible, es importante definir y clasificar las características de las redes viales existentes y proyectadas, para dirigir de manera eficiente las políticas que se deseen implementar.

Una red de movilidad motorizada se encuentra destinada para la movilización de vehículos particulares, transporte público y de carga. Las vías, dependiendo de su infraestructura, se clasifican en vías férreas, nacionales y urbanas, en donde, cada una se divide en vías con distinta jerarquía vial (Tabla 3). Específicamente para las vías urbanas, las cuales se encuentran adscritas al perímetro urbano de las municipalidades, se clasifican en autopistas urbanas o vías de travesía, vías arterias, colectoras, de servicio, privadas y motovías (Alcaldía de Medellín, 2014).

Tabla 3. Jerarquía vial de vías urbanas en el municipio de Medellín

Jerarquía vial	Descripción
Autopistas urbanas o vías de travesía	Permiten la circulación de flujos regionales de larga distancia y metropolitanos de larga o mediana distancia
Vías arterias	Movilizan flujo vehicular de largo recorrido dentro del área urbana, uniendo entre sí diferentes zonas de la ciudad
Vías colectoras (vías secundarias)	Distribuyen y canalizan el tránsito vehicular hacia o desde el sistema arterial hasta diferentes sectores urbanos
Vías de servicio (vías ordinarias)	Facilitan el acceso directo a las propiedades o actividades adyacentes, disminuyendo la importancia de la función de movilidad

Fuente: Alcaldía de Medellín. (2014). Acuerdo 48 de 2014 - Plan de Ordenamiento Territorial del Municipio de Medellín. Medellín. Recuperado de https://www.medellin.gov.co/irj/go/km/docs/pccdesign/SubportaldelCiudadano_2/PlandeDesarrollo_0_17/ProgramasyProyectos/Shared Content/Documentos/2014/POT/ACUERDO POT-19-12-2014.pdf

Según el POT de Medellín, en el numeral 2 - Artículo 186, se especifica que la sección vial, de las vías arterias en la ciudad, debe tener un ancho de veinticinco metros (25 m) desde eje de separador a paramento y para vías colectoras aproximadamente 15 m. Con el fin de mitigar los efectos de las emisiones de contaminantes atmosféricos y el ruido provenientes del parque automotor, se presentan áreas verdes pertenecientes al Subsistema de Movilidad dentro de la sección vial, en las glorietas y separadores (Alcaldía de Medellín, 2014).

4.1.3. Caracterización espacial

En el análisis de variables ambientales, es común que se necesiten datos espacialmente continuos, requerimiento que generalmente se ve limitado por la manera en que los datos son recolectados,

mediante muestreos puntuales. Como consecuencia de esto, existen métodos que permiten estimar una superficie espacial continua a partir de valores puntuales de una variable en específico. Partiendo del principio de autocorrelación espacial, se puede estimar el valor de un dato espacial $Z_i(x_i, y_i)$, definido en un dominio espacial $S(x, y)$ y en un sistema de proyección geográfico, conociendo su valor en distintos lugares en el espacio. Existe una mayor similitud entre datos espaciales más cercanos, en comparación a datos que tienen una mayor separación entre sí, indicando que el dato se correlaciona consigo mismo en función de la distancia. Como consecuencia de este principio, se puede estimar el valor de $Z_0(x_0, y_0)$ en un punto dado, donde se desconoce su valor, por medio de una combinación lineal de $Z_i(x_i, y_i)$ (Londoño, 2018a). A continuación se describen algunos métodos para la estimación de Z_0 , aplicados en el desarrollo de este trabajo.

– Algoritmos de interpolación espacial

Los algoritmos de interpolación espacial son enfoques matemáticos que permiten estimar los valores de un dato espacial; éstos se clasifican en función de diferentes características, descritas en la Tabla 4, como: la cantidad de datos utilizados en la estimación, el valor del error y el método de cálculo de los pesos ponderantes (λ_i) expuestos en la Ecuación 1, la cual puede interpretarse como un promedio ponderado de los datos (J. Li & Heap, 2014; Londoño, 2018a; Webster & Oliver, 2007).

Tabla 4. Características de los métodos de interpolación espacial

Clasificación según	Característica	Definición
Datos usados en la estimación	Global	Utilizan todos los datos disponibles en la región de interés para derivar la estimación y capturar las tendencias generales.
	Local	Operan dentro de un área pequeña alrededor del punto que se está estimando y capturan las variaciones locales o de corto alcance.
Diferencia entre dato observado y estimado	Exacto	Genera una estimación igual al valor observado en un punto muestreado.
	Inexacto	El valor estimado en un punto difiere de su valor conocido.
Incorporación de errores asociados a las estimaciones	Determinístico	No incorporan errores asociados a las estimaciones.
	Estocástico	Incorporan el concepto de aleatoriedad y proporcionan tanto estimaciones como errores asociados (parte estocástica, es decir, incertidumbres representadas como variaciones estimadas).
Superficie resultante	Gradual	Producen una superficie lisa y gradual.
	Abrupto	Producen una superficie discreta y abrupta.
Rango de los resultados	Convexos	Arrojan estimaciones que siempre se valoran entre el mínimo y el máximo de los valores observados.
	No convexos	Pueden arrojar estimaciones fuera del rango de los valores observados.

Clasificación según	Característica	Definición
Cantidad de variables explicativas	Univariado	Utilizan sólo muestras de la variable primaria para derivar la estimación.
	Multivariado	Usan variables explicativas o secundarias para derivar la estimación.
Función del estimador	Lineal	Derivan la estimación utilizando valores observados asumiendo una distribución normal de muestras y el estimador es una combinación lineal de los datos disponibles en la muestra.
	No lineal	Derivan predicciones basadas en los valores transformados de los datos observados, es decir que el estimador no es una combinación lineal de los datos.
Pesos ponderantes	Tradicional	Los pesos ponderantes se asignan a partir de consideraciones geométricas simples.
	Geoestadístico	Los pesos ponderantes se basan en la correlación espacial entre los datos.

Fuente: Li, J., & Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling and Software*, 53, 173–189. <https://doi.org/10.1016/j.envsoft.2013.12.008>.

$$Z_0(x_0, y_0) = \sum_{i=1}^n \lambda_i Z_i(x_i, y_i) \quad (1)$$

Inverso de la Distancia (IDW)

El algoritmo de interpolación espacial Inverso de la Distancia o *Inverse Distance Weighting* (IDW), en la estimación, incluye varios datos cercanos a la ubicación del dato que se desea estimar, es decir que es un método determinístico de estimación local, y asigna, a cada uno de ellos, un peso equivalente al inverso de la distancia al punto que se va a estimar (Londoño Ciro & Valdés Quintero, 2012). El valor a estimar (Z_0), se obtiene mediante la Ecuación 2, donde Z_i es el valor conocido, d_i es la distancia entre la ubicación del punto a estimar (x_0) y el valor disponible (x_i) y p es un exponente, usualmente igual a 2 (J. Li & Heap, 2014; Londoño Ciro & Valdés Quintero, 2012).

$$Z_0(x_0, y_0) = \frac{\sum_{i=1}^n \frac{1}{d_i^p} Z_i}{\sum_{i=1}^n \frac{1}{d_i^p}} \quad (2)$$

Spline

El algoritmo *Spline* es un modelo estocástico local que ajusta funciones polinómicas de base radial con el fin de minimizar la curvatura y suavizar los datos, donde los resultados obtenidos dependen de un parámetro de tensión. Al igual que el IDW, este método no depende de la distribución estadística de los datos observados, sin embargo, sí depende de su distribución espacial, es decir que se encuentra afectada por la densidad de muestreo (área [km²]/número de estaciones de

monitoreo) (J. Li & Heap, 2014; Londoño Ciro & Valdés Quintero, 2012). El valor a estimar (Z_0), se obtiene mediante la Ecuación 3, donde n es la cantidad de puntos, a_1 y λ_i son coeficientes determinados por la solución de un sistema de ecuaciones lineales, r_i es la distancia entre la ubicación del punto a estimar (x_0) y el valor disponible (x_i), φ^2 es un parámetro de peso, K_0 es la función de Bessel modificada y c es 0.577215 (Londoño, 2018a).

$$Z_0(x_0, y_0) = a_1 \sum_{i=1}^n \lambda_i R(r_i), \quad R(r_i) = \frac{1}{2\pi\varphi^2} \left[\ln\left(\frac{r_i}{2}\varphi\right) + c + K_0(r_i\varphi) \right] \quad (3)$$

Kriging

El algoritmo Geoestadístico Univariado (Kriging), es un enfoque estocástico local que permite estimar datos espaciales al modelar el cambio de la varianza de los datos con la distancia, para calcular λ_i (Ecuación 1). Es considerado un método robusto, donde los datos de entrada deben distribuirse normalmente en el espacio, es decir, ser estacionarios, y se debe tener una muestra con una densidad de muestreo baja (J. Li & Heap, 2014; Londoño, 2018a; Oliver & Webster, 2014). Por ende, el primer paso para aplicar el algoritmo Kriging, consiste en un análisis exploratorio o geoestadístico de los datos espaciales, en donde, se busca que los datos cumplan con los siguientes criterios (Londoño, 2018a):

- Similitud en las medidas de tendencia central: *Media* \cong *Mediana* \cong *Moda*
- El histograma de frecuencias de los datos espaciales debe ser monomodal y simétrico respecto a la media
- $|Sesgo \text{ o } skewness| < 0.5$. De lo contrario, los datos deben ser transformados, aplicando el criterio de Webster y Oliver, expuesto en la Ecuación 4.

$$\begin{cases} 0.5 < |Sesgo| \leq 1; & \text{Transformación con raíz cuadrada} \\ |Sesgo| > 1; & \text{Transformación logarítmica} \end{cases} \quad (4)$$

- En caso de que los datos transformados no sean estacionarios, se evalúa la presencia y posible remoción de valores extremos en los datos o *outliers*.

Luego de garantizar la estacionariedad en los datos, se realiza un análisis estructural por medio del cálculo del semivariograma empírico y su ajuste a un modelo de semivariograma teórico. En esta parte del análisis, se busca determinar un modelo matemático que represente el cambio de la varianza de los datos con la distancia. El semivariograma empírico es un gráfico de distancia euclidiana (h) contra semivarianza empírica (γ^*), calculada mediante la Ecuación 5, en donde Z_{i+h} es el valor del dato espacial en las coordenadas (x_{i+h}, y_{i+h}) (Londoño, 2018a).

$$\gamma^* = \frac{\Delta Z_i^2}{2}, \quad Z_i^2 = (Z_{i+h} - Z_i)^2 \quad (5)$$

Un semivariograma es una función monótona creciente que alcanza un valor límite asintótico denominado meseta (*partial sill* - θ_s) y es equivalente a la varianza de los datos, por ende tiene como unidades las del dato espacial al cuadrado (Figura 1). El valor de distancia (h) donde se alcanza la

meseta es llamado el rango (*major range* - θ_r) en unidades de longitud, y determina la distancia hasta el punto donde la autocorrelación espacial es nula. Cuando el origen del semivariograma es diferente de cero, se habla de una discontinuidad en el origen o efecto pepita (*nugget* - θ_n), indicando que la variable estudiada posee valores inferiores al valor mínimo de los datos de la muestra. En la Ecuación 6 se muestra la función del semivariograma teórico (Londoño, 2018a).

$$\gamma = \theta_n + \gamma(\theta_s, \theta_r, \theta_n) \quad (6)$$

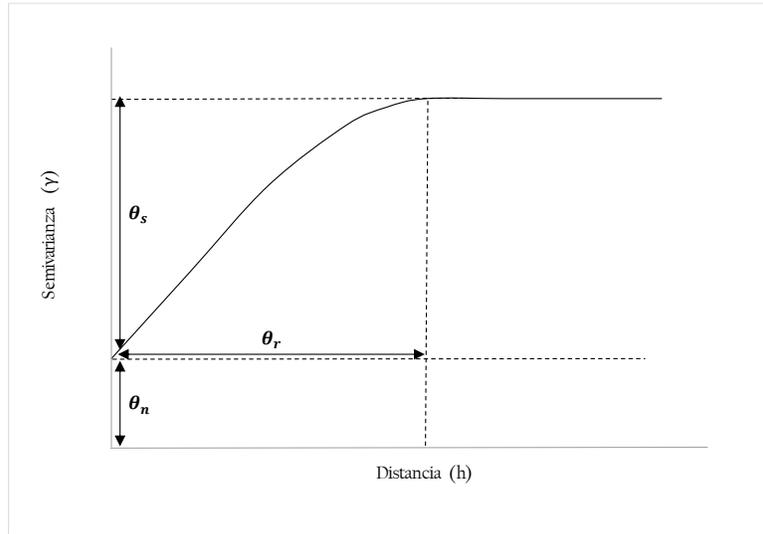


Figura 1. Semivariograma teórico

Luego, para estimar los datos espaciales con el algoritmo Kriging se calculan los pesos ponderados λ a partir de la Ecuación 7, donde previamente se ajusta uno de los siguientes modelos de semivariograma teórico al semivariograma empírico: *Spheric* (SP), *Gaussian* (GS), *Exponential* (EX), *Tetraspheric* (TS), *Pentaspheic* (PS), *Rational Cuadratic* (RC), *Hole Effect* (HE), *J- Bessel* (JB) y *K-Bessel* (KB), con el fin de calcular los elementos de la matriz Γ , que representan el cambio de la semivarianza con la distancia entre los datos conocidos, y g , el cual es un vector que representa el cambio de la semivarianza con la distancia entre el dato a estimar y los datos conocidos (Londoño, 2018a).

$$\lambda = \Gamma^{-1} * g \quad (7)$$

– Gradientes espaciales

Un gradiente espacial ∇Z , es una técnica de análisis espacial que describe la tasa de cambio de un escalar Z respecto a las coordenadas de posición (x_i, y_i) y (x_j, y_j) y está dado por la Ecuación 8, donde ∇Z_{ij} es el cambio del datos entre las posiciones i y j , ∇x_{ij} y ∇y_{ij} representan el cambio de la coordenada x y y entre las posiciones i y j , respectivamente (Londoño, 2018a).

$$\nabla Z = \tan^{-1} \sqrt{\left(\frac{\nabla Z_{ij}}{\nabla x_{ij}}\right)^2 + \left(\frac{\nabla Z_{ij}}{\nabla y_{ij}}\right)^2} * \frac{180}{\pi} \quad (8)$$

– **Regresión espacial multivariada**

El análisis de regresión es un enfoque estocástico que se puede utilizar para explorar relaciones espaciales, para mejorar la comprensión de los factores detrás de los patrones espaciales observados o para hacer predicciones. El algoritmo de regresión espacial permite realizar estimaciones locales de una variable dependiente Z , mediante la implementación del método de Mínimos Cuadrados Ordinarios (*Ordinary Less Squares – OLS*), basándose en la presencia de posibles variables explicativas (Z_j) en el mismo punto (x_i, y_i) , como se observa en la Ecuación 9 (Fischer & Getis, 2010; Londoño, 2018a; Maantay & McLafferty, 2011).

$$Z(x_i, y_i) = \beta_0 + \sum_{j=1}^n \beta_j Z_j(x_i, y_i) + \varepsilon_i, \quad \varepsilon_i \sim N(0, v^2) \quad (9)$$

El modelo conceptual es un componente integral del modelado espacial y debe seleccionarse de manera que represente mejor la estructura de dependencia espacial entre las características que se analizan. Para esto, se debe determinar previamente la mejor combinación de las variables explicativas, mediante la adición de cada variable de manera supervisada por pasos, en donde se seleccionan las variables que cumplan las siguientes condiciones (Eeftens et al., 2012; ESRI, 2019b; Van Nunen et al., 2017):

- Variable significativa (valor $p < 0.05$).
- El signo para el coeficiente de la variable (β_i) en el modelo debe ajustarse a la dirección de efecto esperada.
- Las variables ingresadas más adelante no deben alterar las condiciones anteriores, es decir, no deben cambiar la significancia y el signo de las variables ya seleccionadas.

Además, las variables seleccionadas deben especificar adecuadamente el modelo de regresión, mediante la evaluación de los criterios de especificación del método OLS (ESRI, 2019b; Londoño, 2018a):

- ***Criterios de rendimiento del modelo:*** Determina si las variables explicativas pueden explicar la variación de la variable dependiente. Se espera que el coeficiente de determinación ajustado (AdjR^2) sea mayor a 0.5 y que el criterio corregido de información de Akaike (AICc) sea mínimo.
- ***Criterio de consistencia:*** Determina si los residuos del modelo se distribuyen normalmente, permitiendo conocer la ausencia de alguna variable clave para explicar el comportamiento de la variable dependiente. Se utiliza el valor p del estadístico de Jarque-Bera (JB), donde la hipótesis nula establece que los datos tienen la asimetría y la curtosis de una distribución normal. Se espera que JB sea mayor a 0.1.

- *Criterio de ausencia de redundancia o multicolinealidad:* Permite evaluar si las variables explicativas son independientes entre sí o se pueden expresar como una combinación lineal de otras variables explicativas (son multicolineales). Para ello se usa un factor de inflación de la varianza (VIF) menor a 7.5.
- *Criterio de significancia:* Permite determinar si el coeficiente de las variables explicativas es significativo y si el modelo tiene problemas de heteroscedasticidad, es decir que la varianza de los residuos no es constante. Se utiliza el estadístico de Breusch-Pagan estudentizado de Koenker (K(BP)), donde la hipótesis nula para esta prueba establece que el modelo es estacionario, es decir que un K(BP) menor a 0.05, indica heteroscedasticidad o no estacionariedad estadísticamente significativa.
- *Criterio de ausencia de residuos en la autocorrelación espacial:* Para determinar si los residuos son espacialmente aleatorios o no presentan autocorrelación espacial. Se utiliza el valor p del Índice Global de Moran (SA). Se espera que el valor p sea mayor a 0.1.

En el pronóstico de material particulado en áreas urbanas, se han implementado ampliamente los modelos de Regresión Espacial Multivariada, mostrando un buen desempeño, sin necesidad de información exhaustiva sobre variables atmosféricas, como un modelo de dispersión (Eeftens et al., 2012; Jerrett et al., 2005; Londoño, 2018b; Zhao, Zhan, Yao, & Yang, 2020). Con el método OLS se construye una regresión lineal global para modelar las relaciones entre un conjunto de variables de datos, donde el método supone la estacionariedad espacial de la relación entre la variable dependiente y las variables independientes, generando un solo coeficiente de regresión para cada variable independiente (Fischer & Getis, 2010; Maantay & McLafferty, 2011).

Cuando las relaciones no son consistentes en el área de estudio, es decir que su comportamiento varía, el modelo global se vuelve menos efectivo y puede no ser una buena representación del comportamiento a modelar. Los modelos de regresión espacial locales, como la Regresión Ponderada Geográficamente (*Geographically Weighted Regression - GWR*), examinan las heterogeneidades espaciales en la relación, generando en cada punto un coeficiente de regresión local para cada variable independiente. Las observaciones más cercanas tienen más influencia en la estimación del conjunto local de coeficientes que las observaciones más lejanas, implicando que las relaciones que está tratando de modelar pueden cambiar en el área de estudio. En esencia, el GWR mide las relaciones inherentes alrededor de cada punto de regresión, donde cada observación cercana se pondera mediante una función de la distancia al punto de regresión y los conjuntos de coeficientes se estiman mediante mínimos cuadrados ponderados, debido a que los valores de las varianzas de los datos observados son diferentes, es decir, hay heteroscedasticidad, y no existen correlaciones entre las variaciones observadas (Fischer & Getis, 2010; Lu, Charlton, Harris, & Fotheringham, 2014; Zhao et al., 2020).

El GWR, es un método estocástico localmente lineal, con un inmenso potencial para la interpolación espacial de datos ambientales, basándose en una regresión tradicional e incorporando relaciones espaciales locales de manera intuitiva y explícita (J. Li & Heap, 2011). Esta técnica relaja la suposición de relaciones constantes espacialmente, creando múltiples ecuaciones para describir tales relaciones, capturando de esta manera, relaciones no estacionarias espacialmente entre la variable dependiente y las variables explicativas, mediante la incorporación de información geográfica. En otras palabras, el modelo GWR predice las relaciones entre la variable dependiente y las variables predictoras en cada ubicación (x_i, y_i) , mediante la construcción de ecuaciones de regresión locales o específicas de la misma (Lu et al., 2014; Zhao et al., 2020).

– Validación de los modelos espaciales

Después de estimar datos espaciales para una variable mediante alguno de los métodos anteriores, es necesario evaluar su desempeño, es decir, qué tan parecidas son las predicciones a los datos observados. Esto se realiza mediante la aplicación del método de validación cruzada dejando uno fuera (*Leave One Out*), en donde se retira un valor observado O_i en el sitio de monitoreo i , para luego ejecutar el algoritmo de interpolación o regresión espacial seleccionado y estimar el valor E_i en ese mismo punto. A partir del valor observado y el estimado, se calculan los residuales o el error de estimación r_i , mediante la Ecuación 10 y se procede a calcular los estadísticos de error (Tabla 5), donde σ_{r_i} es el error estándar, μ_{r_i} la media de los residuales, n la cantidad de datos y σ la desviación estándar de r_i (Londoño, 2018a).

$$r_i = O_i - E_i \quad (10)$$

Tabla 5. Estadísticos de error calculados para la validación de los datos estimados

Estadístico de error	Ecuación
Error Estándar Promedio (<i>Average Standard Error</i> - ASE)	$ASE = \frac{\sum_{i=1}^n \sigma_{r_i}}{n}, \sigma_{r_i} = \sqrt{(r_i - \mu_{r_i})^2}$ (11)
Raíz del Error Cuadrático Medio (<i>Root Mean Square Error</i> - RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^n (r_i)^2}{n}}$ (12)
Raíz del Error Cuadrático Medio Estandarizado (<i>Root Mean Square Error Standarized</i> - RMSS)	$RMSS = \frac{RMSE}{\sigma}$ (13)
Error Absoluto Relativo (<i>Absolute Relative Error</i> - ARE)	$ARE = \frac{ r_i }{O_i}$ (14)
Error Absoluto Relativo Promedio (<i>Mean Absolute Relative Error</i> - MARE)	$MARE = \frac{\sum_{i=1}^n ARE}{n_i}$ (15)

Fuente: Londoño Ciro, L. A. (2018). Metodología para caracterizar espacio-temporalmente la concentración de material particulado en valles intramontanos con información escasa. Universidad de Antioquia.

4.1.4. Modelos de regresión de usos del suelo – LUR

Existe una diversidad de modelos matemáticos útiles para predecir el valor de un contaminante, en puntos diferentes a las estaciones de monitoreo, en función de otra serie de covariables meteorológicas y no meteorológicas relacionadas. Algunos modelos aplicados, relacionados con la problemática expuesta son los modelos de dispersión gaussiana, estocásticos, de redes neuronales, de regresión de usos del suelo, entre otros (Beelen et al., 2013, 2010; de Hoogh et al., 2014; Eeftens et al., 2012; Mölter et al., 2010a; Mölter, Lindley, de Vocht, Simpson, & Agius, 2010b).

Un modelo de regresión de usos del suelo (*Land Use Regression* - LUR) se usa a menudo para analizar la contaminación, particularmente en áreas densamente pobladas. El modelo se basa en patrones de contaminación predecibles para estimar concentraciones en un área en particular. Esto requiere algún vínculo con las características ambientales del área, especialmente las que influyen en la intensidad de las emisiones contaminantes y la eficiencia de la dispersión. Los modelos LUR permiten calcular la relación entre una variable respuesta, los valores del contaminante, y otras variables explicativas o covariables, a partir de la construcción de un modelo de regresión con datos espaciales (Eeftens et al., 2012; Londoño & Cañón, 2015).

El modelado se puede realizar utilizando datos vectoriales (puntos, líneas y polígonos) o datos ráster. Para obtener las posibles variables explicativas espaciales, se definen áreas de influencia o **buffers** circulares con un radio fijo y centro en las estaciones de monitoreo, los cuales se intersecan con cada covariable, como la malla vial, usos del suelo, población, entre otras; permitiendo evaluar si el efecto de una variable en específico puede verse influenciado por su proximidad a las estaciones de monitoreo del contaminante definido (Briggs, 2005; Briggs et al., 1997; Eeftens et al., 2012; Mercer et al., 2011).

Los modelos LUR se han utilizado con frecuencia para explicar los contrastes espaciales de pequeña escala en las concentraciones de contaminantes del aire y estimar la exposición individual para los participantes en los estudios de cohortes., utilizando variables explicativas derivadas de sistemas de información geográfica (SIG) (Eeftens et al., 2012). Las variables explicativas más comunes son la densidad de población, los usos del suelo urbano, variables relacionadas con el tráfico (p.ej. redes de vías principales), distancia de cada sitio de monitoreo a la vía principal más cercana, modelos de elevación digital, buffers de diferentes distancias, siendo los más comunes de 50, 100, 300, 500 y 1000 m (Bertazzon et al., 2015; Cordioli et al., 2017; de Hoogh et al., 2016, 2014; Eeftens et al., 2012; Londoño & Cañón, 2015; Marshall et al., 2008; Shi et al., 2017).

Por otro lado, algunos estudios han evaluado la posible adición de variables explicativas alternas a las aplicadas con regularidad en la literatura, como: la disponibilidad del viento representada por variables geomorfométricas (índice de área frontal, índice de barlovento/sotavento, rugosidad superficial) (Cordioli et al., 2017; Shi et al., 2017), velocidades del viento interpoladas (Arain et al., 2007; Bertazzon et al., 2015), estimaciones de las concentraciones de contaminantes a partir de modelos químicos de transporte (Marshall et al., 2008) y de imágenes satelitales (de Hoogh et al., 2016), niveles de ruido y precipitación (Londoño & Cañón, 2015).

Con las variables definidas, se realiza una caracterización espacial multivariada mediante la aplicación de la Ecuación 9 y los criterios expuestos en la sección 4.1.3, obteniéndose un modelo de regresión que describe la asociación entre las variables explicativas o predictoras y las concentraciones medidas, el cual, mediante su ejecución, permite estimar los niveles del contaminante en otras ubicaciones objetivo y/o un mapa de concentraciones estimadas (Londoño & Cañón, 2015; Mercer et al., 2011; Van Nunen et al., 2017).

– Ventajas

Los sistemas de información geográfica ofrecen herramientas potentes para la evaluación de la exposición al proporcionar un medio para capturar y vincular los datos espaciales dentro de una misma estructura geográfica, permitiendo mejorar la integración y la coherencia de los datos; los cuales, por medio de la geomática, son analizados, modelados y administrados mediante la integración de conjuntos de diferentes herramientas y tecnologías (Briggs, 2005; Eeftens et al., 2012). El modelado de LUR, en general, permite explicar la variabilidad espacial intraurbana de los contaminantes a una escala más fina, por la incorporación de las fuentes móviles por medio de variables de tráfico (Eeftens et al., 2012; Mølter et al., 2010b), además, su implementación permite que los cálculos se extiendan a poblaciones y áreas de estudio mucho más grandes (Briggs, 2005).

Una de las grandes ventajas de LUR, en comparación con los modelos de dispersión de contaminantes atmosféricos, está relacionada a que éste es un modelo computacionalmente simple y eficiente, teniendo en cuenta que la carga computacional puede aumentar sustancialmente cuando se utilizan datos vectoriales, en lugar de rasterizados. Gracias a esto, la implementación de los modelos LUR en estudios epidemiológicos cada vez se hace más extensa debido al bajo costo y a la fácil implementación a partir de datos registrados en las redes de monitoreo existentes, o de campañas de monitoreo diseñadas específicamente para el estudio, y predictores geográficos de fuentes de contaminación del aire (de Hoogh et al., 2014).

– Limitaciones

Los modelos LUR desarrollados en la literatura ofrecen un método simple y efectivo para estimar las variaciones de los contaminantes atmosféricos a una escala espacial más fina que la utilizada por un modelo de dispersión o un método de interpolación, sin embargo, dichas estimaciones se encuentran limitadas por la gran dependencia del modelo a la cantidad y distribución uniforme de las estaciones de monitoreo ubicadas en el área de estudio, es decir, la disponibilidad de sólo unas cuantas estaciones de monitoreo para la evaluación de las concentraciones de un contaminante en un área de gran tamaño, con una distribución espacial no uniforme de la red de monitoreo, puede aumentar el error asociado a las estimaciones, dando lugar a un posible sobreajuste del modelo (Dabass et al., 2016; de Hoogh et al., 2014; Eeftens et al., 2012; Londoño & Cañón, 2015).

Por otro lado, el ajuste del modelo LUR también varía según el contaminante evaluado, debido a que se ha observado que el rendimiento para los modelos desarrollados para material particulado es menor que para el dióxido y el monóxido de nitrógeno, observación generalmente asociada a menor variación espacial del material particulado, caracterización menos precisa de la fuente, falta

de descriptores de usos del suelo para este contaminante, y la naturaleza de estos gases, los cuales, en algunos casos, pueden estar más relacionados al proceso de combustión que el mismo material particulado (de Hoogh et al., 2014; Shi et al., 2017). Además, es importante mencionar que en esta metodología no se tienen en cuenta los patrones de actividad de la población o la contaminación del aire en interiores, lo cual representa una gran simplificación desde un punto de vista epidemiológico (de Hoogh et al., 2014).

4.2. Marco legal

4.2.1. Marco jurídico

En la Tabla 6 se presentan los decretos y resoluciones que han estado en vigencia en Colombia, relacionados con la calidad del aire ambiente.

Tabla 6. Decretos y resoluciones relacionados a la calidad del aire extradomiciliario, Colombia, 2019.

Regulación	Fecha de expedición	Entidad	Objetivo
Decreto 948 de 1995	5 de junio de 1995	Ministerio del Medio Ambiente	Reglamento de Protección y Control de la Calidad del Aire
Decreto 979 de 2006	3 de abril de 2006	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Modifica algunos artículos del decreto 948 de 1995 relacionados con las normas de la calidad del aire, prevención de la contaminación y planes de contingencia; áreas fuente de contaminación.
Resolución 601 de 2006	4 de abril de 2006	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Se establece la Norma de Calidad del Aire o Nivel de Inmisión, para todo el territorio nacional en condiciones de referencia
Resolución 909 de 2008	5 de junio de 2008	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Normas y estándares de emisión admisibles de contaminantes a la atmósfera por fuentes fijas.
Resolución 910 de 2008	5 de junio de 2008	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Niveles permisibles de emisión de contaminantes para fuentes móviles terrestres.
Resolución 610 de 2010	24 de marzo de 2010	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Modifica la Resolución 601 de 2006
Resolución 650 de 2010	29 de marzo de 2010	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Se adopta a nivel nacional el Protocolo para el Monitoreo y Seguimiento de la Calidad del Aire
Resolución 651 de 2010	29 de marzo de 2010	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Creación del Subsistema de Información sobre Calidad del Aire – SISAI
Resolución 2154 de 2010	2 de noviembre de 2010	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Modifica el Protocolo para el Monitoreo y Seguimiento de Calidad del Aire, adoptado a través de la Resolución 650 de 2010.
Decreto 1076 de 2015	26 de mayo de 2015	Ministerio de Ambiente y Desarrollo Sostenible	Por medio del cual se expide el Decreto Reglamentario Único del Sector Ambiente. La Sección 2 establece las disposiciones generales sobre normas de calidad del aire, niveles de contaminación, emisiones contaminantes y de ruido

Regulación	Fecha de expedición	Entidad	Objetivo
Resolución 1309 de 2010	13 de julio de 2017	Ministerio de Ambiente y Desarrollo Sostenible	Modifica algunos artículos de la Resolución 909 de 2008
Resolución 2254 de 2017	1 de noviembre de 2017	Ministerio de Ambiente y Desarrollo Sostenible	Se adopta la Norma de calidad del aire ambiente para Colombia

4.2.2. Marco normativo

En la Tabla 7 se presentan los acuerdos, protocolos y otros documentos vigentes en Colombia, relacionados con la calidad del aire ambiente.

Tabla 7. Documentos vigentes relacionados a la calidad del aire extradomiciliario, 2019

Documento	Fecha de expedición	Entidad	Objetivo
Política de Prevención y Control de la Contaminación del Aire	2010	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Impulsa la gestión de la calidad del aire en el corto, mediano y largo plazo, para proteger la salud y el bienestar humano, en el marco del desarrollo sostenible.
Protocolo para el monitoreo y seguimiento de la calidad del aire	Marzo de 2010	Ministerio de Ambiente, Vivienda y Desarrollo Territorial	Manual de operación de sistemas de vigilancia de la calidad del aire
			Manual de diseño de sistemas de vigilancia de la calidad del aire
CONPES 3550 de 2008	24 de noviembre de 2008	Departamento Nacional de Planeación	Lineamientos para la formulación de la política integral de salud ambiental, con énfasis en los componentes de calidad de aire, calidad de agua y seguridad química
Acuerdo 48 de 2014	17 de diciembre de 2014	Alcaldía de Medellín	Se adopta la revisión y ajuste de largo plazo del Plan de Ordenamiento Territorial del Municipio de Medellín
Acuerdo Metropolitano Nro. 16 de 2017	6 de diciembre de 2017	Área Metropolitana del Valle de Aburrá	Se adopta el Plan Integral de Gestión de la Calidad del Aire en el Valle de Aburrá - PIGECA
Resolución metropolitana 912 de 2017	19 de mayo de 2017	Área Metropolitana del Valle de Aburrá	Se adoptan medidas en el sector industrial que contribuyan al desarrollo de una gestión integral de la calidad del aire en la jurisdicción del Área Metropolitana del Valle de Aburrá
Acuerdo Metropolitano Nro. 4 de 2018	19 de febrero de 2018	Área Metropolitana del Valle de Aburrá	Se deroga el Acuerdo Metropolitano Nro. 15 de 2016 y se adopta el nuevo protocolo del Plan Operacional para Enfrentar Episodios de Contaminación Atmosférica - POECA, en la jurisdicción del Área Metropolitana del Valle de Aburrá
Resolución metropolitana 2231 de 2018	31 de agosto de 2018	Área Metropolitana del Valle de Aburrá	Se declaran unas Zonas Urbanas de Aire Protegido -ZUAP, dentro de la jurisdicción del Área Metropolitana del Valle de Aburrá
CONPES 3943 de 2018	31 de julio de 2018	Departamento Nacional de Planeación	Política para el mejoramiento de la calidad del aire

5. Metodología

Diversas metodologías se han implementado para estudiar la variación espacial de contaminantes atmosféricos, en donde, su selección depende del problema que se desee solucionar, la aplicación de los resultados y la información disponible. Partiendo de la limitación que supone el desconocimiento parcial de la variación intraurbana de las partículas finas en el Valle Aburrá, la cual restringe la implementación de estudios epidemiológicos que permitirían afianzar en el entendimiento de los efectos a la salud ocasionados por este contaminante, se decidió explorar la construcción de modelos de regresión de usos del suelo, para evaluar la variación intraurbana del $PM_{2.5}$ en el Valle de Aburrá, teniendo en cuenta que la principal fortaleza de esta metodología es la estructura empírica del mapeo de regresión, que permite la adaptación a áreas locales sin monitoreo adicional o adquisición de datos, generando resultados estadísticamente confiables, con costos relativamente bajos (de Hoogh et al., 2014; Eeftens et al., 2012; Jerrett et al., 2005). Además, en la zona de estudio se cuenta con una red de monitoreo de calidad del aire, una red de cámaras de fotodetección que permiten el acceso a conteos vehiculares y mapas de usos del suelo, para implementar esta metodología.

Con el fin de evaluar la variación espacial intraurbana de las concentraciones de $PM_{2.5}$, se modeló, por medio del método de regresión de usos del suelo, el comportamiento de las concentraciones mensuales de éste contaminante durante el año 2018, en la ciudad de Medellín, a partir de la Red de Monitoreo de Calidad del Aire vigente en el Valle de Aburrá y otras variables explicativas. La metodología aquí descrita es consistente con la desarrollada y estandarizada por el *European Study of Cohorts for Air Pollution Effects* (ESCAPE) para desarrollar modelos LUR (Eeftens et al., 2012) y con la encontrada en diferentes estudios (Allen et al., 2011; Bertazzon et al., 2015; Cordioli et al., 2017; Londoño & Cañón, 2015; Van Nunen et al., 2017).

5.1. Generalidades del estudio

– Tipo de estudio

El presente fue un estudio de tipo correlacional, debido a que se evaluó el grado de asociación entre diferentes variables presuntamente relacionadas (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2010), por medio de la elaboración de regresiones espaciales mensuales para $PM_{2.5}$.

– Periodo de análisis

El análisis se realizó en el periodo comprendido entre el 1ro de enero y el 31 de diciembre de 2018, debido a la disponibilidad de la información requerida, principalmente, a la cantidad de estaciones de monitoreo de $PM_{2.5}$ en el Valle de Aburrá, la cual fue densificada a finales del año 2017.

– Escala temporal

Previamente, en estudios desarrollados en el Valle de Aburrá con el fin de analizar el comportamiento de los niveles de material particulado en el tiempo, se ha observado la presencia

de un ciclo diurno con un comportamiento bimodal de las concentraciones, relacionado con los momentos de mayor flujo vehicular (EAFIT, 2019). Además, se ha evidenciado en el ciclo anual una estacionalidad marcada en las series de material particulado, en donde, se observa un incremento en las concentraciones de PM₁₀ y PM_{2.5} entre febrero y marzo, disminuciones entre junio y julio, y un comportamiento tendiente al promedio anual entre septiembre y octubre (Área Metropolitana del Valle de Aburrá, 2018; EAFIT, 2019; Londoño, 2018b). En estos estudios, se ha indicado que el valor máximo observado se presenta en los meses de transición entre la época seca y el periodo de lluvias en la Región Andina Colombiana, propiciando las condiciones meteorológicas, que, en conjunto con la orografía del valle, inciden en el incremento estacional de los niveles de PM_{2.5} (Área Metropolitana del Valle de Aburrá, 2018; EAFIT, 2019).

Debido a la aplicación de los modelos LUR en la evaluación de la exposición en estudios epidemiológicos, donde no es prudente utilizar una escala horaria para evaluar asociaciones con enfermedades relacionadas a la contaminación atmosférica, se ha observado que la escala temporal predominante para aplicar la metodología es la anual (Área Metropolitana del Valle de Aburrá, 2018; Eeftens et al., 2012). Por esta razón, adicionando las características del estudio, en donde se desea evaluar la influencia de determinadas variables en los niveles de PM_{2.5}, con el fin de aportar al entendimiento del comportamiento del mismo; para esta investigación se estableció la escala mensual para la construcción de los modelos LUR, es decir, se construyeron modelos mensuales durante un año, debido a la estacionalidad mensual marcada que presentan las concentraciones del PM_{2.5} en el Valle de Aburrá (Área Metropolitana del Valle de Aburrá, 2018).

– Sistema de coordenadas utilizadas

Sistema de Coordenadas Proyectadas: PCS MAG Ant Medellín

Sistema de Coordenadas Geográficas: GCS_MAGNA

– Zona de estudio

El Valle de Aburrá se caracteriza por ser un valle estrecho ubicado en la cordillera de Los Andes, con topografía irregular. Éste se encuentra entre 1300 y 2800 metros sobre el nivel del mar y se extiende por aproximadamente 60 kilómetros de longitud con un ancho variable (Zapata, Cano, Ramírez, Rubiano, & Jiménez, 2015). Por su ubicación en la región tropical ecuatorial, presenta un clima subtropical subhúmedo, donde la temperatura oscila entre 16 y 28°C con una media de 22°C. La precipitación acumulada anual varía entre 1400 mm hasta los 2700 mm, presentándose un régimen de lluvias bimodal, con épocas lluviosas en los meses de mayo y octubre (Guerrero Hoyos & Aristizábal Giraldo, 2019).

El Área Metropolitana del Valle de Aburrá se encuentra conformado por 10 municipios, de los cuales, las cabeceras municipales o zonas urbanas de Medellín, Bello, Copacabana, Envigado, Itagüí, La Estrella y Sabaneta, conformaron la zona de estudio (Figura 2). Los municipios seleccionados comprenden un área de 737.12 km², donde 142.56 km² corresponden a las cabeceras municipales. Éstos son centros densamente poblados con 3,556,157 habitantes en total, donde el 66.7% se encuentran ubicados en Medellín (Departamento Nacional de Estadística, 2019).

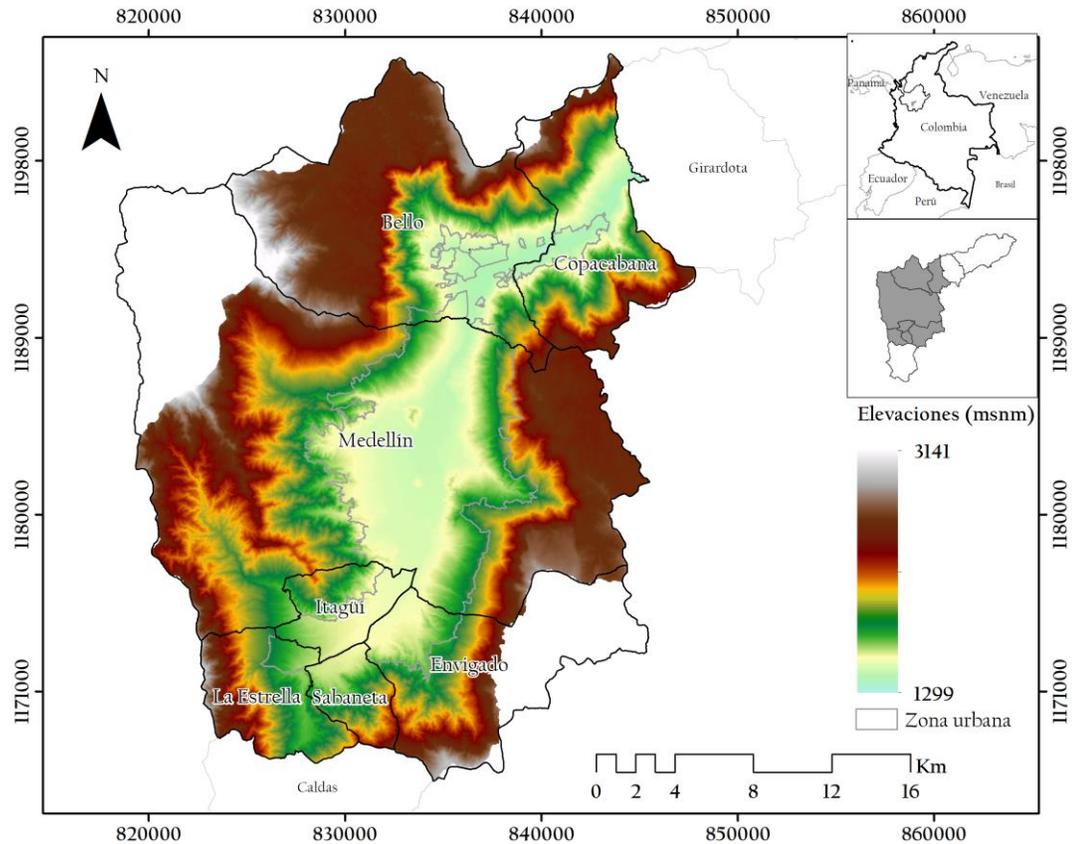


Figura 2. Ubicación y mapa de elevaciones de la zona de estudio

5.2. Diseño experimental y metodológico

Para la construcción de modelos de regresión de usos del suelo, se consideró como variable dependiente, la concentración promedio mensual de $PM_{2.5}$ de las estaciones de monitoreo. Por otro lado, se definieron cinco (5) grupos para clasificar todas las posibles variables explicativas, según la procedencia de las mismas (Tabla 8).

Tabla 8. Grupos de clasificación de las posibles variables explicativas

Grupo	Definición
1	VARIABLES DE POSICIÓN
2	VARIABLES AMBIENTALES
3	VARIABLES DE USOS DEL SUELO
4	VARIABLES DE MOVILIDAD
5	VARIABLE DEMOGRÁFICA

Se definieron diferentes buffers circulares, con radios de 50, 100, 150, 200, 300 y 500 m y centro en las coordenadas de los sitios de monitoreo de la variable dependiente. Con cada buffer se construyó un modelo de regresión de usos del suelo en donde, para cada mes, permanecieron fijos los valores de concentración promedio mensual de $PM_{2.5}$, las variables de posición (Grupo 1) y las variables ambientales (Grupo 2). En cambio, las variables de usos del suelo (Grupo 3), las variables de movilidad (Grupo 4) y la variable demográfica (Grupo 5), fueron calculadas para cada buffer por medio de intersecciones espaciales, teniendo en cuenta la variación mensual del contaminante. Por otro lado, a medida que se cambiaba el mes, se modificaban las variables del Grupo 2 y una variable del Grupo 4 (flujo vehicular), mientras que las demás variables permanecían fijas.

Con el diseño anterior, se construyeron 6 modelos para cada mes del año 2018, correspondientes a cada buffer evaluado. Dando como resultado, 72 modelos para todo el periodo de análisis (Tabla 9).

Tabla 9. Número de modelos construidos teniendo en cuenta los seis buffers explorados

Buffer	Datos fijos para cada buffer	Datos calculados para cada buffer
50 m	<ul style="list-style-type: none"> - Concentración promedio mensual de $PM_{2.5}$ - Variables Grupo 1 - Variables Grupo 2 	<ul style="list-style-type: none"> - Variables Grupo 3 - Variables Grupo 4 - Variable Grupo 5
100 m		
150 m		
200 m		
300 m		
500 m		
Total	6 modelos LUR para cada mes	
Diseño experimental = 6 LUR x 12 meses = 72 modelos LUR		

En general, la construcción y selección del mejor modelo LUR, conlleva un proceso riguroso. Con el fin de facilitar el entendimiento del ejercicio realizado, en este trabajo se optó por dividir la metodología en cuatro momentos principales, en donde, el Momento 1 consistió en el análisis exploratorio de las series de tiempo correspondientes a las variables ambientales y los aforos vehiculares, las cuales debían pasar por un proceso de análisis de calidad del dato y de selección, para garantizar condiciones óptimas de la información disponible en todo el periodo de análisis. Los datos de las series temporales seleccionadas en la escala temporal definida previamente, en conjunto con las demás variables espaciales, como los usos del suelo y la malla vial, pasaron por un preprocesamiento importante, relacionado con la intersección de cada posible variable explicativa con los seis buffers definidos, con el fin de generar las variables predictoras SIG, insumo para la construcción del modelo LUR. Todo este proceso, fue agrupado en el Momento 2. Luego, en el Momento 3, con la información espacial preparada para cada buffer y cada mes, se procedieron a aplicar unos criterios de selección para definir las posibles variables explicativas a evaluar en el modelo y, por consiguiente, construir cada modelo LUR. Por medio de los criterios de selección del método OLS, se seleccionó el mejor modelo para cada buffer en cada mes.

Finalmente, en el Momento 4, evaluando los mismos criterios mencionados con los seis modelos construidos para cada mes, se seleccionó el mejor modelo mensual y con estos, se estimaron los valores para $PM_{2.5}$ para la zona de estudio. En la Figura 3, se muestra un diagrama que relaciona los Momentos descritos con cada uno de los objetivos específicos establecidos en el trabajo de investigación.

Momentos metodológicos

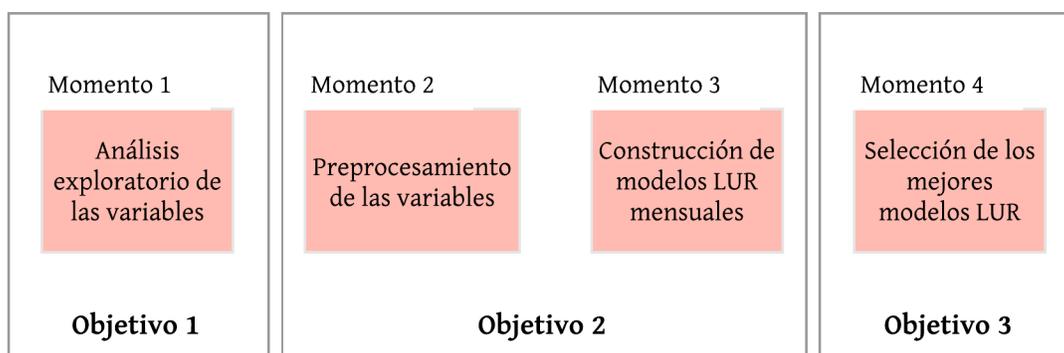


Figura 3. Momentos metodológicos principales

5.3. Momento metodológico uno: Análisis exploratorio de las variables

5.3.1. Recolección de información

Este trabajo se desarrolló con información secundaria, recolectada previamente por diferentes entidades públicas de los municipios que conforman el Valle de Aburrá. La información sobre la variable dependiente y las explicativas fue solicitada a las entidades referidas en la Tabla 10.

Tabla 10. Información necesaria para la construcción de modelos LUR y fuente de información

Entidad	Información solicitada
Secretaría de Movilidad de cada municipio que conforma el Área Metropolitana del Valle de Aburrá	Aforos vehiculares para 2018. Censo de vehículos por una vía durante un día.
	Vehículos matriculados en el año 2018.
	Malla vial actualizada del municipio.
Departamento de Planeación de Medellín	Malla vial y mapas de usos del suelo urbano de Medellín, para 2018
Área Metropolitana del Valle de Aburrá	Malla vial y mapas de usos del suelo urbano para todos municipios del Área Metropolitana del Valle de Aburrá, para 2018
Portal Geográfico del Municipio de Medellín (online) ¹	Malla vial, jerarquía vial, mapas de usos del suelo urbano actualizados, para 2018
Sistema de Alertas Tempranas del Valle de Aburrá – SIATA (online) ²	Concentraciones de $PM_{2.5}$ en el Valle de Aburrá en escala horaria, para 2018

Entidad	Información solicitada
	Ubicación de estaciones de PM _{2.5} para 2018
	Series de variables meteorológicas (temperatura, velocidad del viento y precipitación) en escala horaria o minutal, para 2018
	Ubicación de estaciones meteorológicas para 2018
Sistema Sigma – Universidad Pontificia Bolivariana (online) ³	Concentraciones de PM _{2.5} en el Valle de Aburrá en escala diaria, para 2018
Alcaldía de Medellín (online) ⁴	Proyección de población 2016 a 2020 para el municipio de Medellín
Departamento Administrativo Nacional de Estadística (DANE) ⁵	Población de municipios de Antioquia. Censo Nacional de Población y Vivienda 2018

¹ Fuente: Portal Geográfico del Municipio de Medellín. Recuperado de <https://www.medellin.gov.co/geomedellin/index.hyg>

² Fuente: Sistema de Alertas Tempranas del Valle de Aburrá- SIATA. Recuperado de https://siata.gov.co/descarga_siata/index.php/Index2/

³ Fuente: Sistema SIGMA. Convenio Área Metropolitana del Valle de Aburrá y la Universidad Pontificia Bolivariana. Recuperado de <http://modemed.upb.edu.co/zigma2/>

⁴ Fuente: Alcaldía de Medellín. Proyecciones de Población 2016 a 2020 de Medellín. Recuperado de <https://www.medellin.gov.co/irj/porta/medellin?NavigationTarget=navurl:/06bdb4d91e35cb9d3de717115deedc3>

⁵ Departamento Administrativo Nacional de Estadística - DANE. Resultados Censo Nacional de Población y Vivienda: tercera entrega. Recuperado de <https://www.dane.gov.co/files/censo2018/informacion-tecnica/cnpv-2018-presentacion-3ra-entrega.pdf>

Para Medellín, se encontró información de todas las variables solicitadas, sin embargo, para los otros municipios del Valle de Aburrá, no se pudo obtener información relacionada a aforos vehiculares para todo el periodo de análisis.

– Variable dependiente: PM_{2.5} (µg/m³)

Durante el periodo de análisis estuvieron en funcionamiento 24 estaciones de monitoreo de PM_{2.5} en la Red de Monitoreo de Calidad del Aire del Valle de Aburrá, donde dos de ellas eran semiautomáticas (mediciones cada 72 horas) y el resto automáticas (cada hora). Las series de concentraciones promedio diarias de este contaminante fueron obtenidas a través del Sistema Sigma, convenio establecido entre el Área Metropolitana del Valle de Aburrá y la Universidad Pontificia Bolivariana, y la ubicación geográfica de las estaciones fue obtenida por medio del SIATA.

Debido a la limitación y escasez de datos relacionados con aforos vehiculares para los municipios del Valle de Aburrá, sólo exploraron 18 estaciones de monitoreo de PM_{2.5} ubicadas en los municipios definidos en la zona de estudio (sección 5.1), las cuales se muestran en la Tabla 11. Entre estas estaciones, una era urbana industrial, tres eran urbanas de tráfico, dos eran suburbanas de fondo y el resto eran urbanas de fondo (EAFIT, 2019). Aunque lo ideal sería que la mayoría de las estaciones insumo para la construcción de los modelos LUR fueran de tráfico (Londoño & Cañón, 2015), al tener una cantidad limitada de equipos en la red de monitoreo, se optó por utilizar todas las estaciones, sin importar su clasificación.

Tabla 11. Estaciones de la Red de Monitoreo de Calidad del Aire del Valle de Aburrá exploradas

#	Estación	Nombre	Municipio	Código	X	Y	Altura (msnm)	Altura nivel de piso (m)	Tipo de estación ²
1	BEL.FEVE	I.E. Fernando Vélez	Bello	87	835055.99	1192800.9	1510	8	UF
2	COP.CVID	Ciudadela Educativa La Vida	Copacabana	82	842033.9	1193665.2	1455	11.5	UF
3	ENV.HOSP	E.S.E. Santa Gertrudis	Envigado	88	833487.44	1174121.9	1581	11	UF
4	EST.HOSP	Hospital La Estrella	La Estrella	78	826601.07	1172667.7	1765	15	UF
5	SUR.TRAF	Estación Tráfico Sur	La Estrella	48	828448.25	1172311.5	1502	3	UT
6	ITA.CJUS	Casa de Justicia de Itagüí	Itagüí	28	831800.42	1176001	1536	3	UI
7	ITA.CONC	I.E. Concejo Municipal de Itagüí	Itagüí	38	826581.01	1174102	1645	4	SF
8	MED.ALTA	I.E. Pedro Octavio Amado - Altavista	Medellín	79	830318.47	1180007.5	1595	8	UF
9	MED.AРАН	I.E. Ciro Mencia	Medellín	86	836314.65	1187923.9	1545	6	UF
10	MED.BEME	I.E. Pedro Justo Berrio	Medellín	83	830333.01	1181704.3	1614	5	UF
11	MED.LAYE	Tanques La Ye EPM	Medellín	44	836955.82	1175654.7	1981	4	SF
12	CEN.TRAF	Estación Tráfico Centro	Medellín	12	834859.18	1183399.8	1730	4	UT
13	MED.PJIC ¹	Politécnico Colombiano Jaime Isaza Cadavid	Medellín	6	833952.93	1178578.7	1546	6	UT
14	MED.SCRI	San Cristóbal - Parque Biblioteca Fernando Botero	Medellín	85	827460.08	1186197.8	1683	12	UF
15	MED.TESO	I.E. Inem - Sede Santa Catalina	Medellín	84	835813.36	1177561.6	1670	12	UF
16	MED.UNNV	Universidad Nacional - Sede El Volador	Medellín	25	834037.91	1184595.6	1469	4	UF
17	MED.VILL	Villahermosa - Planta de producción de agua potable EPM	Medellín	80	837219.03	1184102.3	1714	4	UF
18	SAB.RAME	I.E. Rafael J. Mejía	Sabaneta	90	829246.33	1171318.2	1625	8	UF

Fuente: Sistema SIGMA. Convenio Área Metropolitana del Valle de Aburrá y la Universidad Pontificia Bolivariana. <http://modemed.upb.edu.co/zigma2/>. Consultado: 29 de enero de 2018.

Fuente: Sistema de Alertas Tempranas del Valle de Aburrá- SIATA. https://siata.gov.co/descarga_siata/index.php/Index2/. Consultado: 31 de enero de 2018.

¹ Estación semiautomática con registros diarios para la mayoría del periodo evaluado.

² UT: Urbana de tráfico, UI: Urbana industrial, UF: Urbana de fondo, SF: Suburbana de fondo. Fuente: (EAFIT, 2019)

– Variables ambientales

Las variables meteorológicas evaluadas fueron temperatura (°C), velocidad del viento (m/s) y precipitación (mm), cuya información fue adquirida a través del SIATA. Durante el periodo de análisis, estuvieron en funcionamiento 32 estaciones meteorológicas con mediciones cada minuto

de las dos primeras variables, sin embargo, algunas de las estaciones se encontraban por fuera de los municipios del Valle de Aburrá, otras no tenían información disponible en el periodo de análisis, por tanto, en una exploración preliminar no se incluyeron estas estaciones, además de las ubicadas en los municipios de Barbosa, Caldas y Girardota; en total se seleccionaron 22 estaciones para la temperatura y la velocidad del viento.

Adicionalmente, estuvieron en funcionamiento 148 estaciones pluviográficas con registros minuto a minuto de precipitación (mm) y, después de aplicar los mismos criterios previamente mencionados, se seleccionaron 29 estaciones (Tabla 12). Los gradientes de temperatura, velocidad del viento y precipitación fueron estimados en la etapa de preprocesamiento de las variables ambientales (Momento 2), descrita más adelante.

Tabla 12. Estaciones con registros de temperatura (T), velocidad del viento (Vv) y/o precipitación (P), para el año 2018

#	Nombre	Municipio	Código	X	Y	T (°C)	Vv (m/s)	P (mm)
1	I.E Barrio Paris	Bello	121	833559.704	1190174.464			x
2	Centro de Desarrollo Infantil Pequeños Exploradores	Bello	154	837281.8779	1189709.17			x
3	Jorge Eliecer Gaitán - Thies	Bello	271	836101.7676	1192751.379	x	x	x
4	La isla	Bello	282	834491.9012	1189863.134			x
5	Ciudadela Educativa La Vida	Copacabana	73	842070.9742	1193647.287	x	x	x
6	Bomberos Envigado	Envigado	244	831374.1896	1173905.406			x
7	Alcaldía Envigado	Envigado	252	832768.0448	1174263.187	x	x	x
8	I.E Concejo Municipal de Itagüí (ITA.CONC) ¹	Itagüí	38	826560.7253	1174126.229	x	x	
9	Colegio Concejo de Itagüí	Itagüí	206	826589.3825	1174081.894	x	x	x
10	Casa de Justicia de Itagüí (ITA.CJUS) ¹	Itagüí	28	831800.4231	1176000.997	x	x	
11	Alcaldía La Estrella - Thies	La Estrella	229	826469.6459	1173274.656	x	x	
12	Alcaldía La Estrella	La Estrella	230	826469.6459	1173274.656			x
13	Deslizamiento La raya	La Estrella	261	826128.8823	1168076.104			x
14	Himalaya	La Estrella	266	827078.6314	1171763.899			x
15	Colegio Canadiense	La Estrella	267	827213.1467	1170048.767			x
16	Estación Tráfico Centro (CEN.TRAF) ¹	Medellín	12	834866.4248	1183401.541	x	x	
17	Universidad Nacional de Colombia - Sede El Volador (MED.UNNV) ¹	Medellín	25	834037.9054	1184595.574	x	x	
18	Tanques La Ye - EPM (MED.LAYE) ¹	Medellín	313	836941.4733	1175649.791	x	x	
19	ISAGEN	Medellín	59	836028.9374	1179125.933	x	x	
20	Jardín Botánico	Medellín	68	835633.2975	1185262.294	x	x	x
21	Torre SIATA	Medellín	77	832758.6187	1184144.329			x
22	Centro de salud San Javier la Loma	Medellín	83	828051.4456	1185571.89	x	x	x
23	Colegio Divino Salvador	Medellín	129	837207.3338	1181682.486			x
24	Miraflores	Medellín	184	837117.7015	1180901.73			x

#	Nombre	Municipio	Código	X	Y	T (°C)	Vv (m/s)	P (mm)
25	Pluviográfica La Sanín	Medellín	189	837450.3663	1177911.746			x
26	Universidad de Medellín (MED.UDEM) ¹	Medellín	197	830389.1743	1180938.577	x	x	x
27	Politécnico Jaime Isaza Cadavid (MED.PJIC) ¹	Medellín	198	833946.0946	1178581.992	x	x	x
28	Torre SIATA	Medellín	201	832758.6187	1184144.329	x	x	x
29	AMVA	Medellín	202	834340.6524	1182251.424	x	x	x
30	UNAL-Sede Agronomía	Medellín	203	833959.0669	1184310.146	x	x	
31	Santa Elena-Radar	Medellín	205	839495.5805	1176855.199	x	x	x
32	La Ladera	Medellín	211	836587.3565	1183245.129			x
33	Sede Social Picacho Las Vegas	Medellín	241	832995.7058	1188437.05			x
34	IE Jesús María Valle	Medellín	242	833489.1103	1188725.466			x
35	Escuela CEDEPRO	Medellín	249	827476.3641	1179775.55	x	x	x
36	Estación Tráfico Sur (SUR.TRAF) ¹	Sabaneta	48	828423.3629	1172329.72	x	x	
37	Club del Adulto Mayor	Sabaneta	146	829001.9499	1171774.907			x
38	Fuente Clara	Sabaneta	193	830039.207	1170151.236			x
39	Institución Rafael J. Mejía - Thies	Sabaneta	318	829096.6387	1171589.886	x	x	

Fuente: Sistema de Alertas Tempranas del Valle de Aburrá- SIATA. https://siata.gov.co/descarga_siata/index.php/Index2/. Consultado: 20 de febrero de 2018.

¹ Estaciones de la Red de Calidad del Aire del Valle de Aburrá con mediciones de variables meteorológicas.

Aunque en el análisis del comportamiento de los contaminantes atmosféricos existen otras variables meteorológicas que influyen el transporte y dispersión de los compuestos, la complejidad de los patrones de flujo del viento en áreas urbanas, generan limitaciones en las predicciones de la dinámica del contaminante. Debido a esto, en la construcción de los modelos LUR, algunos autores han explorado la inclusión de componentes relacionadas a la dinámica atmosférica en la metodología. En el estudio realizado por Arain et al. (2007), se incluyeron campos de flujo de viento interpolados en la construcción del modelo LUR para NO₂ en dos ciudades canadienses, encontrando incrementos entre 1 y 4% en el rendimiento del mismo, sin embargo, señalan que este es un procedimiento largo y complejo (Arain et al., 2007). La inclusión de la dirección del viento como variable explicativa, aunque se ha sugerido que puede mejorar el desempeño del modelo, no es común en la construcción de modelos LUR, por esta razón, en este trabajo se optó por explorar las variables más comunes utilizadas en la metodología e incluir ésta, y otras variables, en un estudio futuro.

– Variables de usos del suelo

A partir de los planes de ordenamiento territorial y los mapas de usos generales del suelo urbanos y rurales de los municipios seleccionados, cuya información fue adquirida a través del Área

Metropolitana del Valle de Aburrá y el Portal Geográfico del Municipio de Medellín, se obtuvieron las capas de uso industrial, residencial y zonas verdes, observados en la Figura 4.

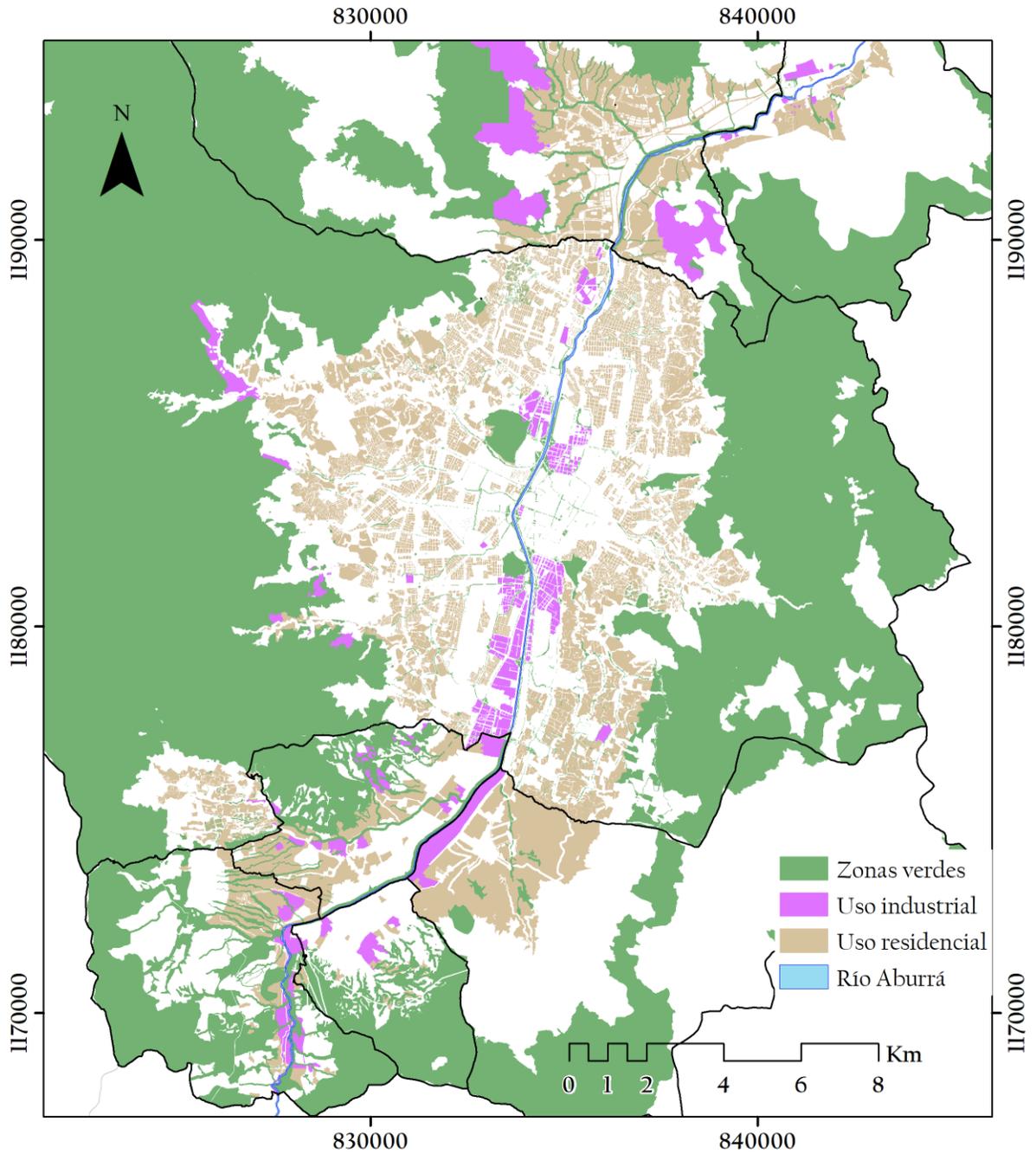


Figura 4. Usos industriales, residenciales y zonas verdes en la zona de estudio

Fuente: Portal Geográfico del Municipio de Medellín. Recuperado de <https://www.medellin.gov.co/geomedellin/index.hyg>

– Variables de movilidad

Con el fin de analizar la influencia de la cercanía al tráfico vehicular en las concentraciones de $PM_{2.5}$, se utilizó como insumo la malla vial de los municipios seleccionados. A partir de la jerarquización vial expuesta en los POT de los municipios, se seleccionaron como vías principales las vías arteria y travesía, y como secundarias, las colectoras, mostradas en la Figura 5.

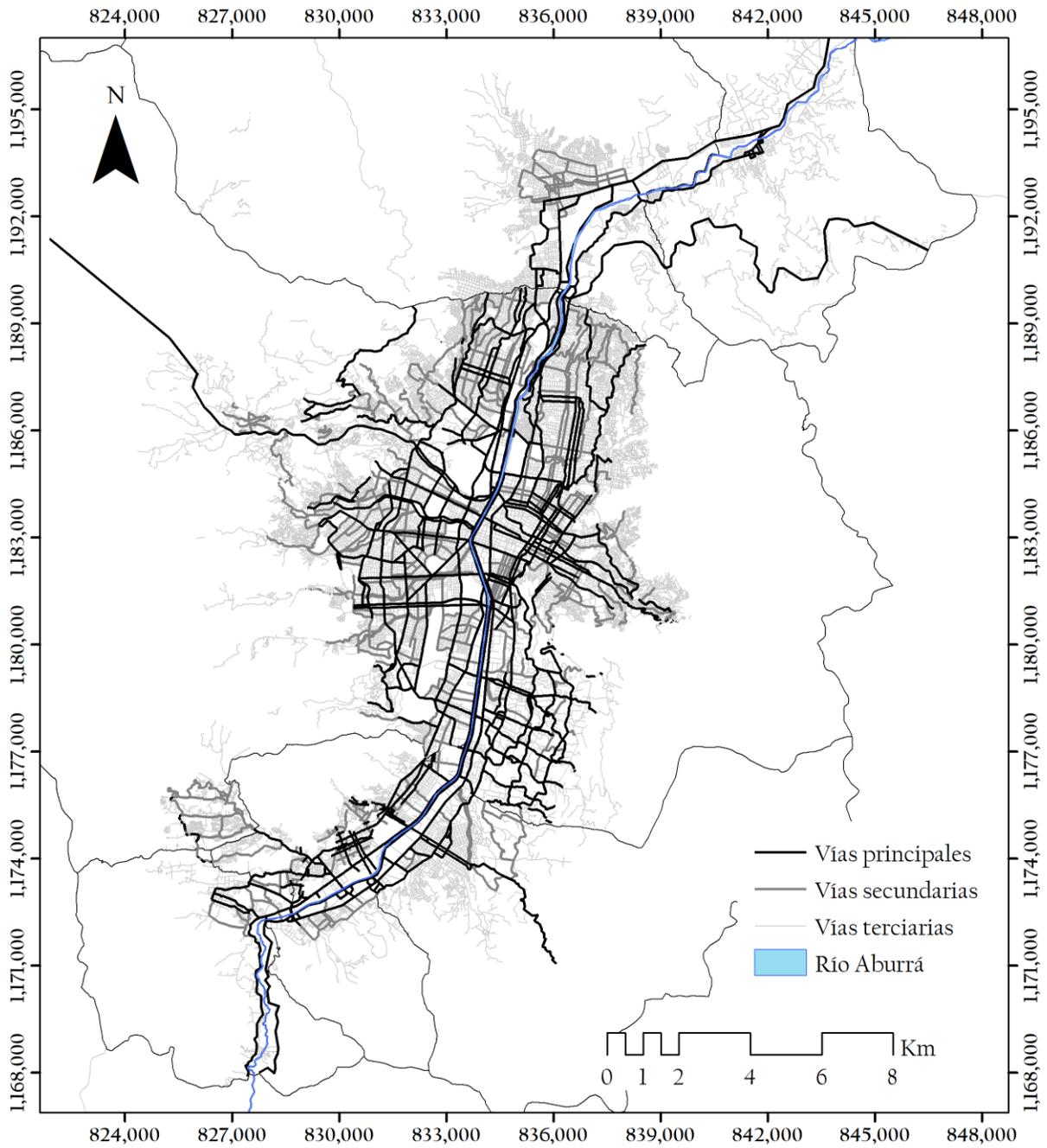


Figura 5. Malla vial de la zona de estudio, insumo para la construcción de los modelos LUR

Por otro lado, para analizar la influencia del flujo vehicular en las vías principales y secundarias cercanas a las estaciones de monitoreo, se obtuvo información de aforos vehiculares acumulados diarios en el año 2018, a partir de los dispositivos del Sistema Inteligente de Movilidad de Medellín

(SIMM) de la Secretaría de Movilidad de Medellín, además de información sobre el parque automotor matriculado para el mismo año. En total, se adquirió información de intensidad dada en vehículos mixtos por día, para 40 puntos distribuidos en la zona urbana del municipio de Medellín (Tabla 13).

Tabla 13. Cámaras de fotodetección insumo para estimación del flujo vehicular

#	Dirección	Nombre	Código	Longitud	Latitud
1	Cr 49 Cl 18	Servientrega	1	834077.615	1179798.262
2	Cr 80 Cl 55	Mediterráneo	2	832012.5372	1185263.626
3	Cr 65 Cl 2A Sur	Campos de Paz	3	832684.9309	1178407.234
4	Cl 33 Cr 76	Los Almendros (Occ)	5	831831.2227	1181891.385
5	Cr 43A Cl 31	Carnavale (N)	9	834815.2549	1181052.197
6	Cr 52 Cl 2 Sur	Cristo Rey (S)	12	833018.8839	1178342.186
7	Cr 52 Cl 2 Sur	Cristo Rey (N-S)	13	833031.4002	1178322.32
8	Cr 65 Cl 80	Cementerio Universal	14	834426.2141	1186031.664
9	Cl 29 Cr 44	Premium Plaza (Ori)	15	834660.8342	1180801.815
10	Cl 44 Cr 70	Éxito 70 (Occ)	18	832724.8089	1183028.287
11	Cr 52 Cl 12A Sur	Guayabal	19	832615.7419	1177314.243
12	Cr 34 Cl 16 A Sur	Mall Campestre	24	834812.6216	1176004.21
13	Cl 44 Cr 53	Alpujarra (Ori)	27	834472.6968	1182649.377
14	Cr 70 Cl 47D	Estación Estadio	31	832824.1876	1183457.732
15	Cr 81 Cr 76	La Mota	33	831571.5824	1179725.432
16	Cr 45 Cl 46	El Palo - El Huevo	35	835194.6543	1182526.839
17	Cl 57B Cr 57	Mínorista	38	834574.7225	1184000.463
18	Cl 10 Cr 43F	Éxito Poblado	40	834291.2706	1178965.338
19	Cr 38 Cl 19	Las Palmas	48	835200.0898	1179820.37
20	Cr 65 Cl 25	Barrio Antioquia	50	833303.6051	1180538.15
21	Cl 30 Cr 66B	Fátima	51	832909.8409	1181107.868
22	Cl 10A Cr 36	10 Hotel	53	835299.7135	1178658.929
23	Cl 10 Cr 40	Mondongos	54	835011.9017	1178669.991
24	Cl 50 Cr 65	Sauces (Occ-Ori)	56	833606.5525	1183792.53
25	Cr 81 Cl 35	Don Quijote	58	831176.9797	1182401.445
26	Cr 48 Cl 16A	Las Vegas	59	834165.5861	1179573.71
27	Cr 65 Cl 95	Parque Juanes	61	834766.2336	1187455.381
28	Cr 50 Cl 4 Sur	Puente 4 Sur	63	833714.9181	1177943.57
29	Cl 18 Cr 23	Chuscalito	65	836737.8715	1179343.199
30	Cr 49 Cl 17A Sur	Tugó II	66	833413.5875	1176489.28
31	Cr 52 Cl 21	Zoológico Santafé	67	833739.7019	1180072.371
32	Cl 66 Cr 45	Palos Verdes	69	836309.5098	1184386.949
33	Cr 80 Cl 76	Doña María	70	832658.7243	1186142.444
34	Cr 52 Cl 17	Bodegas Olaya	71	833680.9159	1179818.183

#	Dirección	Nombre	Código	Longitud	Latitud
35	Cr 25 Cl 9A Sur	Transversal Superior	72	835834.6137	1176449.316
36	Cr 70 Cl 19	Aeroparque Juan Pablo II	73	832378.2753	1180101.074
37	Cl 54 Cr 67B	La Iguaná	74	833397.2434	1184195.835
38	Conexión vial Aburrá - Cauca	Vía Túnel de Occidente	75	826542.6449	1186316.718
39	Conexión vial Aburrá - Cauca	Vía Túnel de Occidente (Occ-Ori)	76	826542.6449	1186316.718
40	Cl 12 Sur Cr 50FF	FLA (N-S)	77	833123.8135	1177205.133

– Población

A partir de las proyecciones de población y el Censo Nacional de Población y Vivienda de 2018 en la zona de estudio, se pudo obtener información de población por barrio, para el año 2018, en Medellín (Alcaldía de Medellín, 2015); por comuna, para el año 2017, en Itagüí (Departamento Administrativo de Planeación, 2018) y por municipio para los otros (Departamento Nacional de Estadística, 2019). En la Figura 6 se muestra el número de habitantes asignado a los centroides de cada entidad (barrio, comuna o municipio).

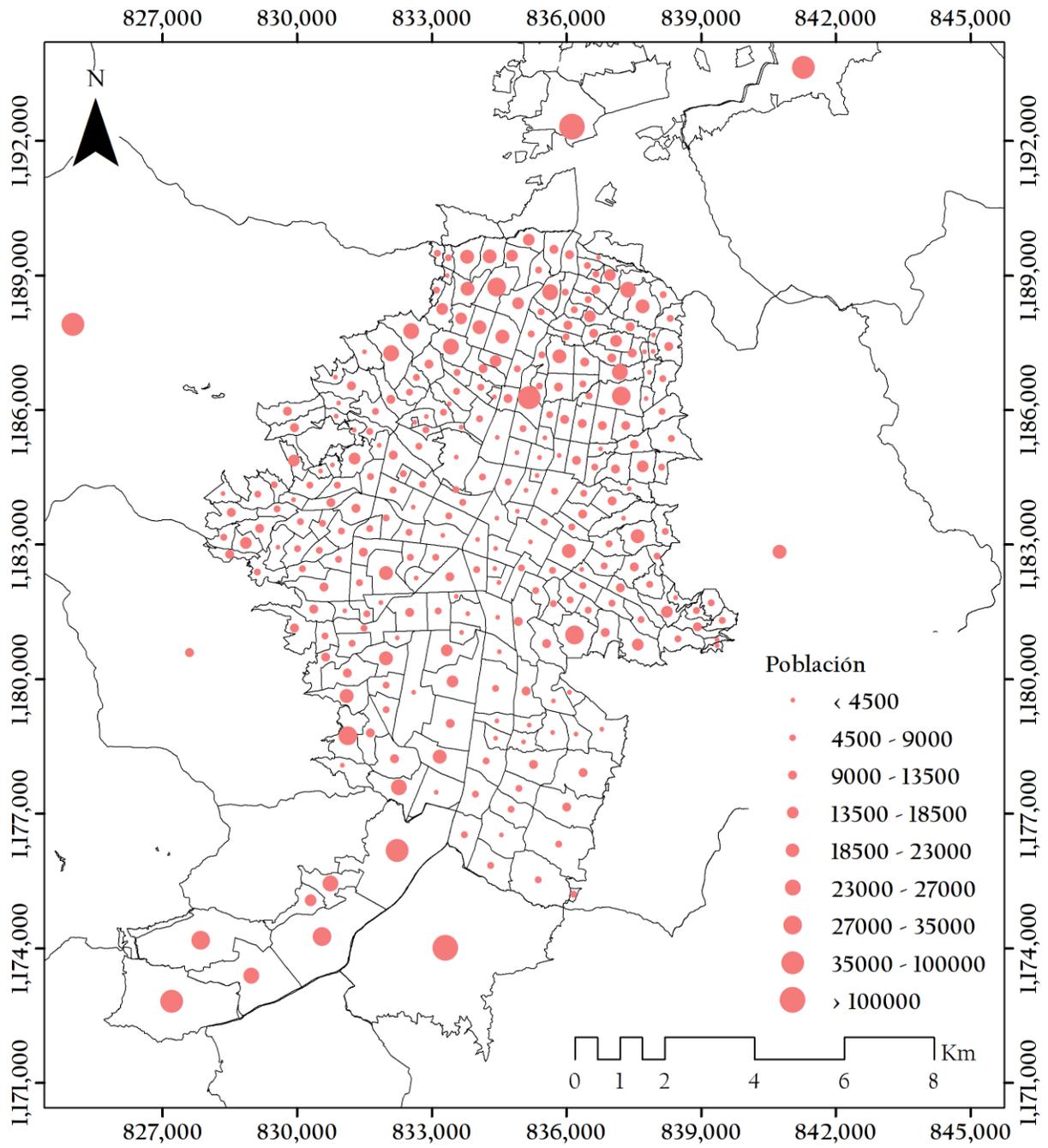


Figura 6. Número de habitantes en la zona de estudio

5.3.2. Análisis exploratorio de variables ambientales y aforos vehiculares

Con el fin de analizar la calidad de la información adquirida y de tener una caracterización de los datos, se realizó una exploración inicial de las series de tiempo de $PM_{2.5}$, temperatura, velocidad del viento, precipitación y aforos vehiculares; permitiendo evaluar la plausibilidad de los datos, además del estado de las series. Se calcularon los estadísticos descriptivos mensuales y para el año completo, y se realizaron gráficos para analizar la dispersión, asimetría y los valores atípicos en todo el periodo de análisis.

Por otro lado, con el propósito de comparar las estimaciones de los modelos LUR mensuales entre sí, se debían mantener condiciones homogéneas en la información de entrada para la construcción de cada modelo, es decir que los datos para cada mes debían provenir de exactamente las mismas estaciones a lo largo del periodo de análisis, tanto para la variable dependiente, como para las variables explicativas (temperatura, velocidad del viento, precipitación y aforos vehiculares). Si para un mes en específico, no se presentaba información en una estación, ésta no podía incluirse en la construcción de los otros modelos; por esta razón, la cantidad de estaciones inicial, indicadas en la Tabla 11, Tabla 12 y Tabla 13, estaba sujeta a reducirse como consecuencia de la calidad y cantidad de la información disponible.

Para analizar la cobertura de los datos analizados, se calculó el porcentaje de valores no disponibles en cada una de las series para todo el periodo de análisis y para cada mes; además, con el fin de evaluar la distribución de los mismos en la serie, se graficó su distribución en todo el periodo de análisis. Para esto, todos los valores de la serie fueron reemplazados por un número único, ocasionando que la serie completa se transformara en una línea recta para visualizar con claridad los periodos sin registros.

5.3.3. Criterios de selección de estaciones

Además de los criterios de inclusión relacionados a la ubicación de las estaciones, mencionados previamente; con el fin de tener homogeneidad en la cantidad de datos a analizar mensualmente y para obtener modelos LUR mensuales comparables, se establecieron criterios para la selección de las estaciones de monitoreo de $PM_{2.5}$, las estaciones meteorológicas y las cámaras de fotodetección, expuestos a continuación:

- Porcentaje de valores perdidos no excedió el 25% de los datos disponibles en las series diarias, para todo el periodo de análisis.
- El gráfico de registros no mostró meses completos sin información disponible.
- El porcentaje de valores perdidos mensual no excedió el 27% de los datos en las series diarias.

Después de explorados estos criterios, sólo se incluyeron en la construcción de los modelos, las estaciones que cumplieron con los tres criterios definidos, es decir que tuvieran mediciones en todo el periodo de medición.

5.3.4. Excedencias a la norma de calidad del aire

Como un ejercicio exploratorio, a las estaciones de monitoreo de PM_{2.5} seleccionadas, se calculó el número de excedencias a la norma, es decir el número de días en donde la concentración del contaminante excedió el nivel diario máximo permitido establecido en la Resolución 2254 de 2017 (Ministerio de Ambiente y Desarrollo Sostenible. República de Colombia, 2017).

Para el cálculo de las excedencias, se categorizaron las series asignándoles un valor nulo o cero cuando las concentraciones no excedieran el umbral diario (Tabla 14) y el valor de uno para los niveles por encima de éste; de esta manera, el valor asignado corresponde a un día de excedencia. Finalmente, se graficó el total de días de excedencias en todo el periodo de análisis, para cada serie.

Tabla 14. Niveles diarios máximos permisibles para PM_{2.5}

Fechas de vigencia	Nivel máximo permisible-PM _{2.5}	Tiempo de exposición
1 de enero de 2018 a 30 de junio de 2018	50 µg/m ³	24 horas
Desde 1 de julio de 2018	37 µg/m ³	24 horas

Fuente: Ministerio de Ambiente y Desarrollo Sostenible (2017). Resolución 2254 de 2017: Por la cual se adopta la norma de calidad del aire ambiente y se dictan otras disposiciones. Bogotá D.C.

5.4. Momento metodológico dos: Preprocesamiento de las variables

Después de definir cuales estaciones de las diferentes variables se utilizarían para el análisis, se procedió a llevar los datos promedio mensuales a un formato espacial (capas vectoriales o *shapes*). Para la precipitación, se utilizó el valor acumulado mensual.

5.4.1. Interpolaciones espaciales

– Temperatura, velocidad del viento y precipitación

Para la construcción de los modelos LUR, se requiere que los valores de las variables del Grupo 2 se encuentren en la misma ubicación de las estaciones de monitoreo de PM_{2.5}, sin embargo, esto sólo ocurre con algunas estaciones de la Red Meteorológica y Pluviográfica del Valle de Aburrá. Para resolver esta situación, se aplicaron algoritmos de interpolación espacial, Inverso de la Distancia (IDW) o Spline, escogiendo la interpolación con menor Error Cuadrático Medio (RMSE), para la temperatura, velocidad del viento y precipitación, en cada mes. Para esto, se utilizó la función *Geostatistical Wizard* del aplicativo *Geostatistical Analyst* del software *ArcGIS 10.7*[®].

Luego, con la superficie interpolada, se estimaron los valores de las tres variables en las coordenadas de las estaciones de monitoreo de PM_{2.5} seleccionadas, por medio de la función *Extract Values to Points* del aplicativo *Spatial Analyst Tools* en *ArcGIS 10.7*[®].

– Gradientes espaciales de temperatura, velocidad del viento y precipitación

En el estudio realizado por Londoño & Cañón (2015), se encontró que los gradientes de temperatura y velocidad del viento mejoraron la descripción de la variabilidad de la concentración del PM₁₀ en los modelos LUR desarrollados. Por ende, además de las variables meteorológicas, a partir de las superficies interpoladas, se calcularon los gradientes de temperatura, velocidad del viento y precipitación, en porcentaje, por medio de la función *Slope* del aplicativo *Spatial Analyst Tools* de *ArcGIS 10.7*[®]. Con estos se calculó un mapa de pendientes para cada variable, el cual representa los gradientes mencionados. Finalmente, se aplicó la función *Extract Values to Points*, para estimar los valores de las nuevas variables en los sitios de monitoreo de PM_{2.5}.

– Aforos vehiculares

La información de aforos vehiculares está relacionada a la ubicación de las cámaras de fotodetección, por ende, no se tiene información para toda la malla vial de la zona de estudio. Para estimar los valores de aforos en las otras vías principales, se utilizó el algoritmo Geoestadístico Univariado (Kriging), que ha mostrado un mejor ajuste que otros métodos de interpolación espacial, para estimar esta variable (Eom, Park, Heo, & Huntsinger, 2006; Musunuru, Wei, & Porter, 2017; Sarlas & Axhausen, 2016; X. Wang & Kockelman, 2009).

Para evaluar la estacionariedad de los datos espaciales de aforos vehiculares, se calcularon las medidas de tendencia central (media, mediana y moda), el sesgo y el histograma de frecuencias, por medio de la función *Histogram* de *Explore Data* en el aplicativo *Geostatistical Analyst*. En caso de no cumplir con el supuesto de estacionariedad, se realizaron las transformaciones sugeridas por los criterios de Webster y Oliver, es decir, se aplicó una transformación de raíz cuadrada, si el valor absoluto del sesgo estaba entre 0.5 y 1.0; o logarítmica, si el valor absoluto del sesgo era mayor a 1.0 (Londoño, 2018a).

Después de transformados los datos, se repetía el procedimiento anterior para evaluar la estacionariedad del nuevo conjunto; de cumplir el supuesto de normalidad, se procedía con la implementación de la metodología Kriging, en caso contrario, se aplicó como método de interpolación espacial un IDW o Spline.

Con los datos estacionarios, se procedió a realizar el análisis estructural, es decir, calcular el semivariograma empírico y a ajustarlo al semivariograma teórico, por medio de la función *Geostatistical Wizard* del aplicativo *Geostatistical Analyst*. Se compararon 9 modelos de semivariograma empíricos disponibles en *ArcGIS 10.7*[®] y se escogió el modelo con el *partial sill*, el modelo más parecido a la varianza de los datos y con el menor valor de ASE, RMSE y RMSS.

Luego, para obtener las estimaciones en toda la malla vial de la zona de estudio, se asignaron los valores de los aforos vehiculares estimados previamente, a las vías principales (Figura 5), por medio de la función *Overlay-Intersect* del aplicativo *Analysis Tools* de *ArcGIS 10.7*[®]. Para la construcción de modelos LUR y el análisis, se utilizó la inversa de la variable transformada.

5.4.2. Población

A cada centroide de los polígonos de los barrios (Medellín), comunas (Itagüí) o municipios (otros municipios de la zona de estudio), se asignó el valor del número de habitantes (Figura 6). Se obtuvo la densidad poblacional en toda la superficie de la zona de estudio, por medio de la función *Kernel density-Density* del aplicativo *Spatial Analyst Tools*, la cual distribuye la magnitud de la población asignada a cada punto en toda el área de estudio, mediante una función de kernel (ESRI, 2019a).

5.4.3. Zonas de influencia o buffers

– Estaciones de monitoreo de PM_{2.5}

Se definieron seis buffers circulares con centro en las coordenadas de los sitios de monitoreo de PM_{2.5}, con radios de 50, 100, 150, 200, 300 y 500 m, con el fin de evaluar la influencia en la variabilidad del PM_{2.5}, de las variables explicativas contenidas en cada una de estas áreas, por medio de la función *Proximity-Buffer* del aplicativo *Analysis Tools*.

– Vías

Con el fin de calcular el área de las vías, basándose en el ancho promedio de estos tipos de vía (Alcaldía de Medellín, 2014), se definieron buffers de 12.5 y 5 m de radio para las vías arteria y colectoras, respectivamente.

5.4.4. Variables predictoras SIG

Cada uno de los 6 buffers definidos se intersecaron con las variables de uso del suelo (Grupo 3), por medio de la función *Extract-Clip*, y las variables de movilidad (Grupo 4), con la función *Overlay-Intersect*, ambas del aplicativo *Analysis tools*. De esta manera, para cada buffer se obtuvo sólo la información que se encontraba disponible en el área definida por los mismos. Estos datos construidos a partir de las intersecciones de los buffers, son las variables SIG predictoras.

Para cada estación de PM_{2.5} y cada buffer, se calcularon las distancias desde la estación hasta el uso o vía más cercana y las áreas intersecadas para cada variable. Con la información adquirida se definieron las posibles variables explicativas, mostradas en la Tabla 15, clasificadas en cada uno de los grupos establecidos previamente (Tabla 8).

Tabla 15. Variables explicativas evaluadas en la construcción de los modelos LUR

#	Variables explicativas	Unidades	Nombre de la variable
Grupo I. Variables de posición			
1	Altura de la estación	msnm	ALTURA
2	Altura a nivel de piso	m	PISO

#	VARIABLES explicativas	Unidades	Nombre de la variable
3	Buffer circular de radio r	m	---
Grupo 2. Variables ambientales			
4	Precipitación acumulada mensual	mm	P
5	Temperatura promedio mensual	°C	T
6	Velocidad del viento promedio mensual	m/s	Vv
7	Gradiente de temperatura promedio mensual	%	GRADT
8	Gradiente de velocidad del viento promedio mensual	%	GRADVv
9	Gradiente de precipitación acumulada mensual	%	GRADP
Grupo 3. Variables de usos del suelo			
9	Distancia a uso industrial	m	DIST_IND
10	Distancia a zonas verdes	m	DIST_VERD
11	Distancia a uso residencial	m	DIST_RESID
12	Área de uso industrial	m ²	AREA_IND
13	Área de zonas verdes	m ²	AREA_VERD
14	Área residencial	m ²	AREA_RESID
Grupo 4. Variables de movilidad			
16	Distancia a vía principal más cercana	m	DIST_VIA
17	Área de vías principales y secundarias	m ²	AREA_VIA
18	Longitud de las vías	m	LONG_VIA
19	Flujo vehicular promedio mensual de la vía más cercana	Vehículos/mes	FLUJO_VEHI
20	Número de vías	Número de vías	NUM_VIA
Grupo 5. Variable demográfica			
21	Densidad poblacional	Número de Habitantes/m ²	DEN_POB

5.5. Momento metodológico tres: Construcción de modelos LUR mensuales

Para cada mes y cada buffer se construyó una tabla con los resultados del preprocesamiento de las variables, expuestas en la Tabla 15, relacionados a cada estación de monitoreo. Si en la construcción de las variables SIG predictoras, una en específico no fue intersectada con un buffer, entonces se le asignó un valor de cero. Si esto ocurría para una variable explicativa en el 90% de las estaciones, ésta no se incluía en la construcción del modelo LUR para este buffer, siguiendo la metodología propuesta por distintos autores (Allen et al., 2011; Londoño & Cañón, 2015; Van Nunen et al., 2017).

El método OLS no requiere que las variables explicativas sean normales. Sin embargo, se puede mejorar la especificación de un modelo al transformar las variables sesgadas. Antes de construir los modelos, se evaluó la normalidad de la variable dependiente ($PM_{2.5}$) y todas las posibles

variables explicativas, por medio del cálculo de la asimetría de los datos, y en caso de no cumplir con este supuesto, se transformaba la variable con raíz cuadrada o logaritmo, siguiendo el criterio de Webster y Oliver (Ecuación 4). Luego, se calculó el coeficiente de correlación de Spearman (r) entre los datos del $PM_{2.5}$ y cada variable explicativa, en donde, para el modelo de regresión, sólo se incluyeron las variables con un valor absoluto de r menor a 0.6 (Allen et al., 2011).

Las variables ambientales variaban mensualmente, pero permanecían fijas para cada buffer; por el contrario, las variables de usos del suelo, las de movilidad (exceptuando el flujo vehicular promedio mensual) y la demográfica, sólo variaban de buffer en buffer, más no mensualmente. La variable flujo vehicular fue la única que variaba tanto para cada mes como para cada buffer. Esta clasificación es ilustrada en la Figura 7.

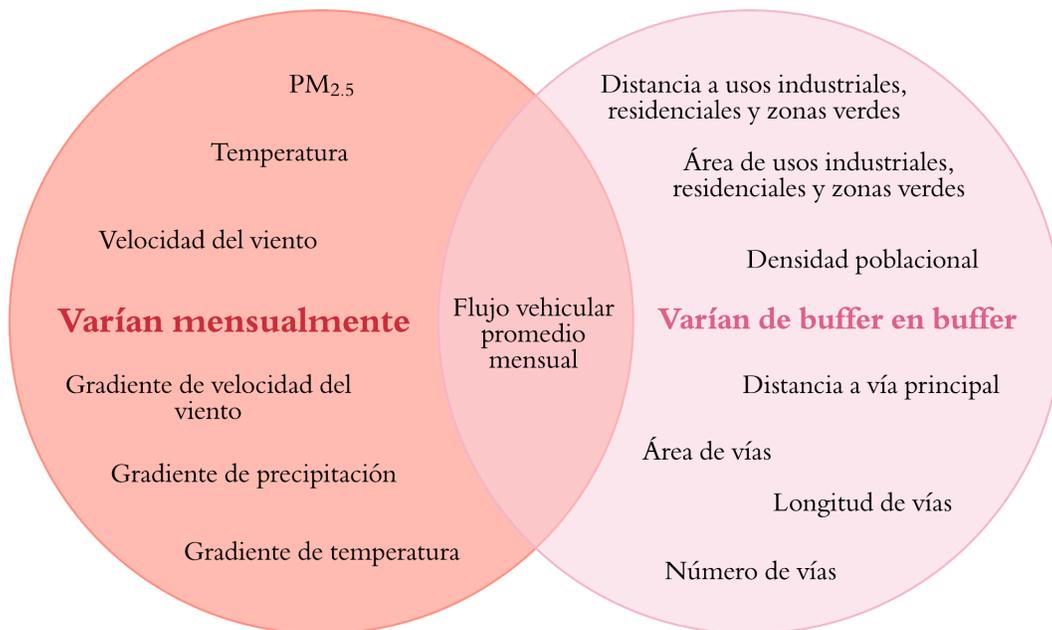


Figura 7. Variables fijas para cada mes y para cada buffer

Para cada mes se obtuvieron 6 modelos LUR diferentes, uno para cada buffer, dando en total 72 modelos para todo el periodo de análisis. Para la construcción de estos, se realizó una caracterización espacial multivariada con algoritmos de regresión espacial (Hoek et al., 2008; Londoño & Cañón, 2015; Mólter et al., 2010a; Van Nunen et al., 2017).

Primero, se especificó cada modelo, analizando cuales variables explicativas y/o sus posibles combinaciones eran significativas y cumplían con los criterios de especificación del método de Mínimos Cuadrados Ordinarios (OLS) (Londoño & Cañón, 2015), enlistados en la Tabla 16. De

esta manera, para cada mes se seleccionaron los 6 modelos, uno para cada buffer, mejor especificados, de acuerdo a los criterios del método OLS.

Tabla 16. Criterios de especificación del método OLS

Criterio de especificación	Valor
AdjR ²	> 0.5
AICc	Mínimo
VIF	< 7.5
JB	> 0.1
K(BP)	< 0.05
SA	> 0.1

Fuente: Londoño, L., & Cañón, J. (2015). Metodología para la aplicación de modelos de regresión de usos del suelo en la estimación local de la concentración mensual de PM10 en Medellín – Colombia. *Revista Politécnica*, 11(21), 29–40.

En donde:

- **Criterios de rendimiento del modelo:** Se escogió el modelo con mayor coeficiente de determinación ajustado (AdjR²) y menor criterio corregido de información de Akaike (AICc).
- **Criterio de consistencia:** Se escogió el modelo con mayor valor del Estadístico de Jarque-Bera (JB). Éste debe ser mayor a 0.1.
- **Criterio de multicolinealidad:** Se escogió el modelo con menor Factor de Inflación de la Varianza (VIF). Siendo 7.5 el valor máximo aceptado.
- **Criterio de significancia:** Se escogió el modelo con menor valor del Estadístico de residuales de Koenker Breusch-Pagan (K(BP)). El valor máximo aceptado es 0.05.
- **Criterio de ausencia de residuos en la autocorrelación espacial:** Se escogió el modelo con mayor Índice global de Moran (SA). Valor mínimo permitido es 0.1.

Para evaluar las variables explicativas y el rendimiento de los posibles modelos LUR, se utilizó la función *Exploratory Regression* en *Modeling Spatial Relationships* del aplicativo *Spatial Statistics Tools* del software *ArcGIS 10.7*®.

5.6. Momento metodológico cuatro: Selección del mejor modelo LUR mensual

Luego de obtener los 72 modelos LUR mejor especificados según el método OLS, se procedió a seleccionar el mejor modelo para cada mes, aplicando los mismos criterios (Tabla 16). Luego, se calculó el término independiente, los coeficientes de las variables explicativas seleccionadas, el valor estimado de PM_{2.5} de cada ecuación del modelo LUR, por medio de una Regresión Geográficamente Ponderada (GWR). Finalmente, se validaron los valores estimados con cada modelo aplicando el método de validación cruzada para obtener los estadísticos de error descritos en la Tabla 5: ASE, RMSE, RMSS, ARE y MARE.

Para obtener los mapas de regresiones, se creó una malla puntos regulares con celdas espaciadas por diámetro del buffer seleccionado, en donde se aplicó la fórmula obtenida por cada modelo de regresión mensual, con el fin de predecir los niveles de PM_{2.5} para cada punto. Luego, se aplicó alguno de los métodos de interpolación espacial (IDW o Spline) para obtener la superficie de concentraciones del contaminante en la zona de estudio. Para esto se utilizó la función *Geographically Weighted Regression* en *Modeling Spatial Relationships* del aplicativo *Spatial Statistics Tools*.

En la Tabla 17 y la Tabla 18 se presenta un resumen con los métodos espaciales y estadísticos utilizados en el desarrollo de este trabajo, además en la Figura 8, Figura 9, Figura 10 y Figura 11 se muestran los flujogramas metodológicos para realizar la construcción de modelos de regresión de usos del suelo implementados en este trabajo.

Tabla 17. Tabla resumen con los métodos espaciales empleados en la construcción de modelos LUR mensuales

Método	Aplicación en la metodología	Características del método	Hipótesis
Inverso de la distancia (IDW)	Imputación espacial de variables meteorológicas	Univariado, determinístico, local, inexacto, gradual convexo	Superficie subyacente es uniforme
Spline	Imputación espacial de variables meteorológicas	Univariado, estocástico, local, inexacto, gradual, convexo	Mejor predictor es el dato más cercano
Kriging	Imputación espacial de aforos vehiculares	Univariado/multivariado, estocástico, local, exacto, gradual, no convexo	Datos estacionarios
Regresión Espacial	Construcción de modelo LUR para PM _{2.5} a partir de <i>i</i> variables explicativas	Univariado/multivariado, estocástico, global, inexacto, abrupto o gradual, no convexo	Muestras independientes, normales y con varianza homogénea
Regresión Ponderada Geográficamente (GWR)	Estimar parámetros del modelo de regresión lineal múltiple	Multivariado, estocástico, local, inexacto, abrupto o gradual, no convexo	Heteroscedasticidad en el modelo de regresión y baja correlación

Tabla 18. Tabla resumen con los estadísticos empleados en la construcción de modelos LUR mensuales

Estadísticos	Aplicación en la metodología	Hipótesis	Parámetros
Coefficiente de determinación ajustado ($AdjR^2$)	Selección del modelo de regresión con mejor rendimiento	No aplica	$AdjR^2$ mayor a 0.5
Criterio corregido de información de Akaike (AICc)	Evaluar la calidad relativa del modelo estadístico	No aplica	Valor mínimo
Estadístico Jarque-Bera (JB)	Determinar si los residuos del modelo de regresión se distribuyen normalmente	Asimetría y curtosis de una distribución normal	JB mayor a 0.1
Factor de inflación de la varianza (VIF)	Evaluar independencia entre variables explicativas	Variables explicativas independientes entre sí	VIF menor a 7.5
Estadístico Breusch-Pagan estudentizado de Koenker (K(BP))	Evaluar significancia de las variables explicativas	El modelo es estacionario	K(BP) menor a 0.05
Índice Global de Moran (SA)	Evaluar si los residuos son espacialmente aleatorios	No hay autocorrelación espacial	SA mayor a 0.1

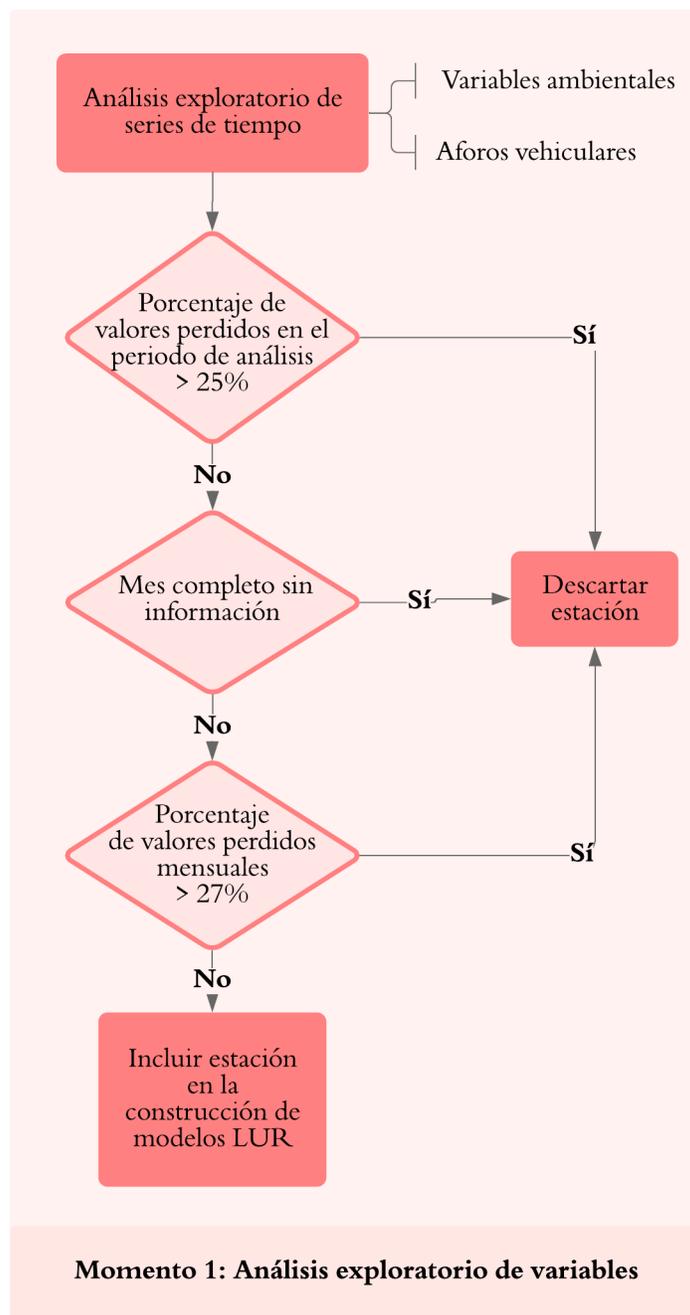


Figura 8. Flujograma metodológico para el Momento 1

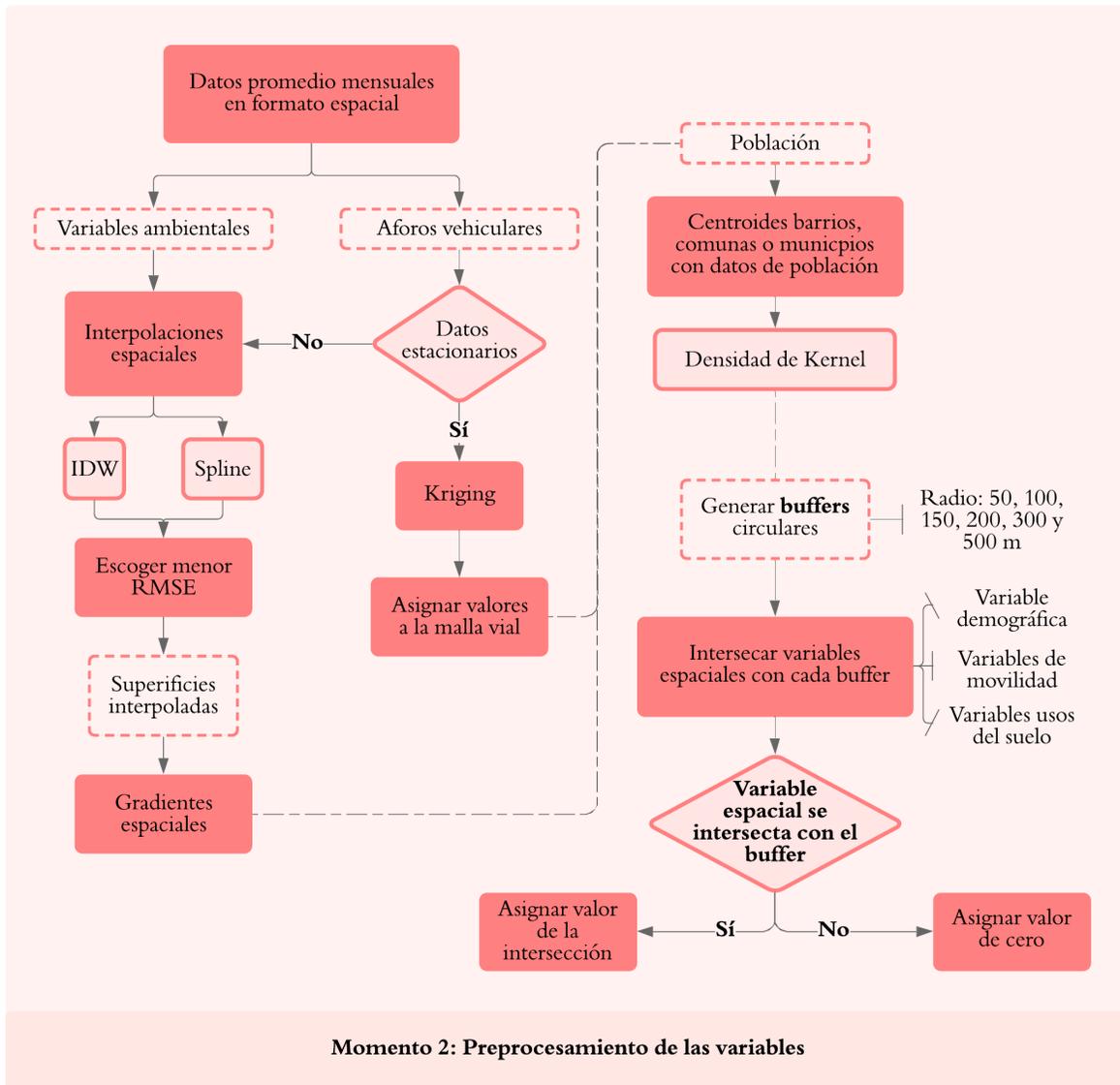


Figura 9. Flujograma metodológico para el Momento 2

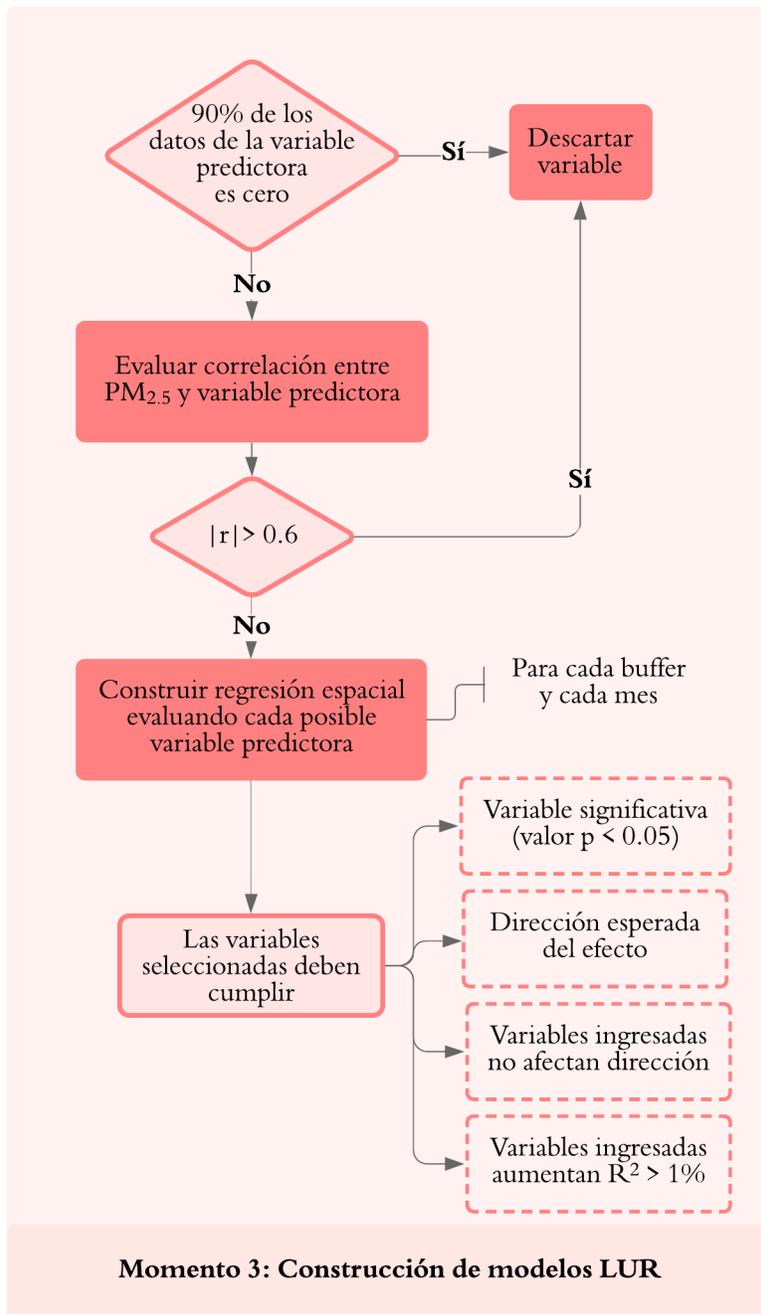


Figura 10. Flujograma metodológico para el Momento 3

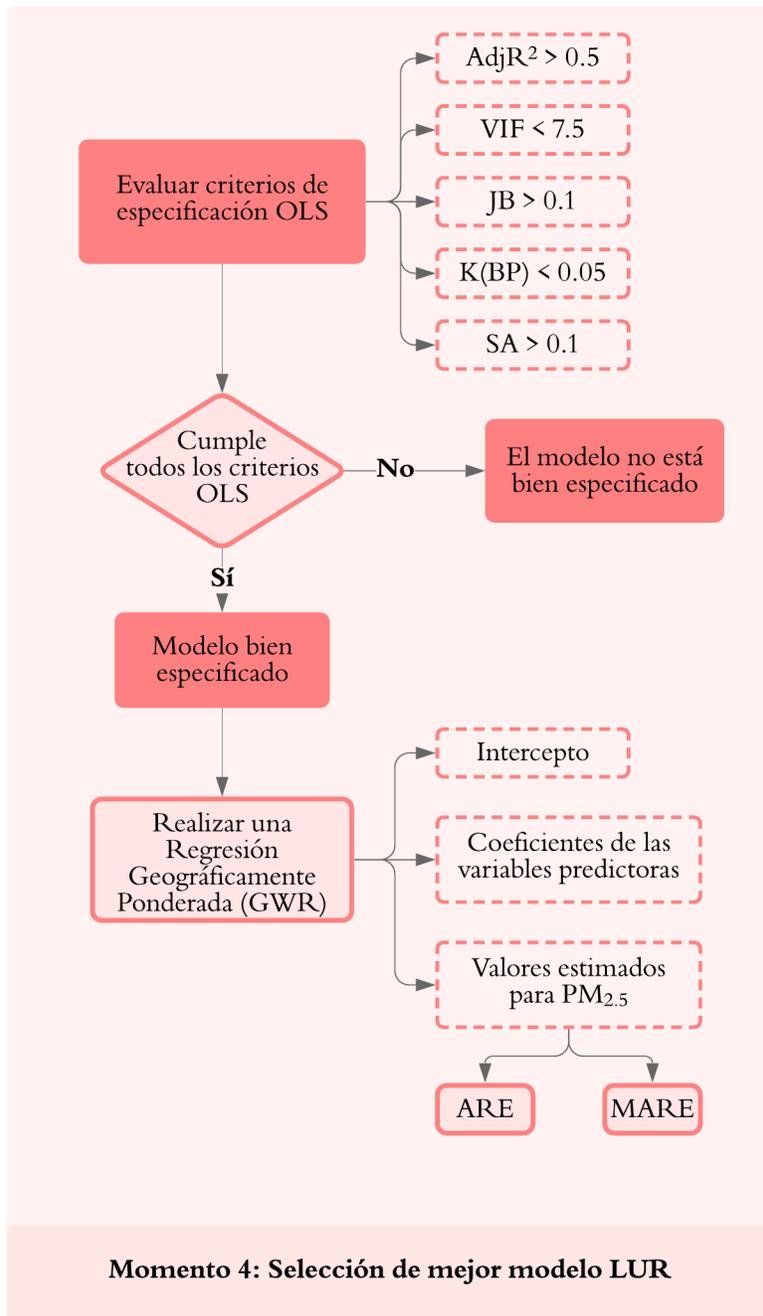


Figura 11. Flujograma metodológico para el Momento 4

6. Resultados

6.1. Momento uno: Análisis exploratorio de las variables

6.1.1. Variable dependiente: PM_{2.5} (µg/m³)

Al realizar una exploración inicial de las series de tiempo de PM_{2.5} disponibles para las 18 estaciones de la Red de Monitoreo de Calidad del Aire, se encontró que los datos presentaban valores plausibles para los niveles del contaminante, por ejemplo, no habían valores negativos ni extremadamente grandes (Tabla 19). Entre todas las estaciones se obtuvo un promedio anual de PM_{2.5} entre 15.4 y 31.9 µg/m³, con media 20.3 µg/m³; con máximos de hasta 69.7 µg/m³. Por otro lado, para el periodo activo de medición de las estaciones, únicamente la estación MED.PJIC presentó un porcentaje de valores perdidos superior al 25%, por ende, esta estación se excluyó de los posteriores análisis.

Tabla 19. Estadísticos descriptivos para las series de PM_{2.5} diarias, en todo el periodo de análisis

Estación	Media (µg/m ³)	Mediana (µg/m ³)	Desviación estándar (µg/m ³)	Máximo (µg/m ³)	Mínimo (µg/m ³)	Valores perdidos (%)
BEL.FEVE	16.8	15.5	6.2	55.8	7.1	0.5
COP.CVID	15.4	13.6	7.0	57.5	6.6	2.5
ENV.HOSP	17.8	16.5	6.8	57.6	7.3	1.4
EST.HOSP	18.7	17.5	6.4	56.5	7.1	1.4
SUR.TRAF	31.9	31.2	8.8	63.4	11.0	19.7
ITA.CJUS	23.9	22.4	7.6	56.5	10.5	3.6
ITA.CONC	21.6	19.3	9.5	64.1	5.7	3.6
MED.ALTA	19.8	18.8	6.2	44.4	8.4	1.9
MED.ALAN	22.4	20.7	7.6	61.5	10.4	1.9
MED.BEME	21.2	20.2	6.6	47.7	9.6	3.3
MED.LAYE	19.5	18.1	7.9	69.1	4.1	11.5
CEN.TRAF	26.7	24.3	10.4	69.7	10.9	24.1
MED.PJIC	20.7	19.8	6.4	38.3	8.5	38.9
MED.SCRI	15.3	13.9	5.6	40.3	6.4	2.5
MED.TESO	17.5	16.0	6.8	59.4	5.7	3.0
MED.UNNV	20.8	18.6	8.3	58.8	7.4	5.8
MED.VILL	18.7	17.5	7.6	58.5	6.8	2.5
SAB.RAME	16.2	15.7	4.9	39.2	5.8	23.0

Todas las demás series presentaron valores perdidos en el periodo de análisis, siendo para la mayoría menor al umbral establecido del 25%. Sin embargo, la distribución de estos valores también fue importante en la selección de las estaciones, debido a que, pueden estar dispersos en todo el periodo de análisis o encontrarse aglomerados en un periodo específico. En la Figura 12 se muestran los patrones de distribución de los valores perdidos en todas las series, donde se puede observar que para todas las estaciones se presentaron desde días sin registros, como en BEL.FEVE,

hasta bloques desde semanas hasta más de un mes de información no disponible, como en las estaciones MED.LAYE, CENT.TRAF, MED.UNNV, SUR.TRAF y SAB.RAME. Al evaluar el porcentaje de valores perdidos mensual para cada serie, se encontró que las estaciones anteriores, exceptuando MED.UNNV, presentaron datos faltantes continuos en un periodo igual o mayor a un mes y fueron descartadas para posteriores análisis (Anexo A).

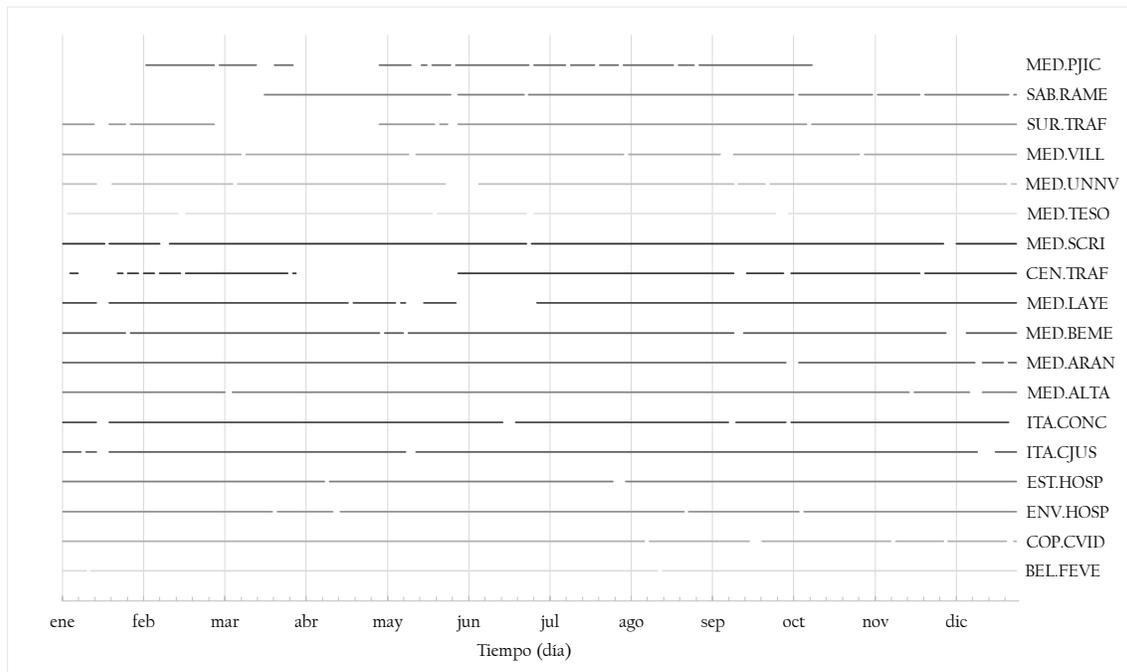


Figura 12. Gráfica de registros de PM_{2.5} en todo el periodo de análisis

Como resultado de este análisis exploratorio, para la construcción de los modelos LUR se seleccionaron 13 estaciones de monitoreo, las cuales cumplieron con los criterios de selección descritos en la sección 5.3.3. En la se Figura 13 observa, a partir de la distribución de los cuartiles, la gran variabilidad en los datos diarios de cada estación seleccionada, donde ITA.CONC presenta la menor uniformidad con una desviación estándar de 9.5 $\mu\text{g}/\text{m}^3$. Todas las series presentan una asimetría positiva o sesgo a la derecha, con la media mayor a la mediana, indicando que las concentraciones entre el 50 y 75 % de la serie, están ligeramente más dispersas que entre el 25 y 50%, es decir que los valores mayores a la mediana se encuentran más separados de la media. Además, todas las series presentan valores atípicos o *outliers*.

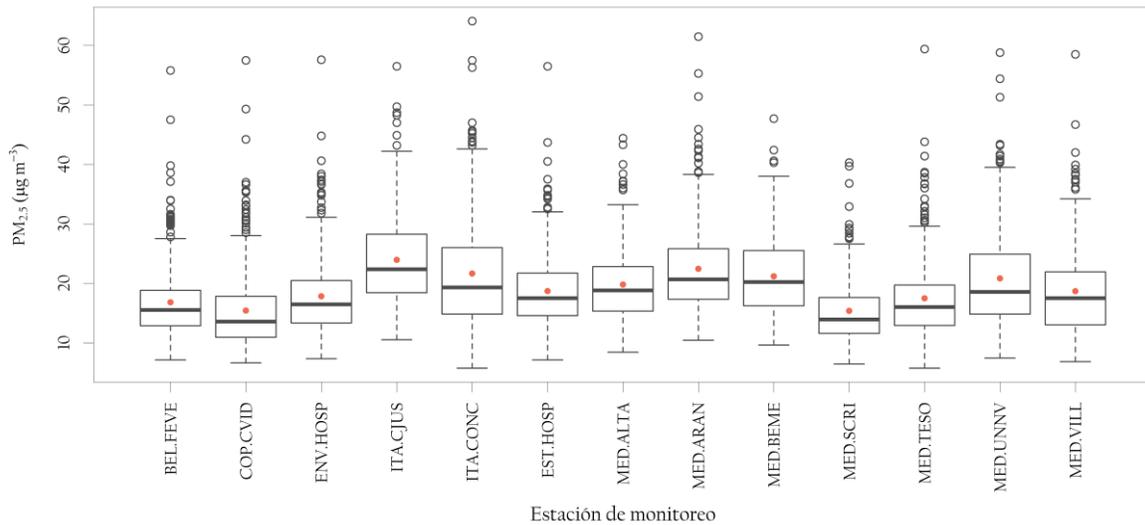


Figura 13. Distribución de los cuartiles para las series de $PM_{2.5}$ diarias seleccionadas

Durante el periodo de análisis, todas las estaciones de monitoreo, exceptuando a MED.SCRI y MED.ALTA, presentaron al menos un día con excedencia al límite máximo permitido en la norma nacional (Tabla 14). MED.AРАН presentó la mayor cantidad de excedencias (5 días), seguida por ITA.CONC y por MED.UNNV (Figura 14).

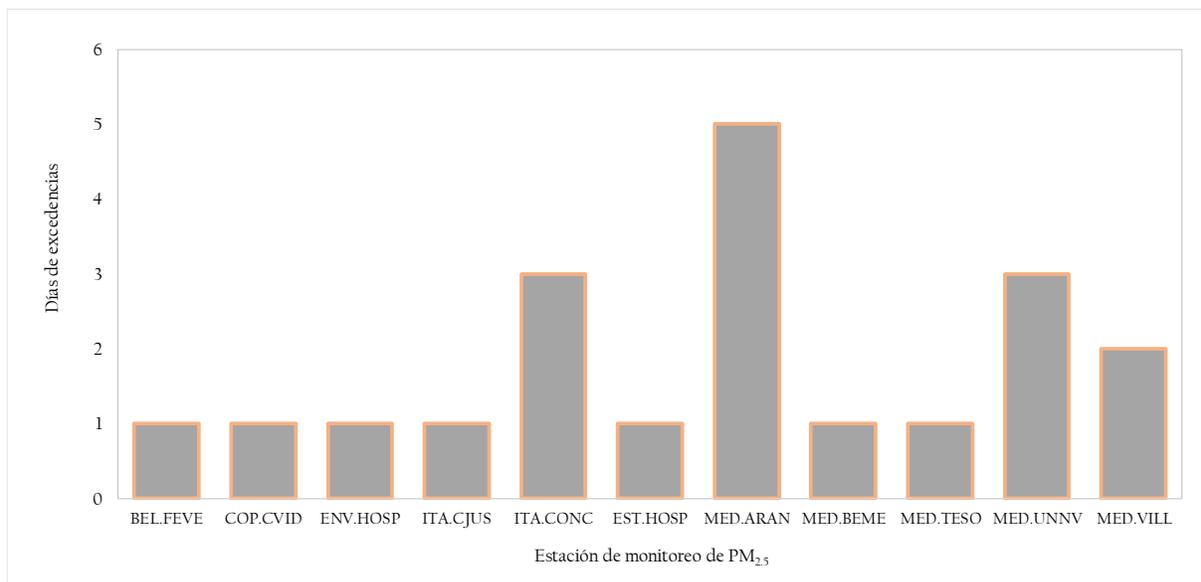


Figura 14. Número de días con excedencias a la norma nacional de los niveles de $PM_{2.5}$

Al evaluar los promedios mensuales de las series seleccionadas (Figura 15), se observa un comportamiento monomodal en los niveles de $PM_{2.5}$ en el Valle de Aburrá, con picos de concentración entre febrero y marzo, menores concentraciones entre junio y julio, y una leve tendencia al aumento en los últimos meses del año. Los niveles más altos fueron registrados en la estación ITA.CONC, seguida por ITA.CJUS, mientras que los niveles más bajos, están asociados a las estaciones MED.SCRI y COP.CVID.

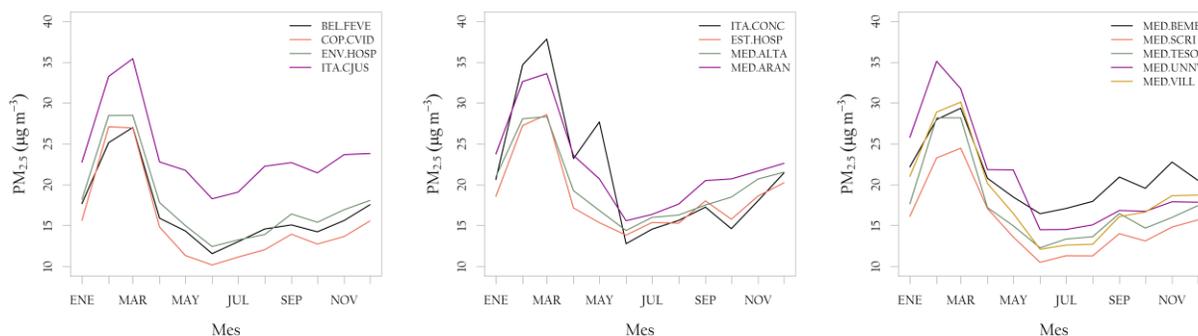


Figura 15. Promedios mensuales de $PM_{2.5}$ para las estaciones seleccionadas

En la Figura 16 se muestra la ubicación de las 13 estaciones de monitoreo seleccionadas, las cuales se encuentran distribuidas a lo largo de la zona de estudio, indicando una densidad de muestreo de 56.70 km^2 por estación de monitoreo. La mayor cantidad de estaciones se localizan en la cabecera del municipio de Medellín y una ubicada en las laderas del municipio. Se puede observar que, en promedio para el año 2018, los niveles de $PM_{2.5}$ más altos fueron registrados en estaciones ubicadas en Medellín y en el sur de la zona de estudio, mientras que en el norte se presentan concentraciones más pequeñas.

Al evaluar la normalidad de los datos espaciales mensuales, aplicando el criterio de Webster y Oliver se encontró que, para marzo, mayo, agosto y septiembre, los datos no eran estacionarios, por ende, se aplicó una transformación de raíz cuadrada para los datos de septiembre y logarítmica para los otros datos espaciales. Los valores utilizados en la construcción de modelos LUR se muestran en la Tabla 20.

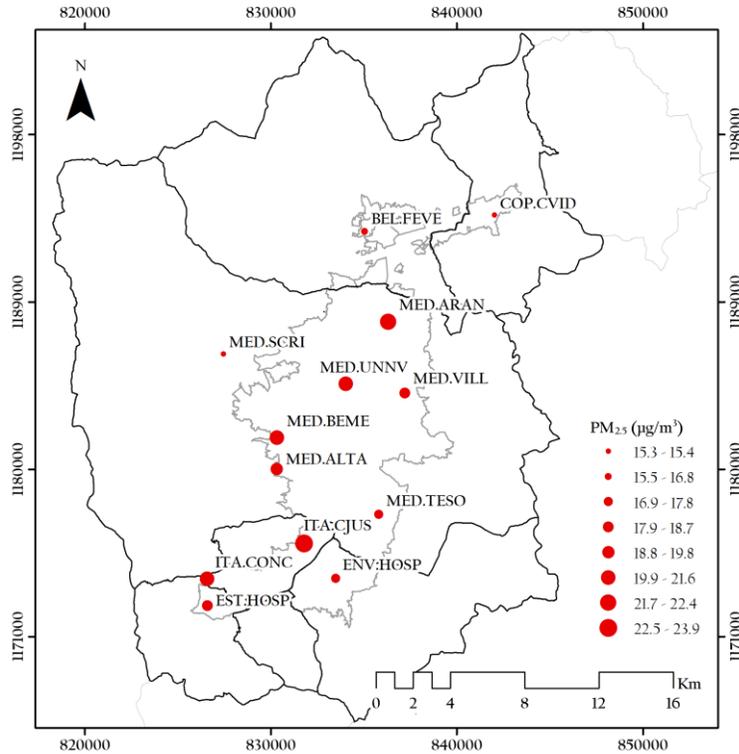


Figura 16. Ubicación estaciones de monitoreo de PM_{2.5} seleccionadas

Tabla 20. Concentraciones de PM_{2.5} promedio mensual para las estaciones de monitoreo, insumo para la construcción de modelos LUR

ID	Estación	Concentración promedio mensual de PM _{2.5} (µg/m ³)											
		1	2	3 ¹	4	5 ¹	6	7	8 ¹	9 ²	10	11	12
0	BEL.FEVE	17.7	25.2	1.4	15.9	1.2	11.6	13.0	1.2	3.9	14.2	15.6	17.6
1	COP.CVID	15.6	27.1	1.4	14.8	1.1	10.2	11.2	1.1	3.7	12.7	13.7	15.6
2	ENV.HOSP	18.2	28.5	1.5	17.8	1.2	12.4	13.3	1.1	4.1	15.4	17.0	18.1
3	ITA.CONC	20.7	34.7	1.6	23.2	1.4	12.8	14.6	1.2	4.2	14.6	18.1	21.5
4	ITA.CJUS	22.8	33.3	1.5	22.8	1.3	18.3	19.1	1.3	4.8	21.5	23.7	23.8
5	EST.HOSP	18.6	27.3	1.5	17.2	1.2	13.8	15.4	1.2	4.2	15.8	18.6	20.2
6	MED.UNNV	25.8	35.1	1.5	21.9	1.3	14.5	14.5	1.2	4.1	16.8	17.9	17.9
7	MED.ALTA	21.0	28.1	1.5	19.3	1.2	14.4	16.0	1.2	4.2	18.5	20.7	21.5
8	MED.VILL	21.0	28.9	1.5	20.2	1.2	12.1	12.6	1.1	4.0	16.7	18.7	18.8
9	MED.BEME	22.2	28.0	1.5	20.8	1.3	16.5	17.1	1.3	4.6	19.6	22.8	20.6
10	MED.TESO	17.7	28.2	1.5	17.2	1.2	12.3	13.4	1.1	4.1	14.7	16.0	17.5
11	MED.SCRI	16.2	23.3	1.4	17.1	1.1	10.5	11.3	1.1	3.7	13.1	14.8	15.7
12	MED.ALAN	23.8	32.7	1.5	23.6	1.3	15.6	16.4	1.2	4.5	20.7	21.7	22.6

¹ Logaritmo de los datos originales

² Raíz cuadrada de los datos originales

6.1.2. Variables ambientales

– Temperatura (°C)

Al realizar una exploración inicial a las series de temperatura disponibles para las 22 estaciones de la Red Meteorológica, se encontró que los datos presentaban valores plausibles de la variable (Tabla 21). Entre todas las estaciones se obtuvo un promedio anual entre 12.1 y 22.2 °C, con media 20.4 °C y máximos de hasta 28 °C; los valores más bajos para esta variable fueron registrados en la estación con código 205 (Tabla 12), ubicada en Santa Elena. Además, durante su periodo activo, las estaciones 12, 48, 313 y 318 presentaron un porcentaje de valores perdidos superior al 25%, siendo removidas para posteriores análisis.

Tabla 21. Estadísticos descriptivos para las series de temperatura promedio diaria, en todo el periodo de análisis

Código	Media (°C)	Mediana (°C)	Desviación estándar (°C)	Máximo (°C)	Mínimo (°C)	Valores perdidos (%)
12	21.6	21.4	1.3	25.0	18.5	70.7
25	22.0	22.1	1.4	25.4	18.0	13.7
28	21.7	21.6	1.3	24.4	18.5	1.4
38	19.8	19.8	1.2	23.0	16.4	3.8
48	20.6	20.4	1.2	23.1	17.5	52.9
59	20.0	20.1	1.6	24.5	15.8	24.9
68	21.8	21.8	1.2	24.7	18.8	0.5
73	21.2	21.2	1.1	23.8	18.3	0.0
83	19.0	19.1	1.1	21.5	15.9	0.8
197	21.5	21.5	1.4	28.0	17.6	5.2
198	22.2	22.2	1.4	26.4	17.7	1.4
201	21.5	21.5	1.1	25.7	18.4	3.0
202	22.0	22.0	1.2	24.9	19.0	1.1
203	21.8	21.9	1.2	24.6	18.5	5.5
205	12.1	12.2	0.7	13.7	9.9	0.3
206	19.2	19.3	1.2	21.6	16.0	4.9
229	19.6	19.5	1.2	22.1	16.4	0.0
249	19.2	19.2	1.1	21.2	16.1	0.0
252	21.2	21.1	1.3	24.9	17.3	3.0
271	22.0	22.0	1.2	25.0	19.0	1.1
313	18.1	18.2	1.1	20.8	14.7	43.0
318	20.5	20.6	1.2	23.3	16.2	50.4

Todas las series presentaron valores perdidos en todo el periodo de análisis, exceptuando la estación 73, siendo para la mayoría menor al umbral establecido del 25%. En la Figura 17 se puede observar que la mayoría de las estaciones presentaron días sin registros, donde la estación 25

presenta el mayor bloque de información no disponible, entre septiembre y noviembre; seguida por la estación 59, que presenta una cantidad importante de valores perdidos más dispersos en el primer semestre del año. Al evaluar el porcentaje de valores perdidos mensual para cada serie, se encontró que las estaciones 25, 38, 59, 203 y 206, presentaron datos faltantes continuos en un periodo igual o mayor a un mes y fueron descartadas de posteriores análisis (Anexo B).

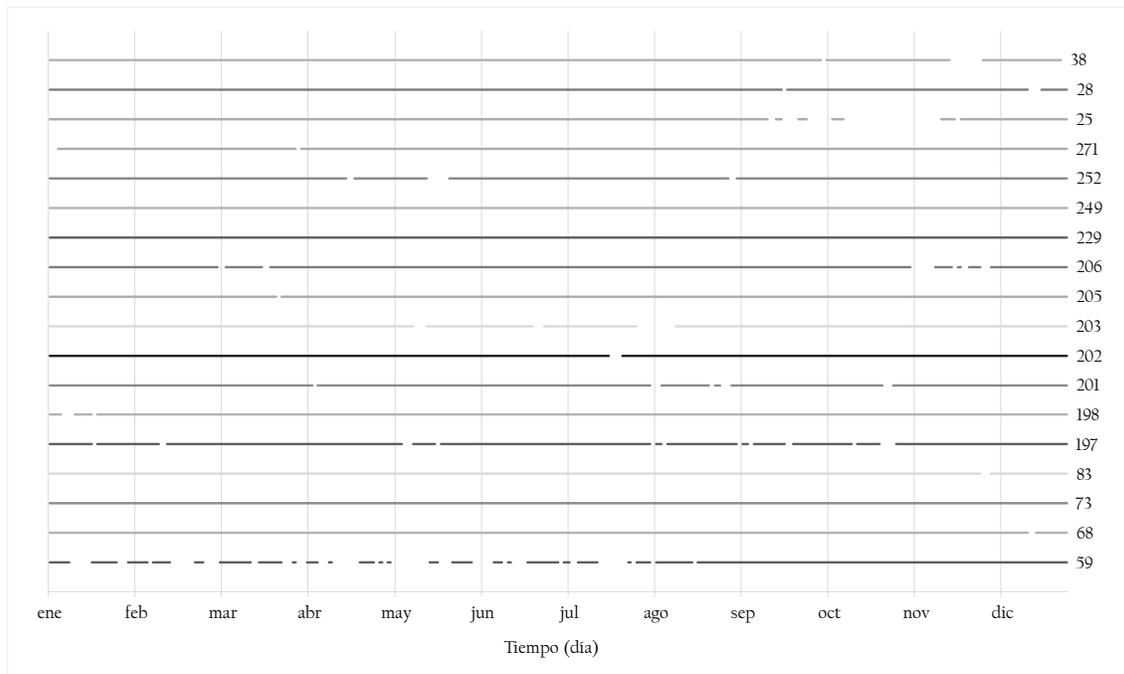


Figura 17. Gráfica de registros de temperatura promedio diaria en todo el periodo de análisis

Como resultado de este análisis exploratorio, se seleccionaron 13 estaciones de monitoreo, aplicando los criterios de selección descritos en la sección 5.3.3. En la Figura 18 se observa que todas las series presentan leves asimetrías y sólo cinco tienen valores atípicos o *outliers*. En general, la variabilidad en los datos diarios de temperatura se mantiene entre 15 y 30 °C, exceptuando la estación 205 que presenta valores muy por debajo de la media de las otras estaciones, por estar ubicada en el corregimiento de Santa Elena, en las laderas de la zona de estudio.

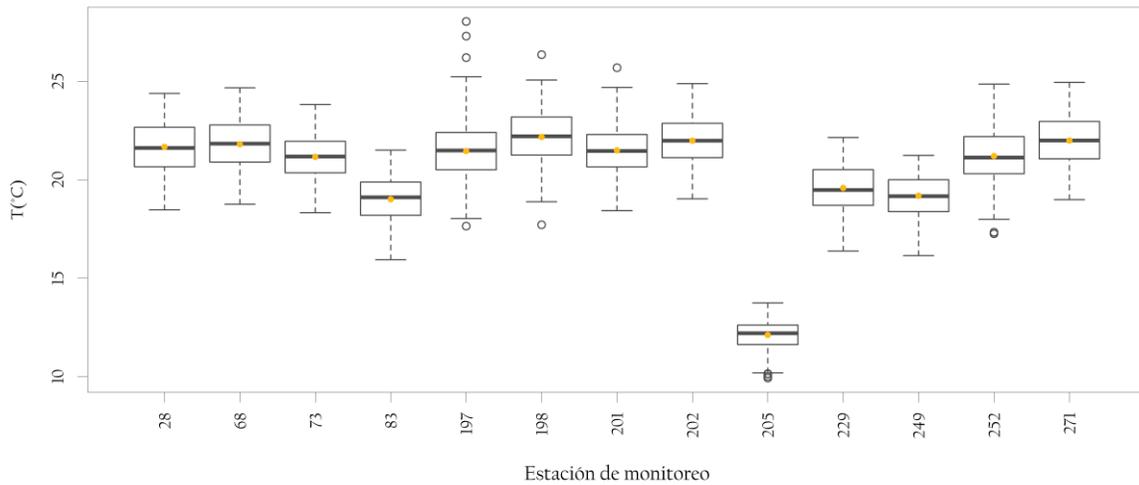


Figura 18. Distribución de los cuartiles para las series de temperatura promedio diaria seleccionadas

Al evaluar los promedios mensuales de las series seleccionadas (Figura 19), se observa que para 2018 se presentó un comportamiento bimodal de la temperatura en el Valle de Aburrá, con valores altos en marzo y entre los meses junio y agosto. Los valores más altos fueron registrados en la estación 198, mientras que los niveles más bajos, como se dijo previamente, se presentaron en la estación 205, sin embargo, el comportamiento general de los datos se conserva.

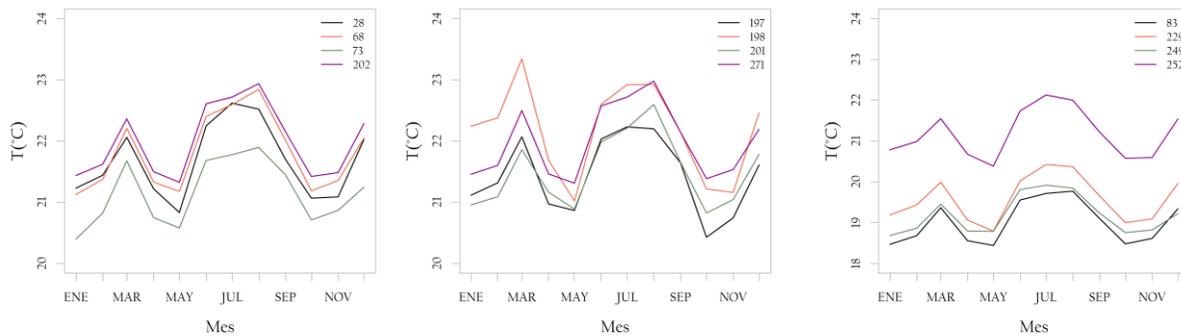


Figura 19. Temperatura promedio mensual para las estaciones seleccionadas

En la Figura 20 se muestra la ubicación de las 13 estaciones seleccionadas, distribuidas a lo largo de la zona de estudio. Se puede observar que, en promedio para el año 2018, las temperaturas del aire más altas fueron registradas en estaciones ubicadas en los centros de las cabeceras

municipales, mientras que las temperaturas más frías estuvieron asociadas a las estaciones en las laderas de la zona de estudio.

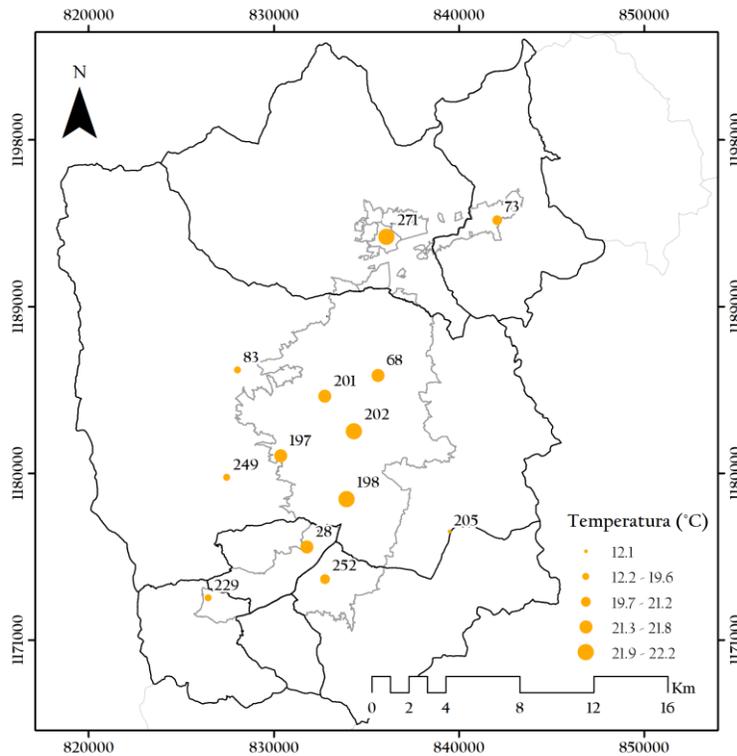


Figura 20. Ubicación estaciones de temperatura seleccionadas

– Velocidad del viento (m/s)

Para las series de velocidad del viento de las 22 estaciones de monitoreo disponibles, se obtuvo un promedio anual que oscilaba entre 0.5 y 3.6 m/s, con media 1.5 m/s. Los valores máximos de la variable fueron registrados en la estación 205, encontrándose muy por encima de la media general de las estaciones (7.7 m/s). Por otro lado, la mayoría de las series presentaron ausencia de datos en el periodo de análisis, siendo éste menor al umbral establecido del 25%; sin embargo, las estaciones 12, 48, 313 y 318, presentaron un alto porcentaje de valores perdidos, superando dicho umbral, por ende, fueron excluidas de posteriores análisis.

Tabla 22. Estadísticos descriptivos para las series de velocidad del viento promedio diaria, en todo el periodo de análisis

Código	Media (m/s)	Mediana (m/s)	Desviación estándar (m/s)	Máximo (m/s)	Mínimo (m/s)	Porcentaje perdidos (%)
12	0.8	0.8	0.2	1.1	0.4	70.7
25	1.2	1.2	0.4	2.2	0.3	13.7
28	0.8	0.8	0.3	2.0	0.3	1.4
38	0.5	0.5	0.2	1.2	0.2	4.9
48	0.9	0.9	0.3	1.9	0.3	76.7
59	1.4	1.3	0.4	3.5	0.4	24.9
68	1.7	1.7	0.5	2.8	0.7	0.5
73	1.8	1.8	0.4	2.7	1.0	0.0
83	1.3	1.2	0.3	2.0	0.8	0.8
197	1.1	1.1	0.2	2.4	0.6	5.2
198	1.1	1.1	0.3	1.8	0.3	1.4
201	2.5	2.5	0.8	4.7	1.0	3.0
202	1.6	1.5	0.5	3.1	0.8	1.1
203	2.0	2.0	0.6	3.9	0.9	5.5
205	3.6	3.2	1.4	7.7	1.3	0.3
206	1.1	1.1	0.2	1.9	0.4	4.9
229	1.8	1.7	0.4	3.5	1.1	0.0
249	1.2	1.2	0.2	2.2	0.8	0.0
252	2.0	1.9	0.4	3.5	1.0	3.0
271	2.1	2.1	0.5	3.7	1.0	1.1
313	1.1	1.1	0.2	1.7	0.7	42.2
318	1.3	1.2	0.2	2.1	0.7	45.8

En la Figura 21 se puede observar que, al ser las mismas estaciones de la Red Meteorológica del Valle de Aburrá, los registros de las series de velocidad del viento presentaron un comportamiento similar a los de las series de temperatura (Figura 17), encontrándose que las estaciones 25, 38, 59, 203 y 206, presentaron datos faltantes continuos en un periodo igual o mayor a un mes y fueron descartadas para posteriores análisis (Anexo C).

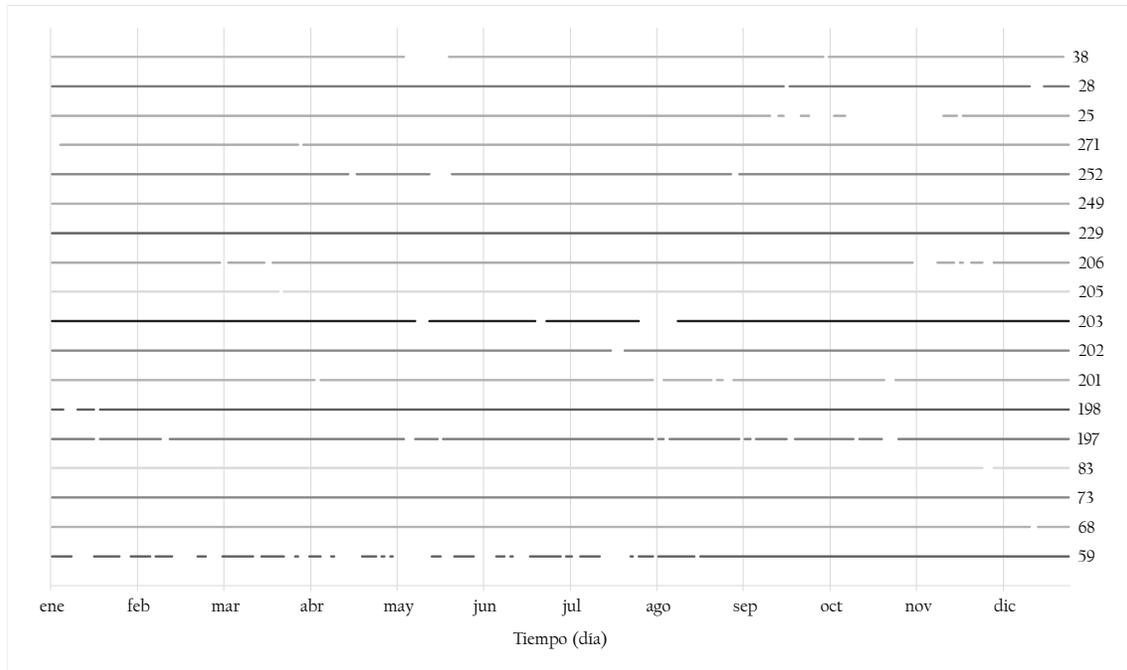


Figura 21. Gráfica de registros de velocidad del viento promedio diaria en todo el periodo de análisis

Como resultado de este análisis exploratorio, se seleccionaron 13 estaciones de monitoreo, aplicando los criterios de selección descritos en la sección 5.3.3. En la Figura 22 se observa que la mayoría de las series presentan leves asimetrías, con presencia de valores atípicos o *outliers* para ocho estaciones. La estación 205 presenta una simetría positiva, indicando que los datos por encima de la media están más dispersos. En general, la variabilidad en los datos diarios de velocidad del viento se mantiene en valores menores a 4 m/s, exceptuando la estación 205 que presenta valores de hasta 8 m/s aproximadamente, estando por encima de la media de las otras estaciones.

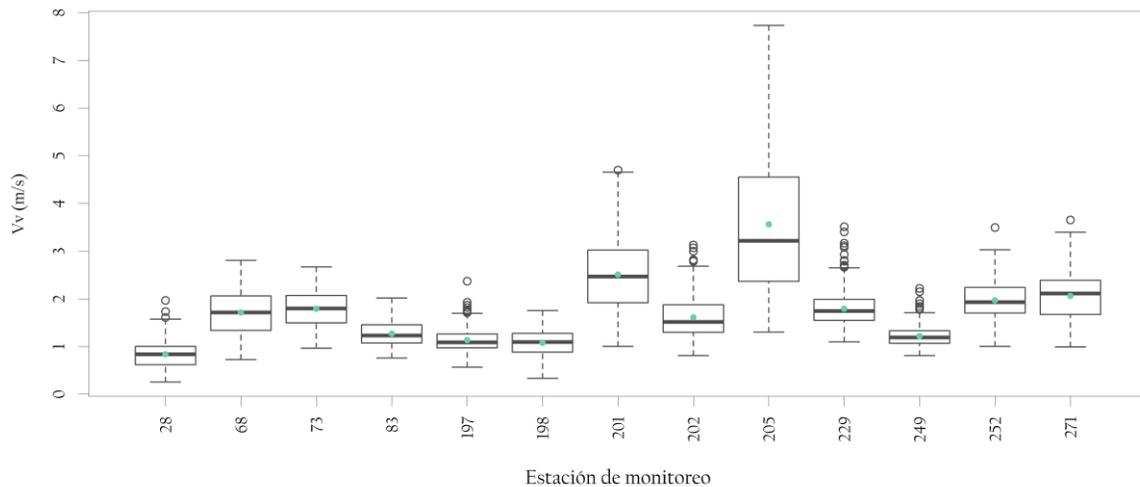


Figura 22. Distribución de los cuartiles para las series de velocidad del viento promedio diaria seleccionadas

Al evaluar los promedios mensuales de las series seleccionadas (Figura 23), se observa que para 2018 se presentó un comportamiento bimodal de la velocidad del viento en el Valle de Aburrá, con valores altos en febrero-marzo y junio-agosto. En promedio, los valores más bajos fueron registrados en la estación 28, mientras que los niveles más altos en la estación 205, la cual presenta promedios muy por encima de las otras estaciones, especialmente entre junio y agosto, en donde se duplican; sin embargo, el comportamiento general de los datos se conserva.

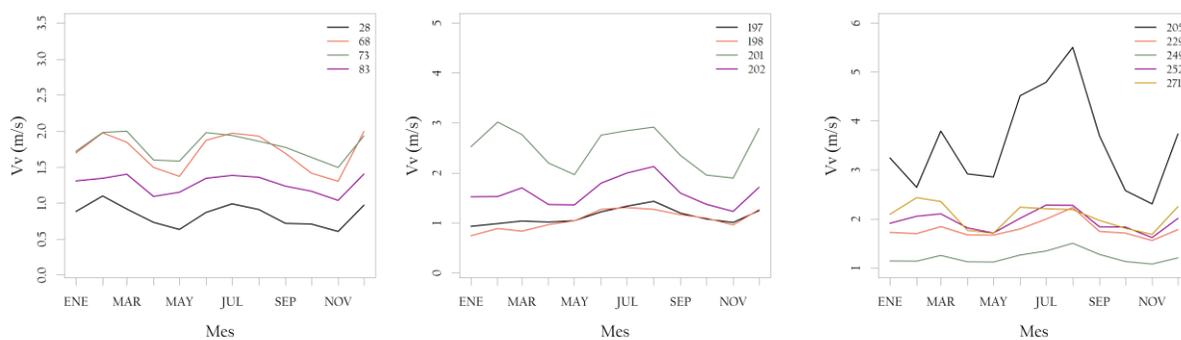


Figura 23. Velocidad del viento promedio mensual para las estaciones seleccionadas

En la Figura 24 se puede observar que, en promedio, las velocidades del viento más bajas fueron registradas en estaciones ubicadas en el centro-occidente, mientras que las magnitudes más altas se presentaron en la estación ubicada en la ladera de la zona de estudio. Existe gran variabilidad en los datos espacial a lo largo de la zona de estudio, poniendo en evidencia la variabilidad local de la variable evaluada.

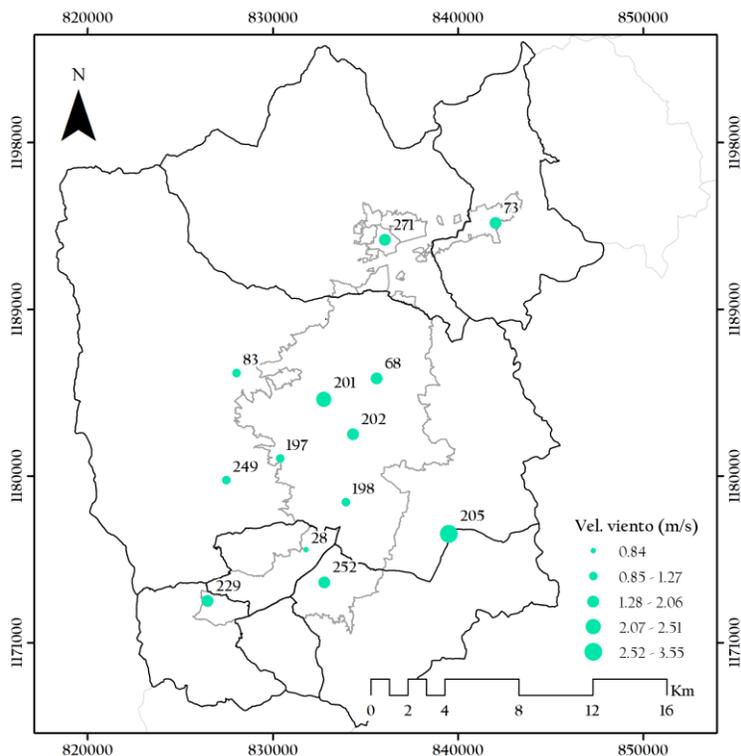


Figura 24. Ubicación estaciones de velocidad del viento seleccionadas

– Precipitación (mm)

De las 29 estaciones pluviográficas, se obtuvo una precipitación acumulada anual entre 498.8 y 3669.2 mm, con media 1839.1 mm (Tabla 23). El promedio de los datos de precipitación acumulada diaria osciló entre 1.4 y 10.1 mm y los valores máximos se presentaron en las estaciones 205 y 206, además, todas las estaciones presentaron días sin lluvia (0 mm). Por otro lado, para el periodo activo de medición de las estaciones, ninguna estación presentó un porcentaje de valores perdidos superior al 25%.

Tabla 23. Estadísticos descriptivos para las series de precipitación diaria, en todo el periodo de análisis

Código	Acumulado (mm)	Media (mm)	Mediana (mm)	Desviación estándar (mm)	Máximo (mm)	Porcentaje perdidos (%)
68	2115.8	5.8	0.6	10.6	67.4	0.5
73	1876.2	5.1	0.4	9.7	85.5	0.0
77	1201.9	3.6	0.6	6.2	47.2	8.2
83	2351.1	6.5	0.4	12.3	87.1	0.8
121	1688.1	5.0	0.5	9.0	74.7	6.6
129	1506.2	4.3	0.5	6.9	42.9	3.6
146	2026.7	5.6	1.5	8.3	44.5	0.5
154	1484.2	4.2	0.3	8.3	67.3	2.7
184	1641.9	4.6	1.0	7.4	47.2	1.9
189	1205.0	3.6	0.0	8.0	66.3	8.5
193	2355.1	6.6	1.5	9.9	49.0	2.5
197	1346.9	3.9	0.5	6.7	53.1	5.2
198	632.3	1.8	0.6	2.4	12.7	1.4
201	2300.2	6.5	1.0	11.2	72.1	3.0
202	2311.5	6.4	1.1	10.8	68.9	1.1
205	3669.2	10.1	1.6	17.0	141.8	0.3
206	3172.4	9.1	2.0	15.4	122.2	4.9
211	1094.5	3.0	0.3	5.4	28.7	0.3
230	2509.3	6.9	2.3	10.5	69.3	0.0
241	1676.9	4.7	0.5	9.0	66.8	2.5
242	1711.5	4.7	0.3	8.9	80.5	0.0
244	2041.1	5.6	1.7	8.5	52.3	0.3
249	1339.5	3.7	1.0	5.7	30.2	0.0
252	1220.2	3.4	0.9	5.3	29.3	3.0
261	2384.0	6.5	0.8	11.4	92.2	8.5
266	1914.4	5.3	1.1	8.6	55.9	10.1
267	2332.7	6.6	1.8	9.4	64.8	3.3
271	498.8	1.4	0.4	1.9	10.5	1.1
282	1727.7	5.5	0.8	10.1	86.1	13.2

En la Figura 25 se puede observar que la mayoría de las estaciones presentaron días sin registros, donde la estación 282 presenta el mayor bloque de información no disponible en los primeros meses del año; seguida por las estaciones 189, 261 y 266, las cuales no presentan datos en un mes completo. Al evaluar el porcentaje de valores perdidos mensual para cada serie, se encontró que

las estaciones 77, 82, 189, 206, 261, 266, 267 y 282 superaron el umbral establecido y fueron descartadas de posteriores análisis (Anexo D).

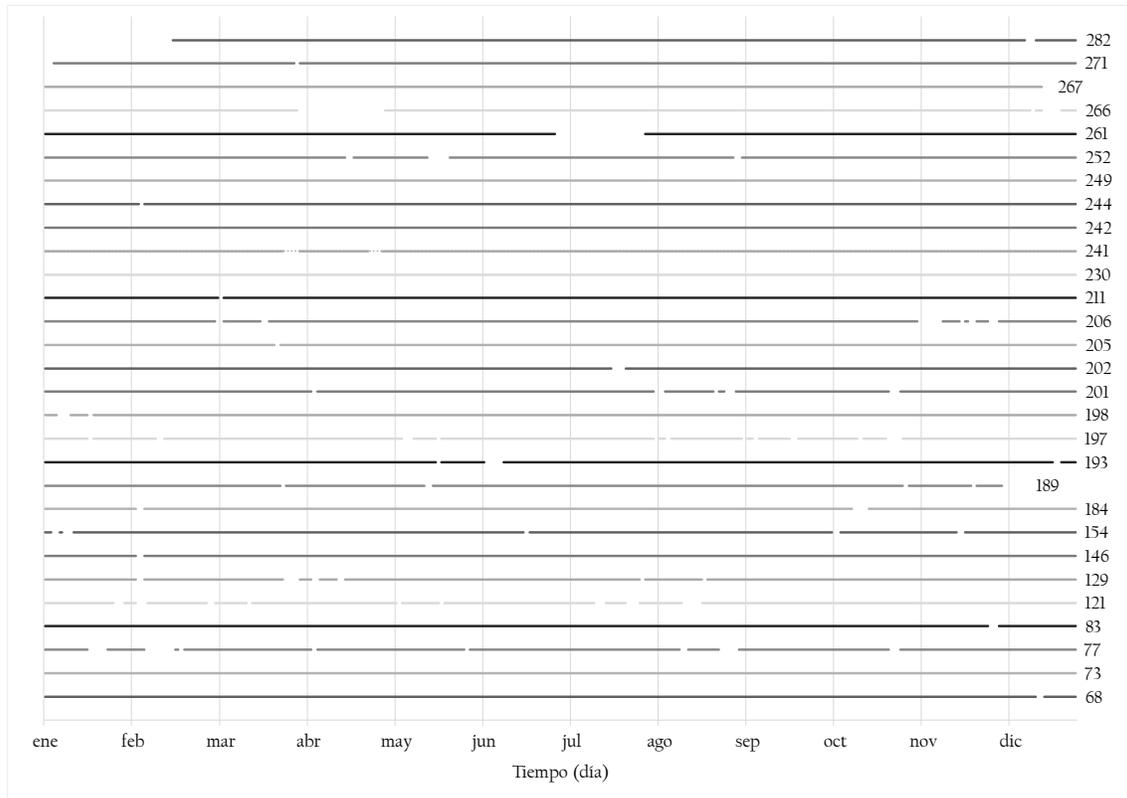


Figura 25. Gráfica de registros de precipitación diaria en todo el periodo de análisis

Como resultado de este análisis exploratorio, se seleccionaron 22 estaciones de monitoreo, aplicando los criterios de selección descritos en la sección 5.3.3. En la Figura 26 se observa que todas las series presentan asimetría positiva, indicando una mayor dispersión en los datos mayores a la mediana; además se observa la presencia de valores atípicos o *outliers* en todas las estaciones. En general, la mayoría de los datos de precipitación acumulada diaria se mantiene en valores menores a 15 mm.

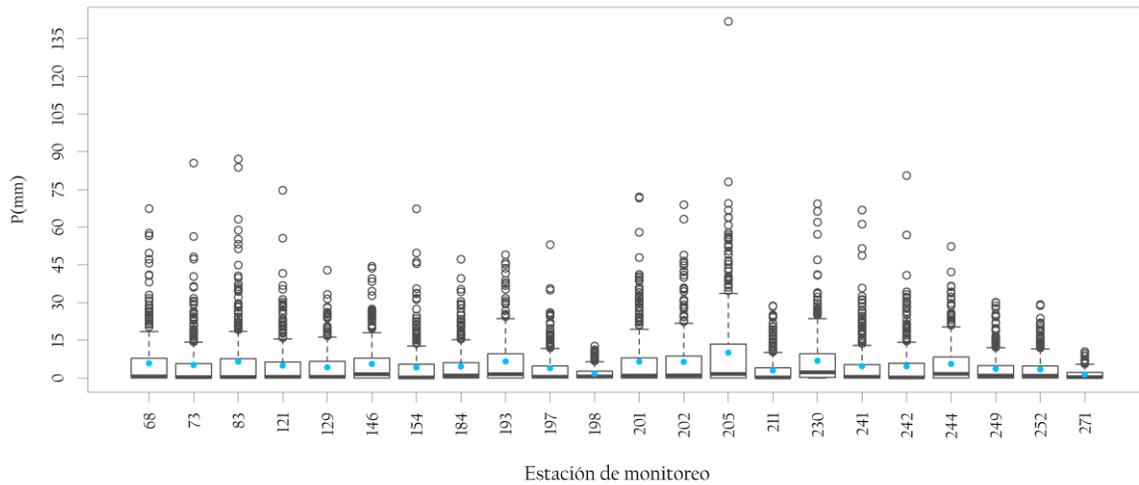


Figura 26. Distribución de los cuartiles para las series de precipitación diaria seleccionadas

Al evaluar la precipitación acumulada mensual de las series seleccionadas (Figura 27), se observa que para 2018 se presentó un comportamiento bimodal en el Valle de Aburrá, con valores máximos en mayo y octubre. Los valores más bajos fueron registrados en las estaciones 198 y 271, mientras que los niveles más altos en la estación 205, especialmente en el primer pico del año; sin embargo, el comportamiento general de los datos se conserva.

En la Figura 28 se puede observar que, los valores de precipitación acumulada promedio más altas fueron registradas en estaciones ubicadas en el sur y centro de la zona de estudio, donde la magnitud más alta se presentó en la estación ubicada en la ladera del municipio de Medellín. Existe gran variabilidad en los datos espaciales a lo largo de la zona de estudio, poniendo en evidencia la variabilidad local de la variable evaluada.

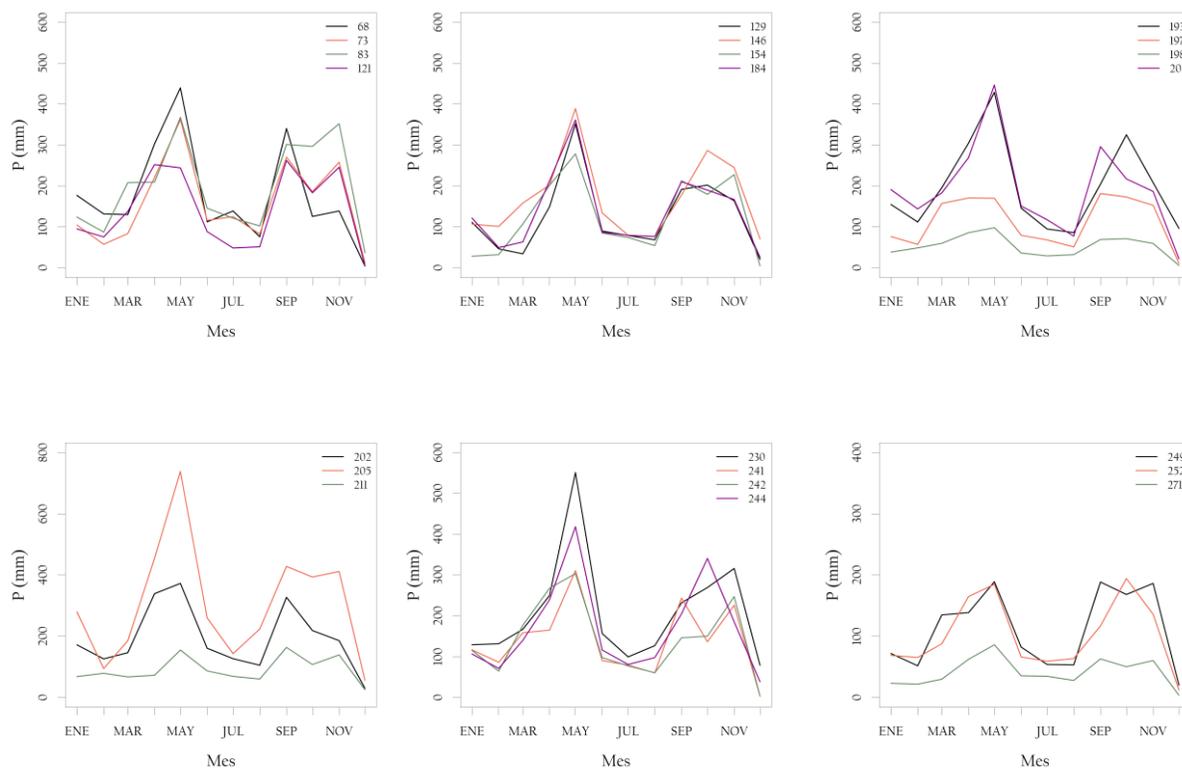


Figura 27. Precipitación acumulada mensual para las estaciones seleccionadas

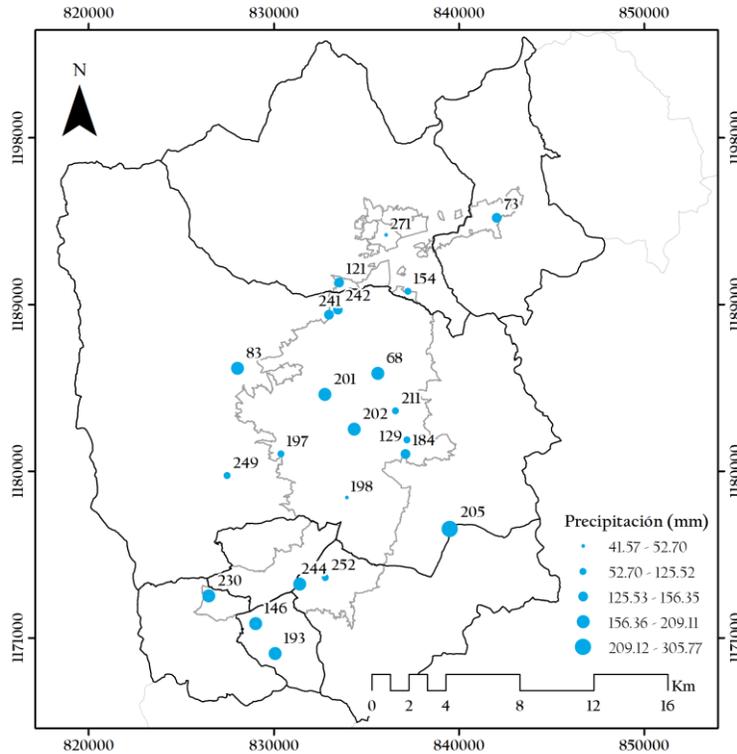


Figura 28. Ubicación estaciones de precipitación seleccionadas

6.1.3. Aforos vehiculares

De las 40 cámaras de fotodetección, se obtuvo un flujo vehicular que, en promedio, osciló entre 5204 y 57537.7 vehículos por día, con media de 20304 veh./día (Tabla 24). Los valores máximos se presentaron en las cámaras 48, 63 y 66, ubicadas en vías arteria, y los valores mínimos en las estaciones 75 y 76, ubicadas en una vía arteria a las afueras de la cabecera municipal del municipio de Medellín. Por otro lado, para el todo el periodo de análisis las cámaras 50 y 77 presentaron un porcentaje de valores perdidos superior al 25% y no se incluyeron en análisis posteriores.

Tabla 24. Estadísticos descriptivos para las series de aforos vehiculares diarios en todo el periodo de análisis

Cámara	Media (vehículo/día)	Mediana (vehículo/día)	Desviación estándar (vehículo/día)	Máximo (vehículo/día)	Mínimo (vehículo/día)	Porcentaje perdidos (%)
1	33709.7	34564.0	12001.5	86368	1063	1.5
2	20815.4	22073.0	6519.9	43594	1561	1.1
3	21671.0	22691.0	7299.6	49052	185	1.1

Cámara	Media (vehículo/día)	Mediana (vehículo/día)	Desviación estándar (vehículo/día)	Máximo (vehículo/día)	Mínimo (vehículo/día)	Porcentaje perdidos (%)
5	22968.6	23964.5	6723.3	48060	1190	7.9
9	23646.2	24307.0	8505.1	63042	235	9.0
12	17171.4	18879.5	5772.3	36934	628	0.8
13	14609.5	15012.0	9517.7	43836	340	0.0
14	10845.3	10855.5	3154.2	23678	1022	16.7
15	17231.1	18647.0	6594.3	40946	225	3.3
18	22027.3	22643.0	5968.3	47398	3422	1.4
19	23669.8	23790.0	7947.7	59256	1645	1.4
24	17658.8	18500.0	4853.6	37450	20	1.1
27	23547.5	24970.5	7324.3	41906	688	24.4
31	17672.4	18161.0	4789.2	37972	813	1.1
33	22275.5	24775.0	10464.2	54910	806	0.0
35	8123.2	8670.0	2751.8	17496	369	10.7
38	10917.5	11526.5	3831.5	25060	205	2.5
40	22293.1	24480.0	8821.4	51958	517	1.6
48	43818.9	48015.0	27095.9	128766	57	0.3
50	12953.9	13386.0	4087.1	28504	2293	32.1
51	16418.7	17326.5	4446.9	33938	1415	0.3
53	18869.8	19871.0	4910.2	44844	3210	9.0
54	25165.7	25988.0	8000.7	58270	2734	0.8
56	13699.3	9383.0	10837.9	51868	576	0.3
58	18268.2	18424.0	4187.3	37070	255	6.0
59	16300.3	16970.5	6435.0	42326	1255	12.9
61	9014.8	8901.0	2130.2	19010	1897	0.5
63	57537.7	58372.5	18520.4	141262	238	0.8
65	26483.6	26405.0	5307.4	47788	12218	0.0
66	48792.9	50943.0	13863.7	101026	864	0.0
67	15456.6	16551.0	5446.8	34676	64	1.6
69	9867.5	9695.0	2523.7	21236	1766	0.5
70	10876.9	9463.5	6505.2	33052	482	1.4
71	27547.6	29605.5	7795.2	47296	4599	0.8
72	19066.3	20478.0	5531.1	41736	82	1.6
73	15259.7	16267.0	5159.3	37944	734	1.1
74	19735.9	19640.0	5765.5	42752	6252	0.0
75	5840.1	4966.0	2920.4	29770	731	0.0
76	5204.0	4550.5	2316.2	22620	982	0.3
77	25148.4	25665.5	7479.8	53392	2755	29.3

En la Figura 29 se puede observar que la mayoría de las cámaras presentaron días sin registros, donde la 27 presenta el mayor bloque de información no disponible en los últimos meses del año; seguida por las estaciones 14 y 59, las cuales no presentan datos en más de un mes. Al evaluar el porcentaje de valores perdidos mensual para cada serie, se encontró que las estaciones 5, 9, 14, 15, 27, 35, 53, 58 y 59 superaron el umbral establecido y fueron descartadas de posteriores análisis (Anexo E).

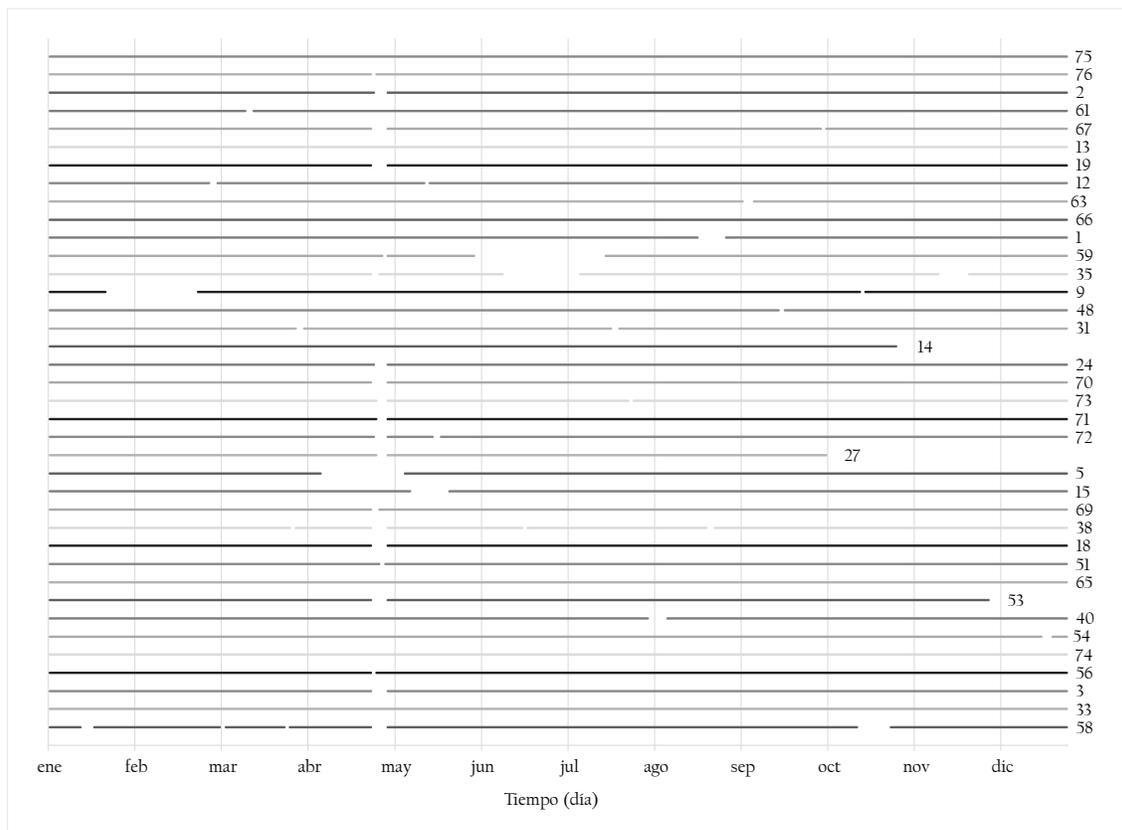


Figura 29. Gráfica de registros de aforos vehiculares diarios en todo el periodo de análisis

Como resultado de este análisis exploratorio, se seleccionaron 29 estaciones de monitoreo, aplicando los criterios de selección descritos en la sección 5.3.3. En la Figura 30 se observa que todas las series presentan asimetría negativa, indicando que los datos comprendidos entre el 25 y el 50% de la serie están ligeramente más dispersas que entre el 50 y el 75%; además se observa la presencia de valores atípicos o *outliers* en todas las estaciones. En general, la mayoría de los datos de flujo vehicular se mantiene en valores menores a 30000 vehículos por día.

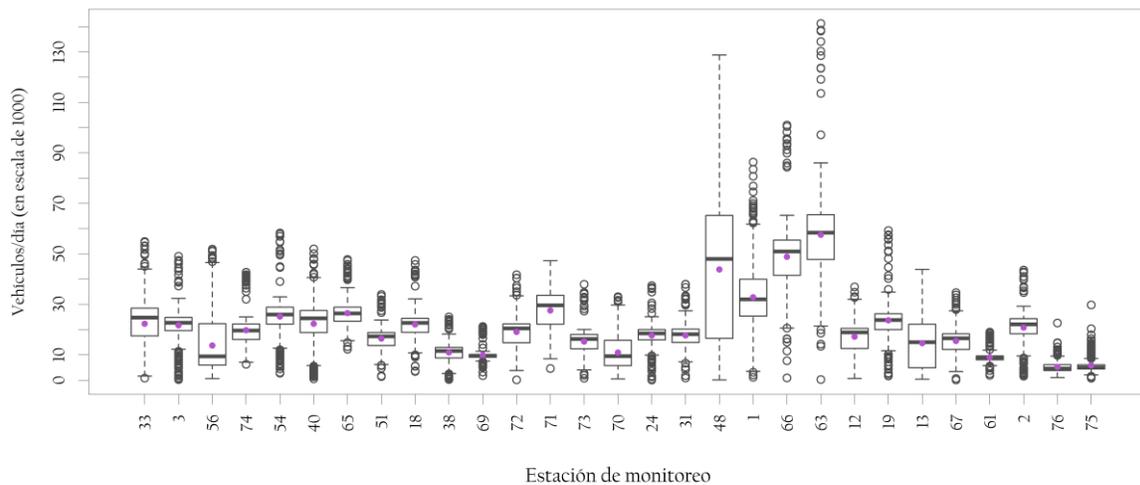


Figura 30. Distribución de los cuartiles para las series de aforos vehiculares diarios seleccionados

Al evaluar el flujo vehicular mensual de las cámaras seleccionadas (Figura 31), se observa que para 2018, el máximo de todas las series se presentó en el mes de enero y, para la mayoría, los valores no oscilaron en gran medida para los otros meses. Sin embargo, en las cámaras 13, 40 y 48, ubicadas al sur de Medellín, se presenta una disminución sustancial de la cantidad de vehículos por día registrados para el último trimestre del año. Por otro lado, la cámara 56, ubicada en la Avenida Colombia con la 65, en el occidente de la cabecera municipal, reportó una disminución en el promedio mensual de vehículos por día desde el mes de mayo.

En la Figura 32 se puede observar que, el flujo vehicular más alto fue registrado en las cámaras ubicadas en el sur de Medellín, donde la magnitud más alta se presentó en la cámara 66, seguida por la 1, ambas ubicadas en la Avenida Regional, sentido sur-norte. Los valores mínimos registrados están asociados a las cámaras 75 y 76, ubicadas en la ladera occidental, por fuera de la cabecera municipal. Existe gran variabilidad en los datos espaciales a lo largo de la zona de estudio, poniendo en evidencia la variabilidad local de la variable evaluada.

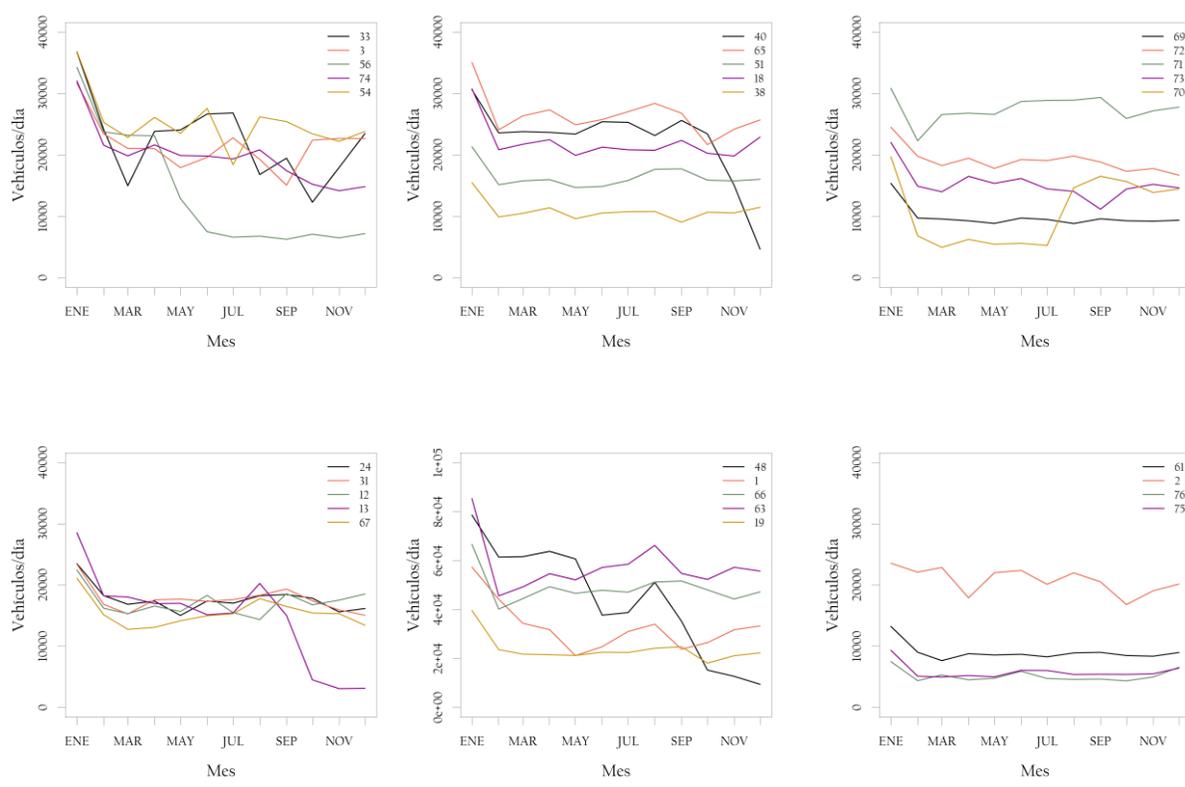


Figura 31. Aforos vehiculares mensuales para las cámaras seleccionadas

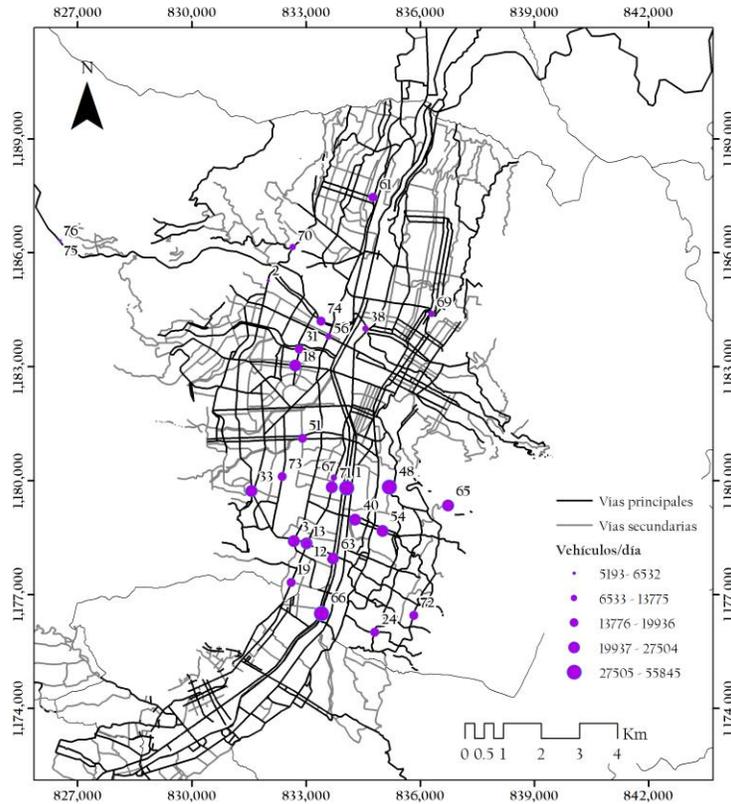


Figura 32. Ubicación de cámaras de fotodetección seleccionadas

6.2. Momento dos: Preprocesamiento de las variables

6.2.1. Interpolaciones y gradientes espaciales de las variables ambientales

Con los valores promedio mensuales de las estaciones meteorológicas seleccionadas, se realizaron interpolaciones espaciales para cada mes y cada variable, con el fin de estimar los valores en las estaciones de $PM_{2.5}$, debido a que sólo dos estaciones meteorológicas median en un mismo punto que una estación del contaminante (COP.CVID: 73 e ITA.CJUS: 28). En la Tabla 25 se muestra el error cuadrático medio asociado con cada algoritmo de interpolación espacial utilizado para cada mes. Se observa también que, para la zona de estudio y el periodo de análisis evaluado, el mejor interpolador difería entre las tres variables ambientales, donde se encontró que, para la temperatura y la velocidad del viento, el método que presentó mejor desempeño fue el IDW con un RMSE promedio de 2.56 y 0.71, respectivamente, disminuyendo el error en la interpolación en 8 de los 12 meses analizados para la primera variable, y en 11 meses para la segunda. Por otro lado, el mejor interpolador para la precipitación fue el Spline, con un RMSE promedio de 65.88, disminuyendo el error de la interpolación en 10 meses.

Tabla 25. RMSE de las interpolaciones espaciales para las variables ambientales

Mes	Temperatura (RMSE)		Velocidad del viento (RMSE)		Precipitación (RMSE)	
	IDW	Spline	IDW	Spline	IDW	Spline
Enero	2.78	2.86	0.68	0.68	59.54	57.79
Febrero	2.84	2.80	0.58	0.66	31.46	24.61
Marzo	2.90	2.99	0.78	0.82	50.10	41.21
Abril	2.66	2.67	0.58	0.61	103.60	102.32
Mayo	2.48	2.47	0.56	0.57	158.17	154.57
Junio	1.06	0.76	0.89	0.94	51.26	51.10
Julio	2.89	2.93	0.94	0.98	34.84	34.52
Agosto	2.92	2.94	1.13	1.17	39.25	42.42
Septiembre	2.70	2.71	0.72	0.74	99.04	96.51
Octubre	2.55	2.57	0.51	0.52	87.87	82.24
Noviembre	2.50	2.48	0.46	0.48	91.50	91.80
Diciembre	2.84	2.88	0.71	0.76	17.36	14.96

Con las superficies interpoladas se estimaron los valores de cada variable ambiental en las estaciones de monitoreo de PM_{2.5}, mostrados en la Tabla 26, Tabla 27 y Tabla 28. Además, a partir de las mismas superficies, se calcularon los gradientes espaciales de cada variable para los mismos puntos mostrados en la Tabla 29, Tabla 30 y Tabla 31, con la metodología descrita en la sección 5.4.1, obteniendo seis posibles variables explicativas para la construcción de los modelos LUR. Al evaluar la normalidad de los datos espaciales mensuales, aplicando el criterio de Webster y Oliver se encontró que sólo 23 de las 72 variables ambientales eran estacionarias, debido a la presencia de valores atípicos que no fueron removidos para no comprometer el rendimiento de los modelos LUR, por la disminución del número de estaciones de PM_{2.5}.

Tabla 26. Datos de temperatura promedio mensual en cada estación de monitoreo de PM_{2.5}

Estación	Temperatura promedio mensual (°C)											
	1	2	3	4	5	6	7	8	9	10	11	12
0	21.34	21.29	22.38	21.36	21.02	22.25	22.60	22.85	22.03	21.27	21.23	22.08
1	20.40	20.83	21.68	20.75	20.58	21.69	21.78	21.90	21.45	20.71	20.87	21.25
2	20.73	20.44	21.51	20.63	19.87	21.17	22.05	21.94	21.17	20.51	20.07	21.48
3	19.28	19.52	20.09	19.17	18.94	20.16	20.52	20.47	19.75	19.09	19.20	20.05
4	21.23	21.44	22.06	21.22	20.83	22.24	22.62	22.52	21.70	21.06	21.08	22.02
5	19.24	19.56	20.05	19.12	18.93	20.17	20.48	20.43	19.71	19.05	19.22	20.01
6	20.95	21.24	21.91	21.09	21.03	22.20	22.25	22.52	21.67	20.86	21.19	21.77
7	20.83	21.07	21.76	20.72	20.61	21.80	21.96	21.95	21.35	20.29	20.56	21.36
8	20.78	20.11	21.78	20.89	19.93	21.05	22.10	22.29	21.52	20.73	20.10	21.59
9	20.91	20.97	21.85	20.81	20.58	21.73	22.05	22.04	21.45	20.34	20.52	21.45
10	20.16	19.17	21.09	19.97	18.62	19.78	21.12	21.13	20.47	19.71	18.77	20.64

Estación	Temperatura promedio mensual (°C)											
	1	2	3	4	5	6	7	8	9	10	11	12
11	18.69	18.96	19.59	18.78	18.70	19.83	19.94	20.00	19.32	18.68	18.87	19.54
12	20.79	20.94	21.80	20.91	20.72	21.90	22.11	22.31	21.53	20.74	20.91	21.59

Tabla 27. Datos de velocidad del viento promedio mensual en cada estación de monitoreo de PM_{2.5}

Estación	Velocidad del viento promedio mensual (m/s)											
	1	2	3	4	5	6	7	8	9	10	11	12
0	2.00	2.38	2.31	1.75	1.69	2.21	2.19	2.18	1.95	1.78	1.66	2.22
1	1.72	1.98	2.00	1.60	1.59	1.98	1.94	1.86	1.78	1.64	1.50	1.94
2	1.89	1.96	2.00	1.72	1.63	1.93	2.18	2.18	1.76	1.74	1.53	1.93
3	1.64	1.68	1.81	1.64	1.63	1.78	1.97	2.19	1.72	1.67	1.53	1.77
4	0.90	1.10	0.92	0.74	0.64	0.87	0.99	0.92	0.72	0.71	0.61	0.98
5	1.71	1.69	1.83	1.66	1.66	1.79	1.99	2.22	1.73	1.70	1.55	1.78
6	2.03	2.28	2.17	1.75	1.61	2.21	2.32	2.36	1.93	1.63	1.53	2.29
7	0.96	1.14	1.18	1.10	1.11	1.33	1.45	1.54	1.27	1.14	1.06	1.35
8	1.85	1.91	1.89	1.54	1.45	1.96	2.07	2.11	1.74	1.47	1.35	1.99
9	1.13	1.15	1.20	1.11	1.12	1.35	1.46	1.56	1.29	1.15	1.08	1.37
10	1.69	1.46	1.59	1.41	1.40	1.87	1.99	2.09	1.64	1.41	1.26	1.78
11	1.35	1.39	1.44	1.14	1.18	1.40	1.45	1.44	1.29	1.20	1.07	1.46
12	1.88	2.00	1.95	1.57	1.49	1.99	2.08	2.09	1.77	1.52	1.40	2.04

Tabla 28. Datos de precipitación acumulada mensual en cada estación de monitoreo de PM_{2.5}

Estación	Precipitación acumulada mensual (mm)											
	1	2	3	4	5	6	7	8	9	10	11	12
0	50.03	40.53	68.82	126.18	158.53	57.92	47.92	37.49	135.61	100.25	109.41	5.36
1	104.15	57.32	83.36	220.13	361.01	115.58	124.75	82.43	269.90	185.53	257.89	11.46
2	80.71	67.68	95.05	177.92	218.59	76.82	63.04	68.78	134.35	198.08	146.59	18.47
3	121.43	118.97	161.89	235.58	493.25	144.98	93.05	118.71	222.53	260.71	297.41	69.51
4	81.56	66.79	112.80	180.69	261.17	84.91	64.63	77.14	155.34	218.98	166.21	23.11
5	126.76	125.86	165.81	244.73	520.26	151.79	96.41	121.11	224.15	270.83	303.66	76.54
6	170.30	127.48	153.76	278.41	405.52	133.12	119.80	74.76	301.37	178.98	172.84	16.34
7	78.30	60.37	146.22	170.02	195.29	82.28	66.60	55.69	179.55	178.11	157.28	15.29
8	95.42	79.72	82.71	148.53	260.49	95.37	85.25	66.47	210.89	137.34	152.49	19.12
9	91.91	68.50	161.28	183.43	217.70	92.47	76.84	56.01	204.76	186.46	161.43	14.78
10	117.22	65.75	98.87	215.25	317.21	105.18	73.64	76.59	196.04	193.52	155.34	23.92
11	117.32	83.91	191.91	202.91	343.47	133.65	109.31	96.33	282.11	268.71	327.70	31.95
12	91.73	70.07	118.35	229.63	317.06	94.32	90.61	61.81	236.01	155.07	199.09	6.20

Tabla 29. Datos de gradiente de temperatura promedio mensual en cada estación de monitoreo de PM_{2.5}

Estación	Gradiente de temperatura promedio mensual (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
0	0.020	0.029	0.020	0.018	0.027	0.030	0.020	0.022	0.019	0.019	0.029	0.019
1	0.000	0.006	0.000	0.000	0.006	0.007	0.000	0.000	0.000	0.000	0.005	0.000
2	0.017	0.090	0.013	0.016	0.083	0.093	0.022	0.019	0.017	0.019	0.084	0.021
3	0.023	0.019	0.024	0.025	0.024	0.024	0.022	0.023	0.023	0.023	0.020	0.020
4	0.005	0.026	0.005	0.006	0.018	0.021	0.007	0.007	0.006	0.007	0.017	0.007
5	0.015	0.023	0.016	0.016	0.026	0.027	0.015	0.015	0.015	0.015	0.024	0.014
6	0.008	0.015	0.014	0.006	0.014	0.019	0.015	0.008	0.014	0.015	0.014	0.010
7	0.045	0.059	0.049	0.040	0.050	0.054	0.042	0.041	0.047	0.025	0.047	0.040
8	0.026	0.066	0.030	0.030	0.064	0.070	0.033	0.037	0.033	0.031	0.064	0.032
9	0.040	0.065	0.043	0.032	0.057	0.060	0.036	0.030	0.038	0.017	0.051	0.031
10	0.149	0.150	0.158	0.134	0.119	0.138	0.144	0.144	0.133	0.124	0.120	0.143
11	0.039	0.037	0.040	0.038	0.035	0.036	0.039	0.040	0.038	0.035	0.034	0.035
12	0.007	0.013	0.007	0.009	0.013	0.014	0.010	0.011	0.009	0.008	0.012	0.010

Tabla 30. Datos de gradiente de velocidad del viento promedio mensual en cada estación de monitoreo de PM_{2.5}

Estación	Gradiente de velocidad del viento promedio mensual (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
0	0.010	0.010	0.008	0.004	0.004	0.005	0.003	0.003	0.004	0.005	0.005	0.006
1	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.017	0.021	0.022	0.020	0.018	0.015	0.020	0.018	0.016	0.021	0.018	0.016
3	0.012	0.005	0.008	0.009	0.010	0.006	0.007	0.011	0.008	0.010	0.009	0.005
4	0.030	0.004	0.006	0.006	0.006	0.007	0.007	0.008	0.006	0.006	0.005	0.007
5	0.006	0.002	0.004	0.005	0.005	0.002	0.003	0.006	0.004	0.005	0.005	0.002
6	0.032	0.051	0.044	0.034	0.028	0.042	0.041	0.045	0.031	0.025	0.028	0.044
7	0.006	0.019	0.018	0.011	0.008	0.014	0.014	0.012	0.009	0.008	0.006	0.012
8	0.010	0.006	0.002	0.002	0.003	0.004	0.005	0.009	0.003	0.002	0.002	0.002
9	0.028	0.034	0.032	0.020	0.016	0.026	0.026	0.025	0.019	0.015	0.014	0.026
10	0.038	0.030	0.046	0.030	0.027	0.046	0.050	0.060	0.036	0.022	0.021	0.037
11	0.007	0.008	0.007	0.008	0.005	0.010	0.011	0.013	0.008	0.006	0.006	0.009
12	0.003	0.003	0.005	0.002	0.004	0.003	0.002	0.003	0.003	0.004	0.004	0.001

Tabla 31. Datos de gradiente de precipitación acumulada mensual en cada estación de monitoreo de PM_{2.5}

Estación	Gradiente de precipitación promedio mensual (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
0	2.321	1.653	3.371	5.452	6.018	1.901	1.080	1.353	6.281	4.300	7.092	0.170
1	0.634	0.233	0.142	0.809	1.639	0.497	0.633	0.030	1.176	0.690	0.097	0.064
2	2.323	0.518	1.633	3.075	7.347	2.196	0.899	0.957	3.585	2.680	2.082	1.148

Estación	Gradiente de precipitación promedio mensual (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
3	1.059	1.722	0.697	1.904	7.741	1.562	0.875	1.691	1.352	1.300	3.899	1.246
4	1.062	0.419	1.785	2.216	6.127	1.504	0.757	0.632	2.547	4.445	1.219	0.806
5	0.639	1.233	0.329	1.225	6.200	1.013	0.666	1.756	1.585	0.420	3.669	0.532
6	0.936	0.731	2.051	1.199	1.648	1.535	0.551	0.550	1.114	3.099	3.060	0.642
7	0.391	0.525	1.184	0.478	3.136	0.461	0.086	0.692	0.565	0.991	1.015	0.527
8	1.782	1.638	1.830	5.836	7.235	0.624	1.356	1.013	3.700	3.347	2.065	0.539
9	2.149	1.625	0.507	1.995	6.258	1.697	1.160	1.030	2.928	1.849	2.008	0.522
10	3.279	0.540	1.580	5.035	9.044	2.889	1.622	2.263	5.019	4.536	4.455	0.733
11	0.804	0.434	1.797	0.816	3.058	1.387	1.375	1.007	2.321	3.197	4.087	0.545
12	2.970	1.876	1.281	2.135	2.993	0.602	1.170	0.521	2.127	0.879	1.611	0.122

6.2.2. Interpolación espacial de aforos vehiculares

Al evaluar la normalidad de los datos espaciales mensuales de las 29 cámaras seleccionadas por medio de los criterios descritos en la sección 5.4.1 se encontró que, para todos los meses, exceptuando febrero, se consiguió la estacionariedad en los datos por medio de una transformación logarítmica. Por ende, para once meses fue posible aplicar la metodología Kriging, en donde, para cada mes, los mejores modelos de semivariograma teóricos ajustados, con un RMSE promedio de 0.6, se muestran en la Tabla 32. Se observa que, para los datos trabajados, el modelo ajustado con mayor frecuencia fue el K-bessel. Por otro lado, para el mes de febrero no se evaluó el desempeño de los algoritmos de interpolación espacial, donde se encontró que el mejor interpolador fue el IDW, con un RMSE de 11485.3.

Tabla 32. Modelos de semivariograma teóricos ajustados para los aforos vehiculares

Mes	Modelo	Desviación estándar	Varianza	<i>Nugget</i> (θ_n)	<i>Major range</i> (θ_r)	<i>Partial sill</i> (θ_s)	ASE	RMSE	RMSS
1	K-bessel	0.6	0.3	0.0	3035.5	0.2	0.4	0.4	0.9
3	Exponencial	0.6	0.4	0.2	4661.9	0.1	0.5	0.5	1.0
4	Exponencial	0.6	0.4	0.2	4593.9	0.2	0.5	0.5	1.0
5	K-bessel	0.6	0.4	0.0	2004.5	0.3	0.5	0.6	1.1
6	J-Bessel	0.6	0.4	0.2	2465.3	0.0	0.5	0.6	1.1
7	J-Bessel	0.7	0.4	0.2	2574.1	0.0	0.5	0.6	1.0
8	K-bessel	0.6	0.4	0.3	2148.6	0.1	0.6	0.7	1.1
9	Esférico	0.6	0.4	0.2	15332.1	0.4	0.5	0.6	1.1
10	K-bessel	0.6	0.4	0.1	1911.2	0.2	0.6	0.7	1.1
11	K-bessel	0.6	0.4	0.2	1911.2	0.1	0.7	0.7	1.1
12	Hole effect	0.7	0.4	0.4	2004.5	0.1	0.7	0.8	1.0

Con las superficies interpoladas, se asignaron los valores estimados de flujos vehiculares mensuales a las vías principales y secundarias de la zona de estudio. En la Figura 33 se muestran

los resultados de este proceso para los meses enero, junio y diciembre. En general, los mapas son congruentes con el comportamiento observado en los datos originales, conservando los valores más altos para el flujo vehicular en el mes de enero, la baja variación entre los otros meses del año, y el alto flujo vehicular en el sur de la zona de estudio. Además, se puede observar la disminución del flujo vehicular en el mes de diciembre, relacionada a la época de vacaciones en la zona de estudio.

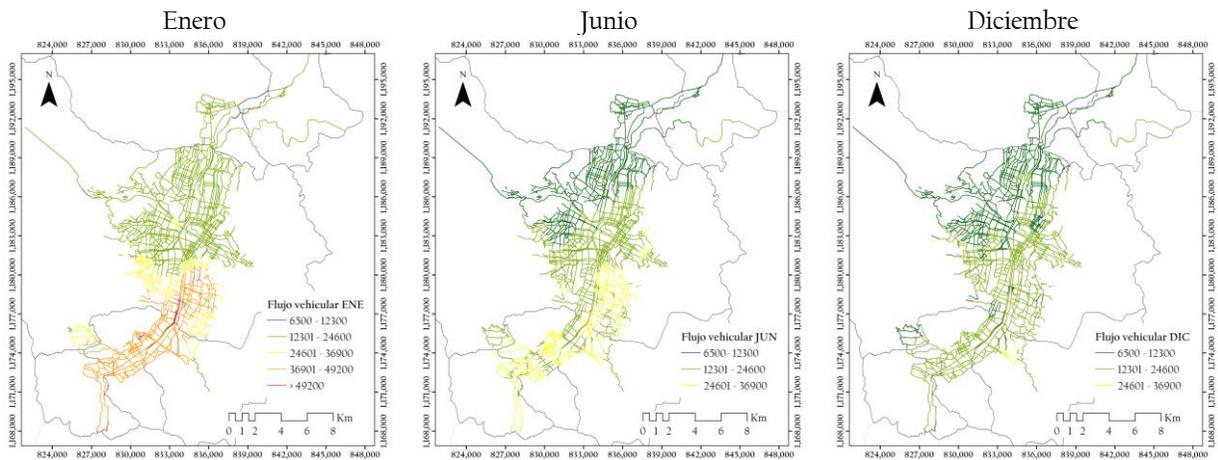


Figura 33. Flujos vehiculares estimados por el método Kriging para enero, junio y diciembre

6.2.3. Densidad poblacional

A partir de la información adquirida, se calculó la densidad poblacional por medio de una densidad de kernel. En la Figura 34 se puede observar que la mayor densidad se encuentra en los municipios de Bello y Envigado. Como en Medellín e Itagüí se pudo obtener información poblacional distribuida por barrio y comuna, respectivamente, se observa una mayor distribución espacial de los resultados de densidad poblacional.

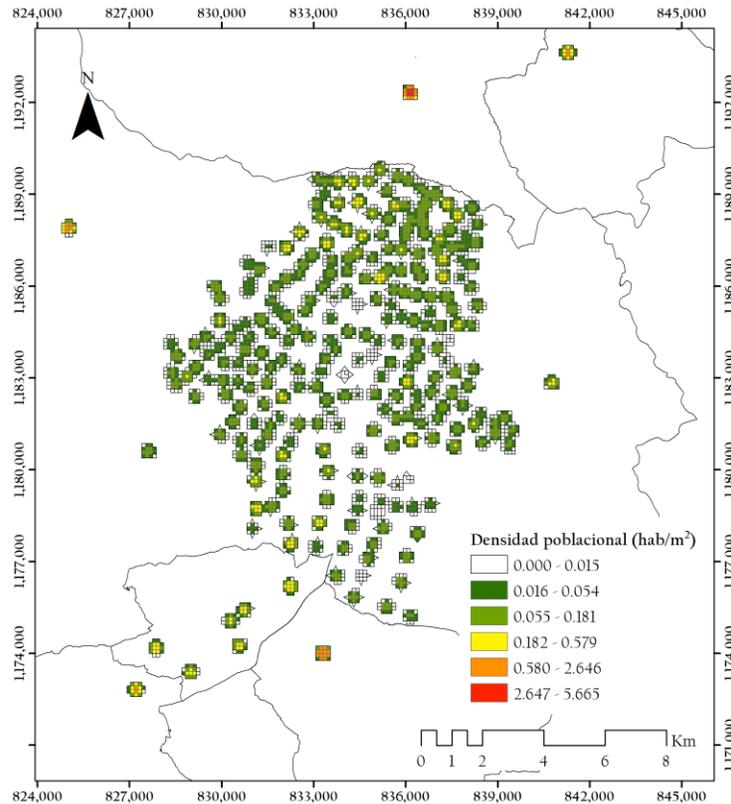


Figura 34. Densidad poblacional para la zona de estudio

6.2.4. Buffers e intersecciones espaciales

Con centro en las 13 estaciones de monitoreo de $PM_{2.5}$ seleccionadas (Figura 16), se construyeron buffers circulares de seis radios diferentes: 50, 100, 150, 200, 300 y 500 m. Para cada tamaño de buffer, se realizaron intersecciones espaciales con las variables de usos del suelo (Grupo 3), de movilidad (Grupo 4) y demográfica (Grupo 5). En la Figura 35 se muestran tres ejemplos de intersecciones espaciales. En la estación MED.UNNV, el buffer de 500 m se intersecó con todas las variables espaciales disponibles, sin embargo, para los buffers entre 50 y 150 m de radio, sólo hubo intersección con la variable de densidad poblacional. Por otro lado, la estación MED.AРАН está rodeada, en su mayoría, por uso residencial, variable que se presenta para todos los buffers, en los buffers mayores a 100 m de radio se pueden encontrar intersecciones con zonas verdes y vías secundarias. La estación ITA.CJUS también presenta una gran área de uso residencial alrededor, sin embargo, para el buffer más grande, se presentan usos industriales y zonas verdes. En general, las variables intersecadas difieren entre cada estación, debido a la heterogeneidad de las zonas circundantes a los puntos de monitoreo.

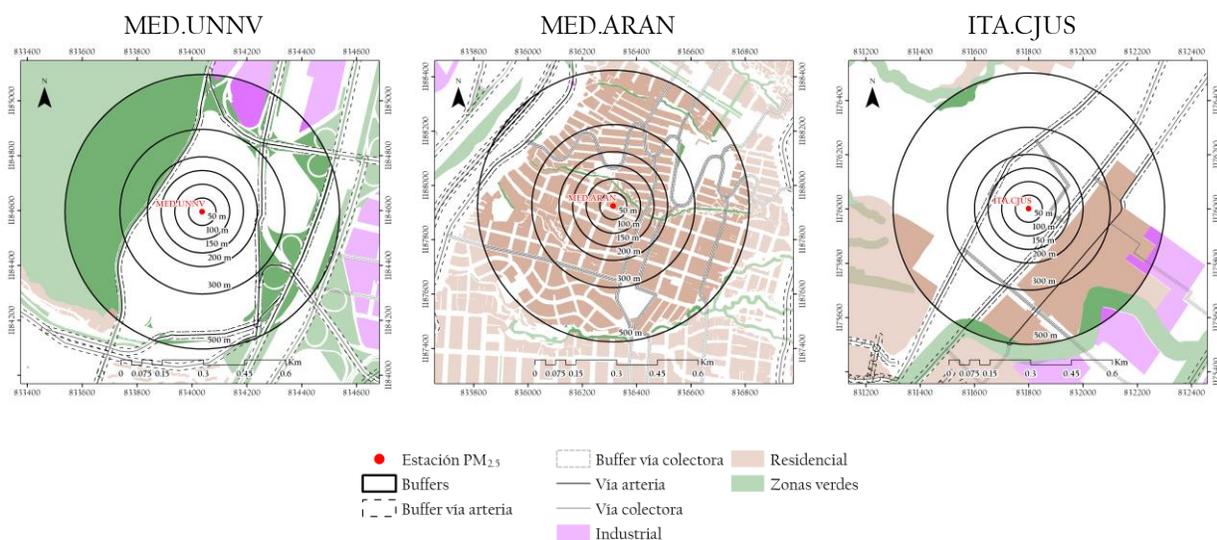


Figura 35. Intersecciones espaciales para los seis buffers evaluados, en tres estaciones de PM_{2.5}

6.3. Momento tres: Construcción de modelos LUR mensuales

Después de la intersección, las variables que tuvieran información igual a cero para más del 90% de las estaciones fueron excluidas del modelo de regresión espacial, dando como resultado que, para el buffer de 50 m, sólo se utilizaran las variables distancia y área de zonas verdes y de usos residenciales, además de las variables ambientales. Además, sólo se incluyó la distancia y área de usos industriales en el buffer de 500 m. Por otro lado, al evaluar la normalidad de todas las variables intersecadas y las variables mensuales, la mayoría no cumplió el criterio de Webster y Oliver, por la presencia de valores atípicos, los cuales no fueron removidos debido a que estos estaban asociados a diferentes puntos de monitoreo, lo que supondría una afectación en la construcción de los modelos LUR.

Para incluir una variable explicativa en la construcción del modelo LUR, ésta debía presentar una relación de dependencia con la variable dependiente. En la Tabla 33 se muestran las variables que no cumplieron este requisito, es decir, las variables explicativas excluidas de la construcción de los modelos de regresión espacial.

Tabla 33. Variables explicativas excluidas de la construcción de los modelos LUR

Mes	Variables descartadas*
1	DEN_POB50, DIST_VIA200, PISO
2	LONG_VIA100, NUM_VIA100, DIST_VIA200, PISO
3	DIST_VERDI100, PISO
4	DIST_VIA200, PISO
5	DIST_VIA200, PISO

Mes	VARIABLES DESCARTADAS*
10	Vv
12	DIST_VIA150

*El significado de las siglas en variables descartadas se encuentra en la Tabla 15

Aplicando la metodología descrita en la sección 5.5, se construyeron 72 modelos de regresión espacial (Anexo F), donde para el mes de febrero y buffer de 50 m de radio, no se encontró algún modelo que cumpliera con los criterios establecidos. En la Tabla 34 se muestran las frecuencias de selección de las variables explicativas en estos modelos. Se puede observar que las variables más frecuentes en los modelos LUR fueron PISO para las variables de posición (Grupo 1); Vv para las ambientales (Grupo 2); para las variables de usos del suelo (Grupo 3), AREA_RESID; y para las variables de movilidad (Grupo 4), fue FLUJO_VEHI. Con los 71 modelos construidos, se explica entre el 19 y 79% de la variabilidad del PM_{2.5}, y en promedio, el 49%. Se encontraron modelos con una a cinco variables explicativas. En general, al aumentar el número de variables se presentaba un incremento en el AdjR² y el valor p del estadístico Jarque Bera (JB), es decir que se mejoraba en el rendimiento y consistencia del modelo, sin embargo, al mismo tiempo, se incrementaba el VIF y el AICc, indicando una reducción en la calidad del modelo.

Tabla 34. Frecuencias de las variables explicativas en los modelos LUR

Variable explicativa	Frecuencia absoluta	Frecuencia relativa
PISO	33	0.16
V	31	0.15
AREA_RESID	17	0.08
FLUJO_VEHI	17	0.08
DIST_RESID	16	0.08
DIST_VERD	15	0.07
GRADT	14	0.07
GRADP	11	0.05
AREA_VERD	11	0.05
AREA_VIA	11	0.05
P	8	0.04
DEN_POB	8	0.04
LONG_VIA	5	0.02
NUM_VIA	5	0.02
T	4	0.02
GRADV	4	0.02
ALTURA	2	0.01
Total	212	1.00

6.4. Momento cuatro: Selección de los mejores modelos LUR mensuales

Al evaluar los criterios de especificación del método OLS (Tabla 16), se determinaron las mejores combinaciones de variables para cada mes, expuestos en la Tabla 35. Los modelos presentaron entre dos y cuatro variables, donde sólo dos parejas de modelos son explicadas por las mismas variables; sin embargo, para todos los meses se presentan algunas en común, donde la altura a nivel de piso de la estación de monitoreo de PM_{2.5}, el flujo vehicular, el gradiente de temperatura y la velocidad del viento son las variables más frecuentes. El tamaño de buffer más frecuente entre los 12 modelos fue el de 200 m, donde no se seleccionó ningún modelo para un buffer de 50 m. Al evaluar los valores del criterio de rendimiento (AdjR²), se observa que los modelos seleccionados explican entre 26 y 79% de la variabilidad del PM_{2.5}, con un valor promedio del 48%. El modelo que mejor explica la variabilidad del PM_{2.5} es el del mes de agosto, seguido por diciembre.

Tabla 35. Criterios de especificación OLS para los mejores modelos LUR mensuales

Mes	Buffer	AdjR ²	AICc	JB	K(BP)	VIF	SA	Variables explicativas
1	200	0.26	71.65	0.79	0.13	1.30	0.39	- GRADT, GRADVv
2	200	0.33	74.76	0.69	0.33	1.25	0.55	-GRADT, FLUJO_VEHI
3 ¹	200	0.31	-35.07	0.80	0.34	1.16	0.54	-GRADT, FLUJO_VEHI
4	200	0.42	71.44	0.86	0.75	1.30	0.64	-GRADT, -AREA_VERD, FLUJO_VEHI
5 ¹	100	0.49	-16.83	0.56	0.37	1.15	0.55	-Vv, GRADVv, DIST_VERD
6	500	0.54	62.9	0.97	0.31	1.05	0.14	-PISO, -Vv, +AREA_VIA
7	150	0.47	60.08	0.8	0.11	1.05	0.83	-Vv, -DIST_RESID
8 ¹	300	0.79	-29.78	0.30	0.08	1.41	0.76	-PISO, -Vv, -DIST_RESID, +FLUJO_VEHI
9 ²	100	0.58	9.2	0.77	0.07	1.42	0.59	-PISO, -Vv, +AREA_RESID
10	500	0.5	64.34	0.74	0.52	1	0.18	-PISO, LONG_VIA
11	200	0.46	67.92	0.74	0.10	1.09	0.56	-PISO, -Vv
12	200	0.62	62.93	0.85	0.15	1.70	0.62	-PISO, -Vv, AREA_RESID

¹Logaritmo de los datos espaciales de PM_{2.5}

²Raíz de los datos espaciales de PM_{2.5}

6.4.1. Modelos LUR mensuales

Al aplicar el procedimiento descrito en la sección 5.6, se encontraron los coeficientes de los mejores modelos LUR para PM_{2.5} para las 13 estaciones de monitoreo seleccionadas, dados por las Ecuaciones 15 a 26.

– Enero

La Ecuación 15 muestra el modelo seleccionado para el mes de enero, el cual explica el 26% de la variabilidad del PM_{2.5} en la zona de estudio. Este modelo está en función de los gradientes de temperatura y velocidad del viento promedio mensual. La variable dependiente tiene una relación

positiva con $GRADV_v$, indicando que, al aumentar la magnitud de esta variable, se incrementan los niveles de $PM_{2.5}$, comportamiento contrario a la otra variable.

$$PM_{2.5_{ENE}} = 19.0136 - 45.5418 * GRADT + 159.3971 * GRADV_v \quad (15)$$

– Febrero

La Ecuación 16 muestra el modelo seleccionado para el mes de febrero, el cual explica el 33% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 36 muestra las variables espaciales escogidas para un buffer de 200 m. Este modelo está en función del gradiente de temperatura promedio mensual y el flujo vehicular promedio mensual de la vía más cercana. La variable dependiente tiene una relación positiva con la variable de movilidad, indicando que, al aumentar la magnitud de ésta, se incrementan los niveles de $PM_{2.5}$, comportamiento contrario a la variable ambiental.

$$PM_{2.5_{FEB}} = 25.6637 - 50.8231 * GRADT + 0.0003 * FLUJO_VEHI_{200} \quad (16)$$

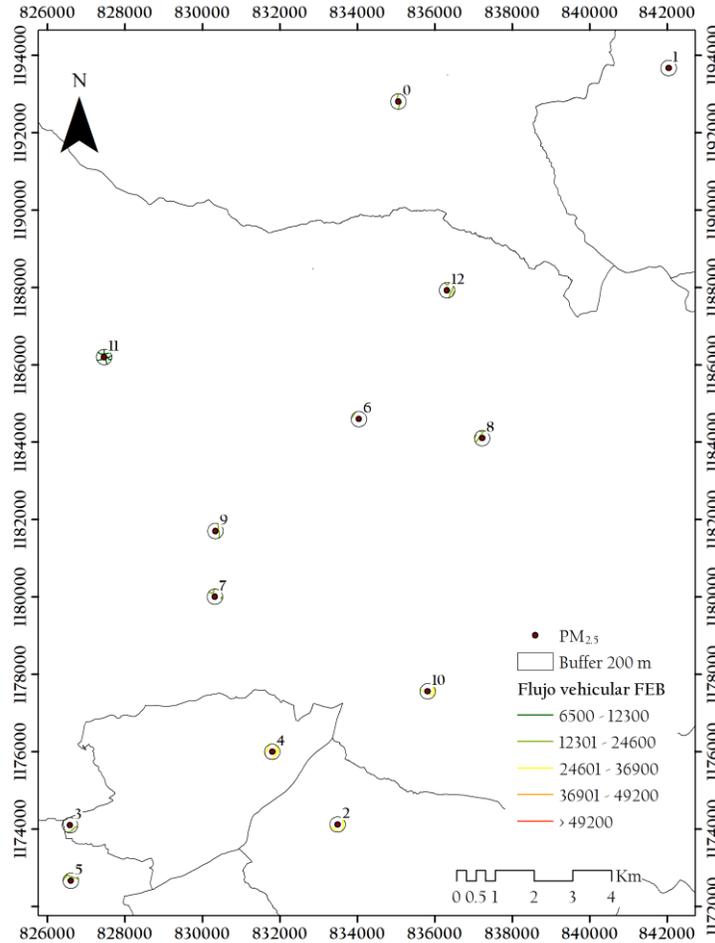


Figura 36. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de febrero

– Marzo

La Ecuación 17 muestra el modelo seleccionado para el mes de marzo, el cual explica el 31% de la variabilidad del $PM_{2.5}$ en la zona de estudio. Al igual que el modelo de febrero, para marzo, la ecuación resultante se encuentra en función del gradiente de temperatura promedio mensual y el flujo vehicular promedio mensual de la vía más cercana, para un buffer de 200 m (Figura 37).

$$\log(PM_{2.5_{MAR}}) = 1.4268 - 0.6259 * GRADT + 0.00001 * FLUJO_VEH1_{200} \quad (17)$$

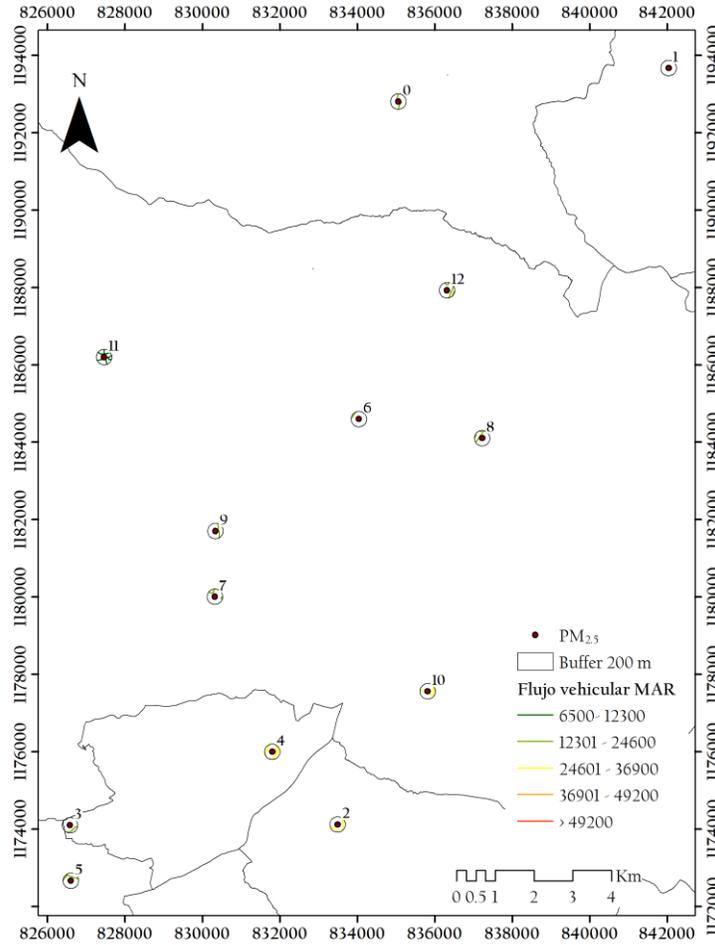


Figura 37. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de marzo

– Abril

La Ecuación 18 muestra el modelo seleccionado para el mes de abril, el cual explica el 64% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 38 muestra las variables espaciales escogidas para un buffer de 200 m. Este modelo está en función del gradiente de temperatura promedio mensual, el área de zonas verdes y el flujo vehicular promedio mensual de la vía más cercana. La variable dependiente tiene una relación positiva con la variable de movilidad, indicando que, al aumentar la magnitud de ésta, se incrementan los niveles de $PM_{2.5}$, comportamiento contrario a las otras dos variables.

$$PM_{2.5_{ABR}} = 17.9654 - 52.3663 * GRADT - 0.0001 * AREA_VERD_{200} + 0.0002 * FLUJO_VEHI_{200} \quad (18)$$

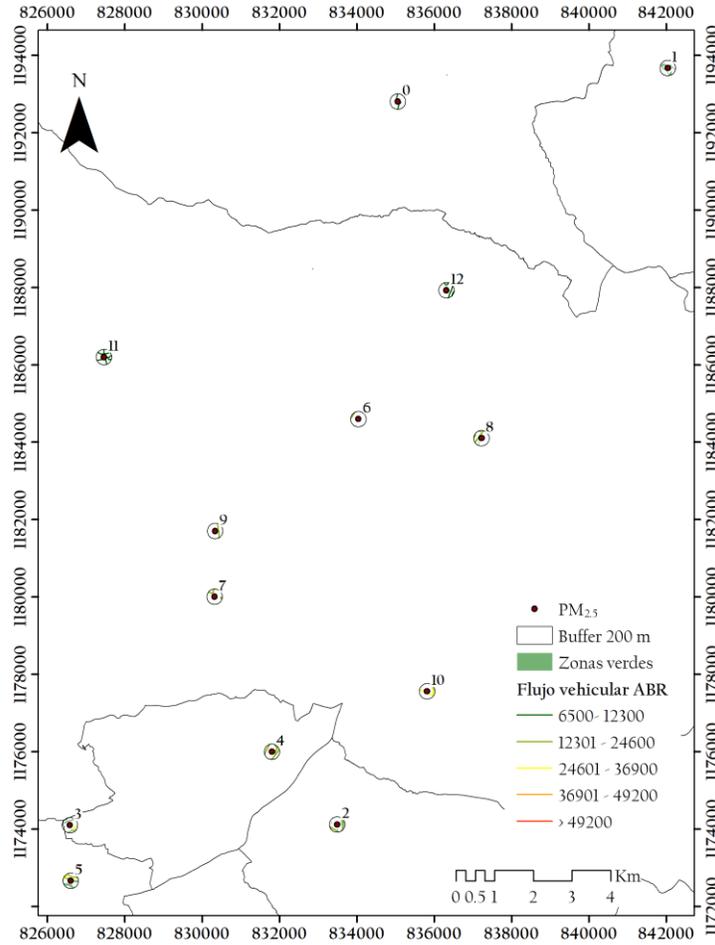


Figura 38. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de abril

– Mayo

La Ecuación 19 muestra el modelo seleccionado para el mes de mayo, el cual explica el 55% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 39 muestra las variables espaciales escogidas, donde este modelo se encuentra en función la precipitación y el área de zonas verdes, en un buffer de 100 m. La variable dependiente tiene una relación positiva con el gradiente de velocidad del viento y la distancia de zonas verdes, indicando que, al aumentar la magnitud de éstas, se incrementan los niveles de $PM_{2.5}$, comportamiento contrario a la velocidad del viento.

$$\log(PM_{2.5_{MAY}}) = 1.3562 - 0.1529 * Vv + 4.6475 * GRADVv + 0.0027 * DIST_VERD_{100} \quad (19)$$

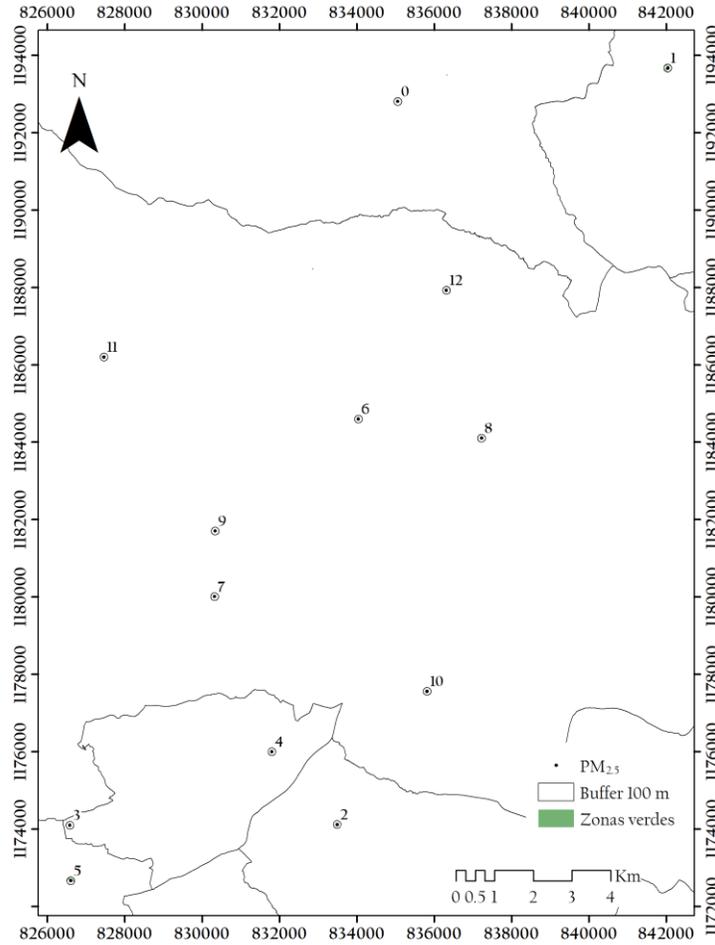


Figura 39. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de mayo

– Junio

La Ecuación 20 muestra el modelo seleccionado para el mes de junio, el cual explica el 54% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 40 muestra las variables espaciales escogidas para un buffer de 500 m. Este modelo está en función de los gradientes de precipitación acumulada mensual y temperatura promedio mensual, la distancia a zona verde, densidad poblacional y el área de las vías.

$$PM_{2.5JUN} = 18.2699 - 0.2871 * PISO - 2.4789 * V + 0.00003 * AREA_VIA_{500} \quad (20)$$

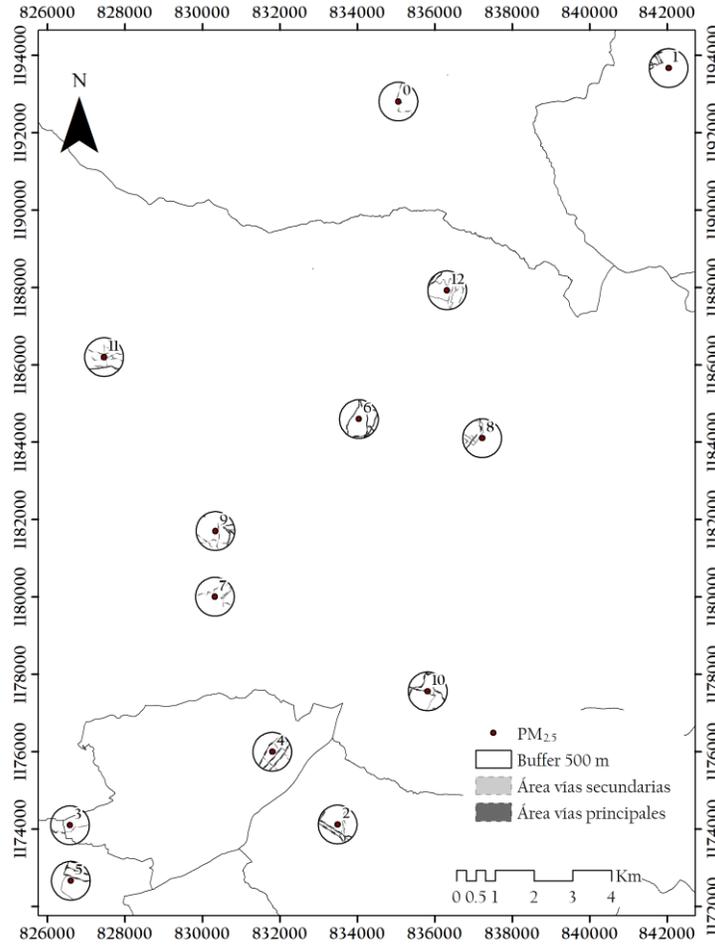


Figura 40. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de junio

- Julio

La Ecuación 21 muestra el modelo seleccionado para el mes de julio, el cual explica el 47% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 41 muestra las variables espaciales escogidas para un buffer de 150 m. Este modelo está en función de la velocidad del viento promedio mensual y la distancia a uso residencial, ambas con una relación inversa al contaminante.

$$PM_{2.5_{JUL}} = 22.3229 - 3.7425 * Vv - 0.0313 * DIST_RESID_{150} \quad (21)$$

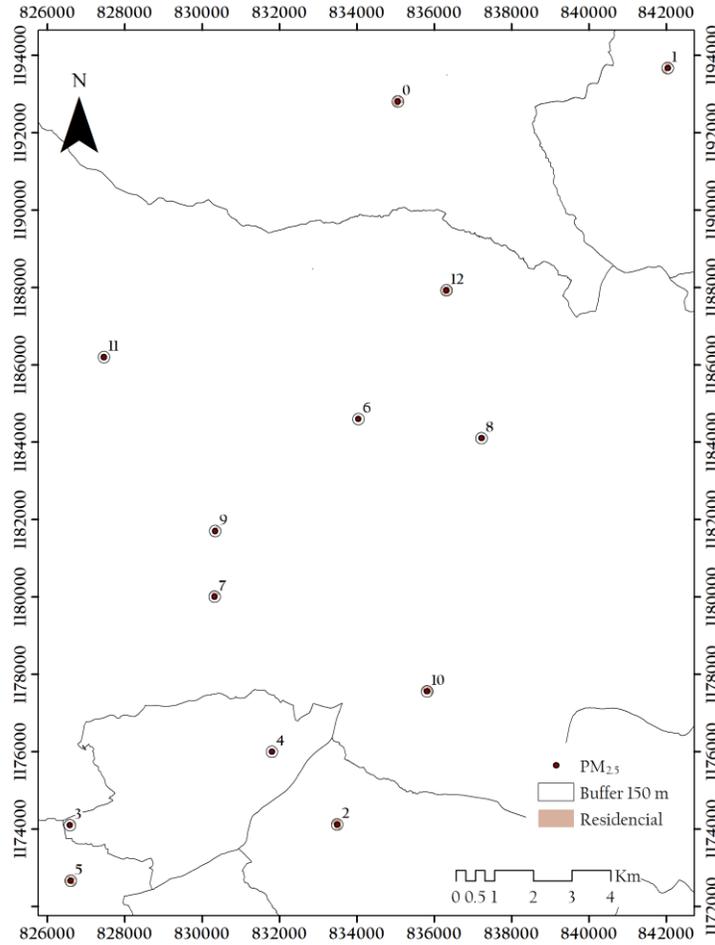


Figura 41. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de julio

– Agosto

La Ecuación 22 muestra el modelo seleccionado para el mes de agosto, el cual explica el 79% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 42 muestra las variables espaciales escogidas para un buffer de 300 m. Este modelo está en función de la altura a nivel de piso de la estación de monitoreo, la velocidad del viento promedio mensual, la distancia a uso residencial y el flujo vehicular promedio mensual de la vía más cercana. El contaminante presenta una relación inversa con las variables PISO, Vv y DIST_RESID, y directa con FLUJO_VEHl.

$$\log(PM_{2.5_{AGO}}) = 1.3676 - 0.0130 * PISO - 0.0957 * Vv - 0.0007 * DIST_RESID_{300} + 0.00001 * FLUJO_VEHl_{300} \quad (22)$$

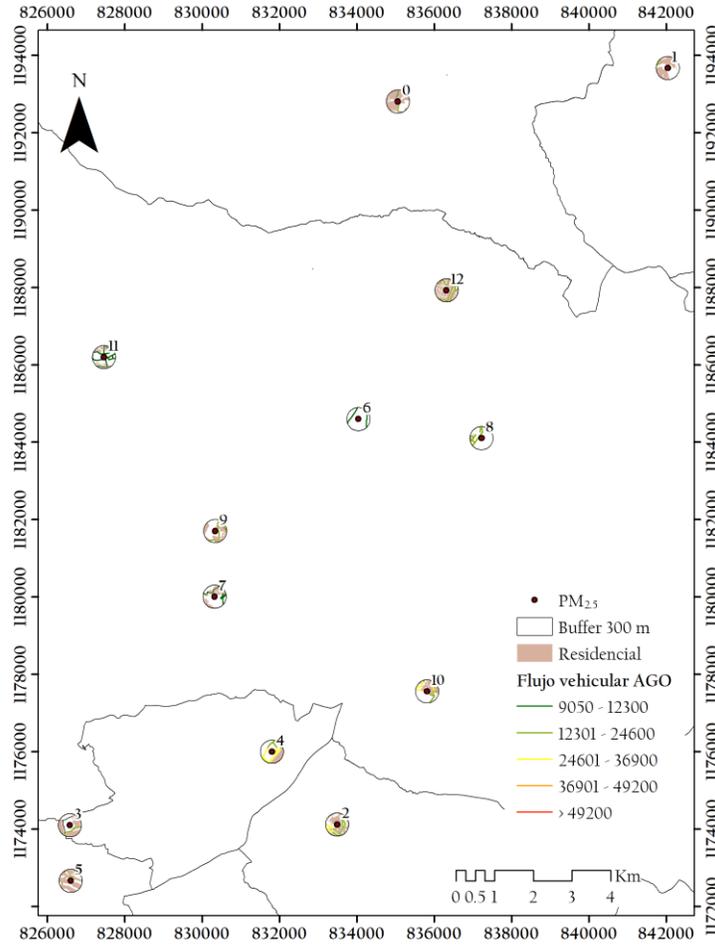


Figura 42. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de agosto

– Septiembre

La Ecuación 23 muestra el modelo seleccionado para el mes de septiembre, el cual explica el 58% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 43 muestra las variables espaciales escogidas para un buffer de 100 m. Este modelo está en función de la altura a nivel de piso de la estación de monitoreo, la velocidad del viento promedio mensual y el área del uso residencial, presentándose una relación directa entre la variable de uso del suelo y el contaminante.

$$\sqrt{PM_{2.5_{SEP}}} = 5.2618 - 0.0511 * PISO - 0.5352 * Vv + 0.0002 * AREA_RESID_{100} \quad (23)$$

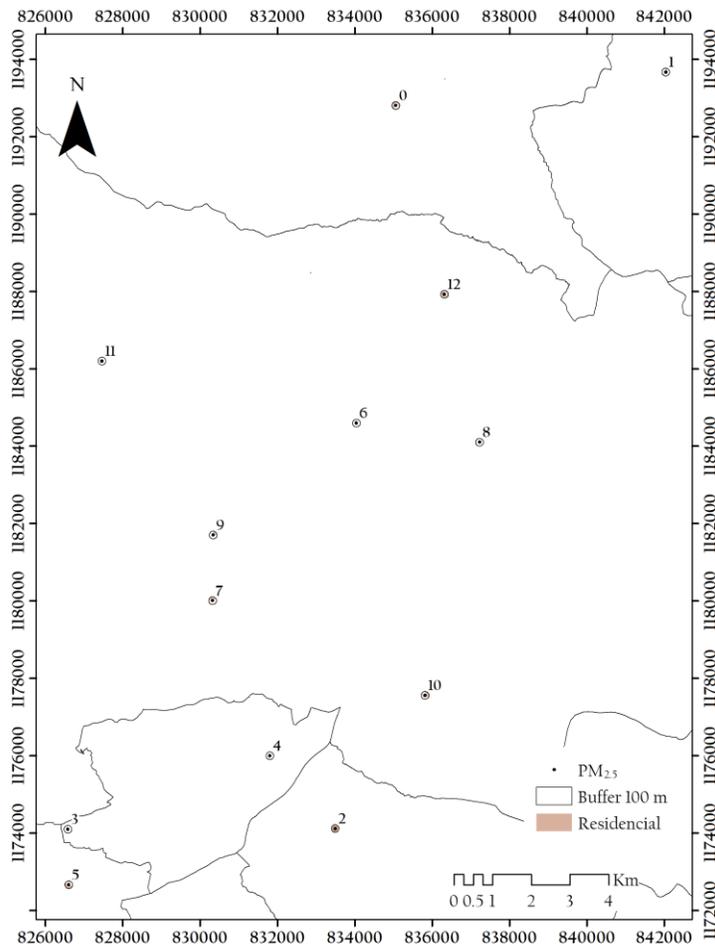


Figura 43. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de septiembre

– Octubre

La Ecuación 24 muestra el modelo seleccionado para el mes de octubre, el cual explica el 50% de la variabilidad del PM_{2.5} en la zona de estudio. La Figura 44 muestra las variables espaciales escogidas para un buffer de 500 m. Este modelo está en función de la altura a nivel de piso de la estación de monitoreo y la longitud de las vías intersecadas en el buffer, presentándose una relación directa entre la variable de movilidad y el contaminante.

$$PM_{2.5_{OCT}} = 16.2774 - 0.4041 * PISO + 0.0010 * LONG_VIA_{500} \quad (24)$$

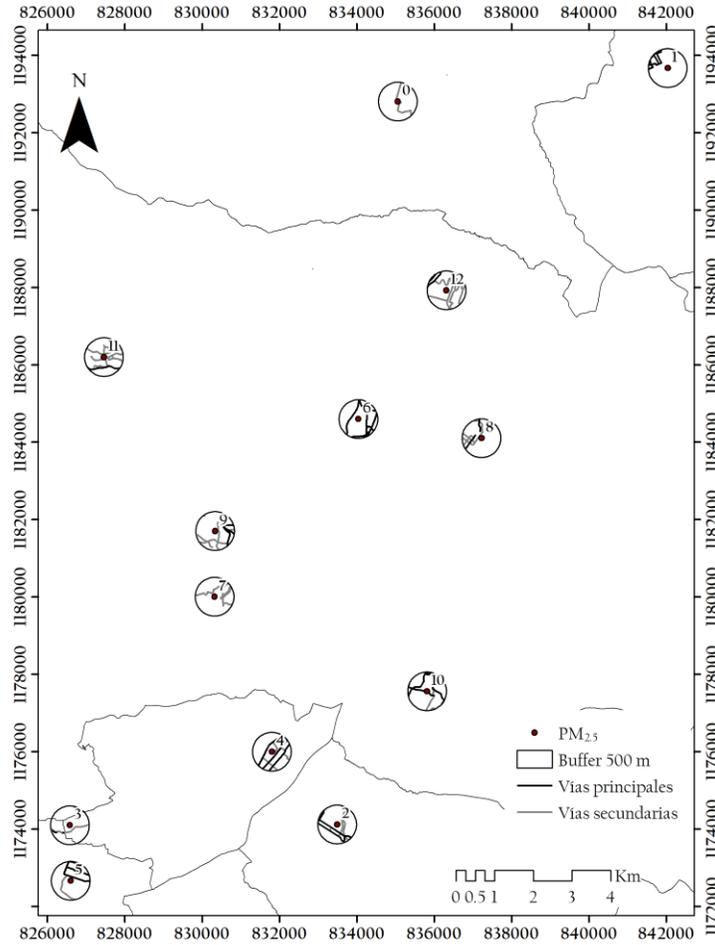


Figura 44. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de octubre

– Noviembre

La Ecuación 25 muestra el modelo seleccionado para el mes de noviembre, el cual explica el 46% de la variabilidad del $PM_{2.5}$ en la zona de estudio. Se escogió el modelo para el buffer de 200 m, estando en función de la altura a nivel de piso de la estación de monitoreo, la velocidad del viento promedio mensual.

$$PM_{2.5_{NOV}} = 27.5662 - 0.3708 * PISO - 4.6895 * V \quad (25)$$

– Diciembre

La Ecuación 26 muestra el modelo seleccionado para el mes de diciembre, el cual explica el 62% de la variabilidad del $PM_{2.5}$ en la zona de estudio. La Figura 45 muestra las variables espaciales escogidas para un buffer de 200 m. Este modelo está en función de la altura a nivel de piso de la

estación de monitoreo, la velocidad del viento promedio mensual y el área de uso residencial, presentándose una relación directa entre la variable de uso del suelo y el contaminante.

$$PM_{2.5} = 28.9382 - 0.5539 * PISO - 4.1656 * Vv + 0.0001 * AREA_RESID_{200} \quad (26)$$

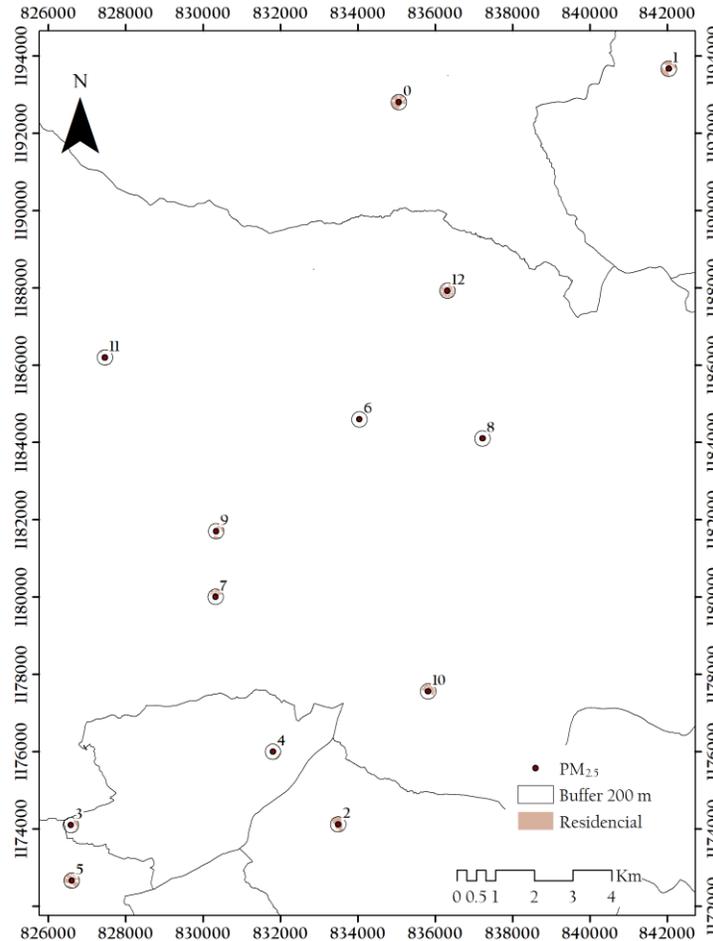


Figura 45. Mapa de las variables explicativas espaciales seleccionadas para el mejor modelo de diciembre

6.4.2. Mapas de predicción y validación de modelos LUR

Los mapas de regresión fueron obtenidos a partir de la ejecución del modelo LUR en una malla de puntos definida y aplicando un método de interpolación con los valores estimados en estos puntos. Evaluando el error cuadrático medio de los algoritmos de interpolación, en la Tabla 36 se observa que el mejor interpolador para la mayoría de los meses fue el algoritmo Spline, con un RMSE promedio de 1.50; sin embargo, para los meses de febrero, marzo y octubre, el IDW presentó el

mejor rendimiento para los datos de PM_{2.5} estimados, con un RMSE promedio de 1.11, indicando la aplicación de diferentes interpoladores a lo largo del año para los valores estimados de PM_{2.5}, resultado de la ejecución de los modelos LUR mensuales.

Tabla 36. RMSE de las interpolaciones espaciales de los valores estimados de PM_{2.5} con los modelos LUR

Mes	RMSE	
	IDW	Spline
Enero	0.54	0.51
Febrero	1.73	1.78
Marzo	1.23	1.25
Abril	2.98	2.56
Mayo	3.26	3.23
Junio	0.63	0.56
Julio	1.40	1.26
Agosto	2.67	2.43
Septiembre	1.23	0.96
Octubre	0.37	0.36
Noviembre	1.02	0.84
Diciembre	1.60	1.11

En la Figura 46 y Figura 47 se muestran los ejemplos de mapas de regresión de PM_{2.5} para cada mes, donde, en los meses de junio y octubre no se evidencia la misma variación espacial, dado que la zona de influencia que permitió obtener algún efecto de las variables explicativas en el modelo, fue de radio 500 m y, por ende, se utilizó una malla de 1000m x 1000m. Aunque no todos los modelos explicaron más del 50% de la variabilidad espacial del PM_{2.5}, la mayoría de los mapas muestran detalles locales considerables.

Al evaluar las concentraciones máximas estimadas, se pueden observar meses donde los niveles del contaminante son altos en comparación a los otros momentos del año, y otros en donde son mucho menores. Los meses de febrero y marzo, cuyos modelos fueron definidos por las mismas variables (GRADT y FLUJO_VEHI) presentan los momentos de mayor contaminación por PM_{2.5} en Medellín. Por otra parte, el mes de junio presenta los niveles más bajos del contaminante.

El modelo para el mes de abril incluyó una variable que representa la fuente de emisión del contaminante, una variable de remoción y otra de dispersión. Aunque la explicabilidad del modelo fue pequeño, permite visualizar la influencia de la cercanía a las vías y a las zonas verdes en los niveles del contaminante. Así mismo, en el modelo para el mes de mayo se incluyó la influencia de la cercanía a las zonas verdes, además de las variables relacionadas a la dirección del viento, con un buffer de menor tamaño, evidenciando la influencia de estas variables en los niveles del contaminante, y manteniendo cierto grado de similitud en los patrones, principalmente, en los niveles altos encontrados en el sur y sur-occidente de la zona de estudio.

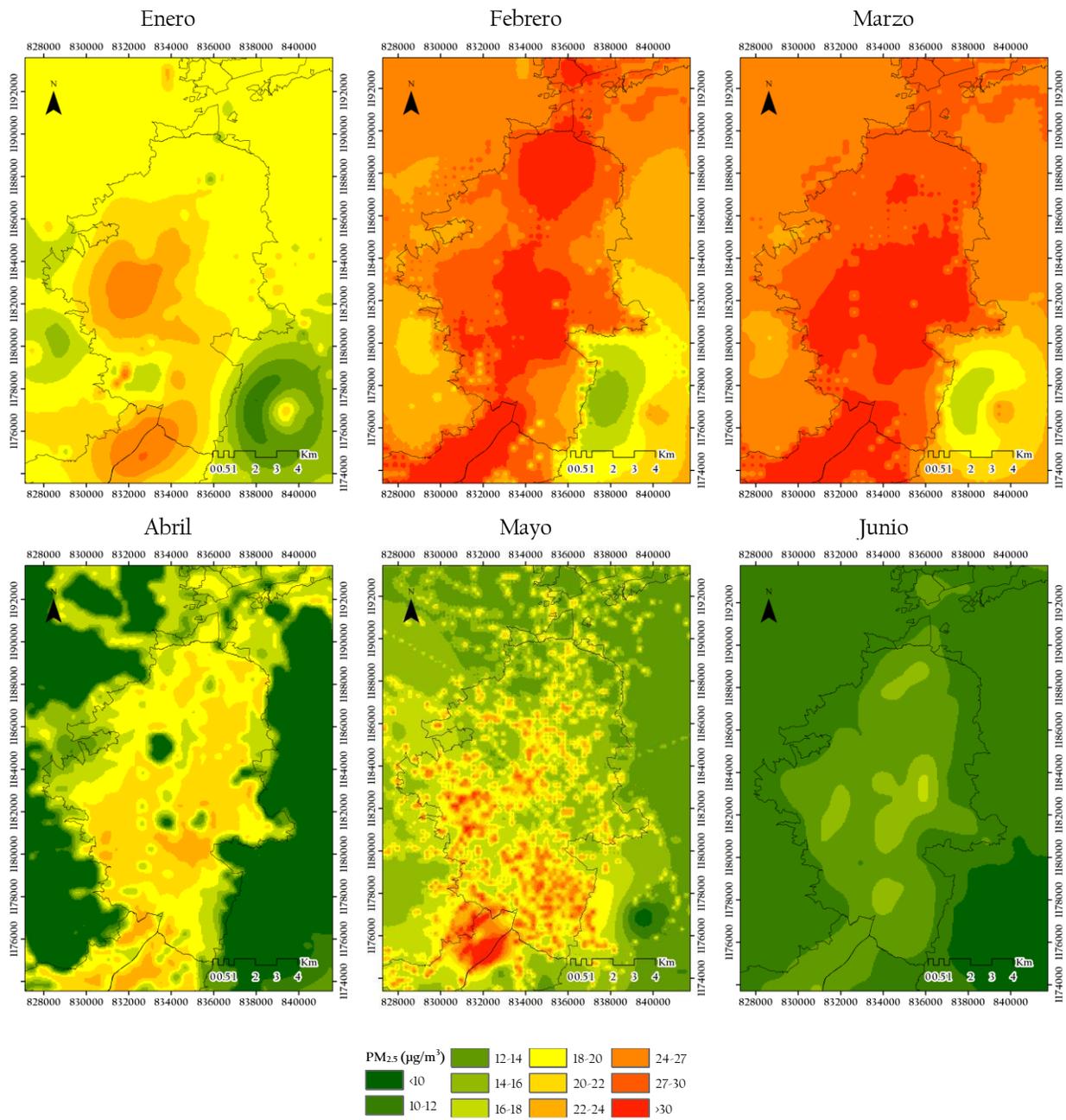


Figura 46. Mapas de regresión de PM_{2.5} promedio mensual, para el primer semestre del año

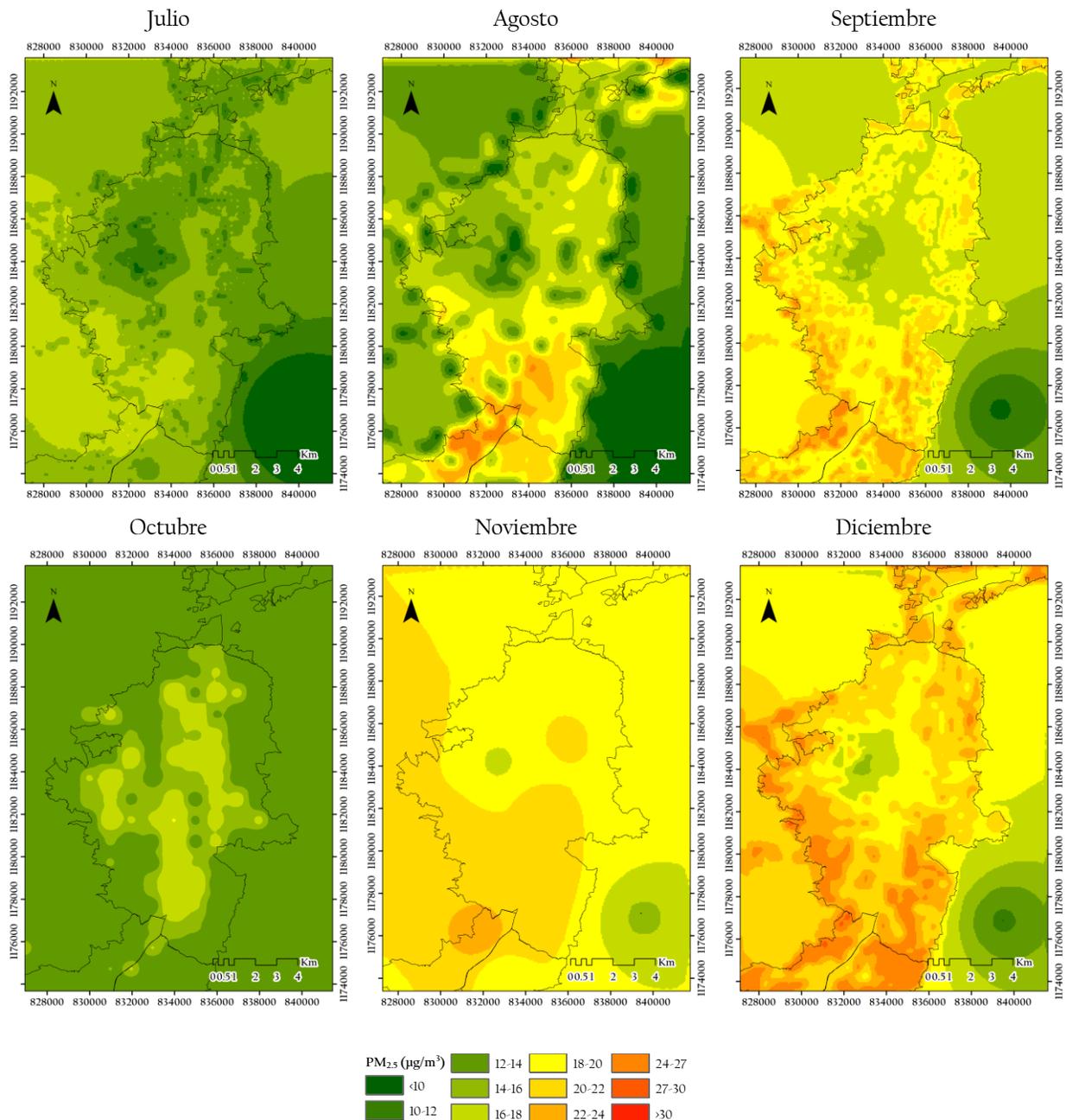


Figura 47. Mapas de regresión de $PM_{2.5}$ promedio mensual, para el segundo semestre del año

En la Figura 48 se muestran las pendientes de las relaciones entre las concentraciones promedio mensual observadas y las estimadas, en cada estación de monitoreo. Al ser el modelo con el mejor rendimiento, la pendiente para los datos espaciales del mes de agosto (línea roja) es la más cercana a la unidad (línea punteada), donde se observa una leve tendencia de sobreestimación de las concentraciones más bajas. Por otro lado, los modelos con menor rendimiento, cuya pendiente se

aleja de la unidad, presentan una tendencia a subestimar las concentraciones más altas del contaminante.

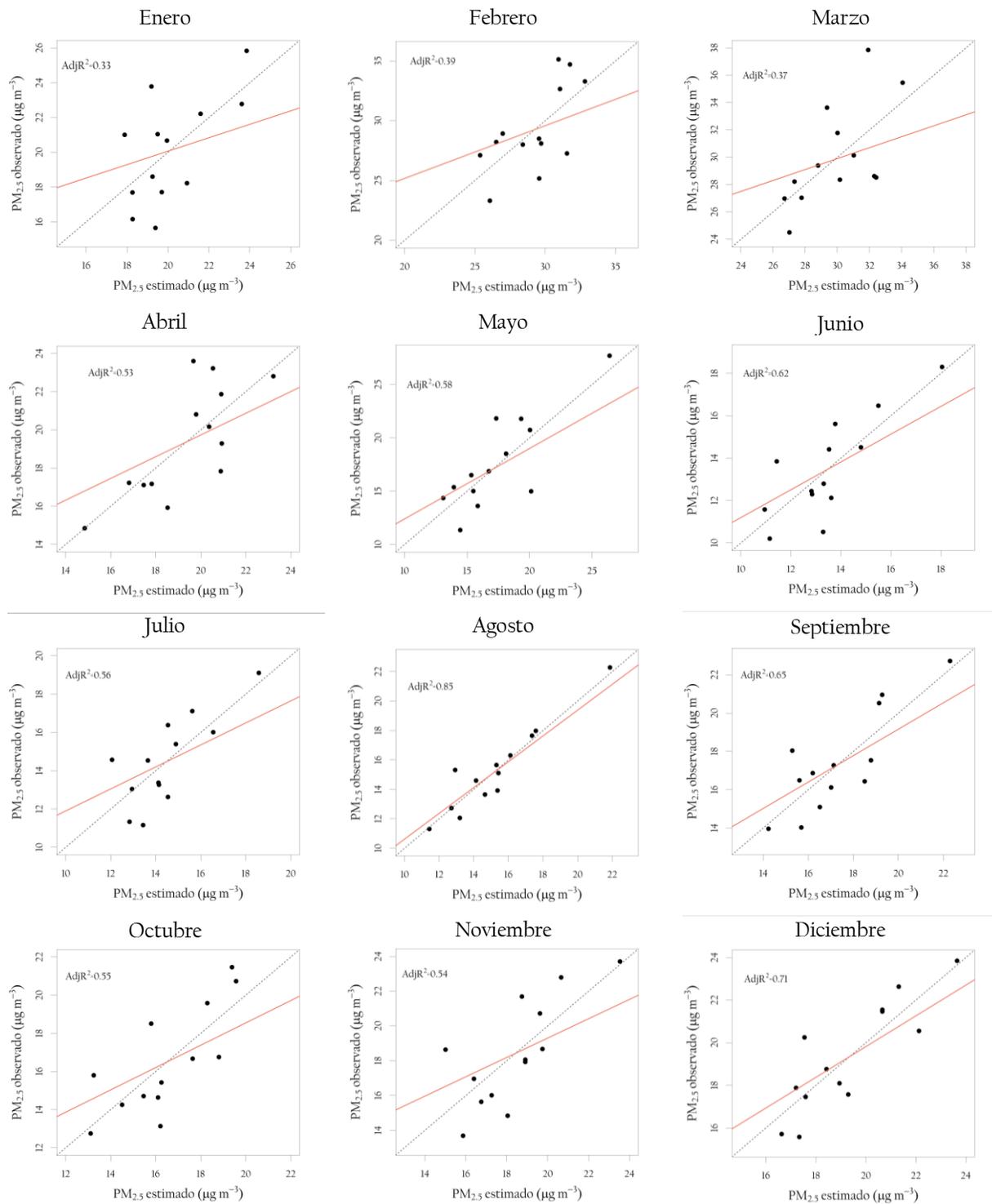


Figura 48. Concentraciones de $PM_{2.5}$ estimadas de los mapas de regresión y valores promedio mensuales observados

En la Tabla 37 se muestran los estadísticos de error de los datos estimados de $PM_{2.5}$ para cada modelo LUR mensual construido, en la Tabla 38 el estadístico de error MARE para cada sitio de monitoreo y en el Anexo G se muestra el estadístico de error ARE para cada sitio de monitoreo y cada mes, con los cuales se calculó el MARE para cada mes. Al evaluar este estadístico de error, el menor rendimiento se presentó para el mes de enero, lo cual es congruente dado que es el modelo que menos explica la variabilidad de la variable estudiada; sin embargo, al evaluar el RMSS, los modelos de los meses de enero, febrero y marzo presentaron una menor capacidad para reestimar un dato observado. Por otro lado, todos los sitios de monitoreo presentaron un MARE menor a 0.11, siendo la estación MED.SCRI (ID=11) la del menor rendimiento, es decir que, en general, la diferencia entre el valor estimado y el valor observado era mayor en este punto, contrario a ITA.CJUS que presentó el mejor rendimiento (ID=4).

Tabla 37. Estadísticos de error estimados de $PM_{2.5}$ para cada modelo mensual

Mes	ASE	RMSE	RMSS	MARE
1	1.94	2.31	0.75	0.10
2	2.24	2.60	0.72	0.08
3	0.03	0.04	0.73	0.02
4	1.38	1.85	0.63	0.07
5	0.05	0.06	0.59	0.04
6	1.07	1.33	0.56	0.08
7	1.21	1.41	0.61	0.09
8	0.02	0.03	0.36	0.02
9	0.14	0.17	0.54	0.03
10	1.51	1.74	0.62	0.09
11	1.63	1.94	0.63	0.09
12	1.07	1.28	0.49	0.06

Tabla 38. Estadísticos de error para cada sitio de monitoreo

ID estación de monitoreo	ASE	RMSE	RMSS	MARE
0	1.08	1.64	0.21	0.07
1	1.06	1.55	0.20	0.09
2	0.71	1.26	0.14	0.06
3	1.04	1.43	0.14	0.06
4	0.38	0.67	0.06	0.02
5	1.40	2.01	0.23	0.10
6	0.98	1.49	0.14	0.05
7	1.04	1.40	0.15	0.06
8	0.75	1.05	0.11	0.05
9	0.71	1.02	0.10	0.04
10	0.47	0.70	0.08	0.04

ID estación de monitoreo	ASE	RMSE	RMSS	MARE
11	1.07	1.82	0.25	0.11
12	1.08	2.11	0.20	0.08

7. Discusión

Actualmente, la exposición a diferentes contaminantes atmosféricos es un problema ambiental de gran interés por sus posibles afectaciones a la salud, donde la magnitud de ésta se encuentra en función de diferentes factores presentes en la zona. El Valle de Aburrá está caracterizado por ser un valle estrecho, densamente urbanizado, donde la topografía de la zona da lugar a condiciones de estabilidad atmosférica, particularmente de inversión térmica, inhibiendo procesos de mezcla en la vertical (Wallace & Hobbs, 2006; Zapata et al., 2015). Lo anterior, en conjunto con la existencia de fuentes de emisión de diferentes contaminantes, especialmente del $PM_{2.5}$, deterioran la calidad del aire en ciertos periodos del año, donde las condiciones meteorológicas y la estacionalidad climática de la región, propician eventos de inversión persistentes en el tiempo.

Las diferentes series de $PM_{2.5}$, evaluadas para el año 2018, presentaron la estacionalidad tan característica de este contaminante en el Valle de Aburrá, donde se esperan incrementos en los niveles durante el mes de marzo (Área Metropolitana del Valle de Aburrá, 2018; Londoño, 2018a). Estos se encuentran relacionados al régimen bimodal de lluvias en Colombia, que durante los meses con altas precipitaciones (marzo-mayo y septiembre-noviembre) se presentan condiciones que favorecen una atmósfera estable en el Valle de Aburrá, limitando la dispersión de contaminantes en la atmósfera e incrementando las concentraciones de los mismos. Durante el periodo de análisis, se presentaron días excedencias a la norma nacional, que en comparación con años anteriores, permiten evidenciar una reducción en las mismas (Área Metropolitana del Valle de Aburrá, 2018).

Al evaluar las demás variables ambientales utilizadas en este estudio, se encuentra que el mejor interpolador para los datos de temperatura fue el IDW, caso contrario al presentado en el estudio realizado por Londoño (2018), donde el mejor método fue el Spline. Por otro lado, se encuentra un leve desplazamiento de los valores máximos de la temperatura promedio mensual para el año 2018, donde se presentó un máximo en marzo y otro en agosto en 2018, mientras que, en el otro estudio, los máximos se presentaron en abril y julio en los años 2013 y 2014. Sin embargo, para ambos estudios, se redujo el error de interpolación de la velocidad del viento con el método Spline y la variable mantiene un comportamiento similar, con magnitudes bajas en mayo y altas en julio, presentándose leves variaciones en las magnitudes de los datos. Esto puede estar atribuido, además de la diferencia en el periodo de análisis y la posible influencia de El Niño Oscilación del Sur (ENSO), a las estaciones meteorológicas seleccionadas, las cuales difieren en ambos estudios, implicando la importancia de las estaciones de la red de monitoreo, la cual puede reducir o aumentar la incertidumbre en las interpolaciones.

En relación con los aforos vehiculares, los datos obtenidos por las cámaras de fotodetección ubicadas en el municipio de Medellín presentaron información valiosa sobre el flujo vehicular diario en la zona de interés. Para algunos sitios de monitoreo se observó un decrecimiento atípico en la magnitud de los aforos en el último trimestre del año, probablemente asociado con la

ampliación de la estación Metro El Poblado, obras viales, de acueducto y alcantarillado, ejecutadas durante el año 2018¹, en el municipio de Medellín.

Al igual que en otros estudios, el Kriging tuvo un mejor desempeño para interpolar los valores de flujo vehicular que otros métodos evaluados, siempre que los datos fueran estacionarios (Sarlas & Axhausen, 2016; X. Wang & Kockelman, 2009; D. Zhang & Wang, 2013). En este trabajo se evaluaron otros modelos de semivariograma teóricos, en comparación a lo encontrado en la literatura, en donde es común que se evalúen los modelos exponencial, gaussiano y esférico. A diferencia de otros estudios, el mejor modelo de semivariograma ajustado para los datos de aforos vehiculares promedio mensual fue el k-bessel, aunque para algunos meses se mejoró la precisión con el modelo exponencial o esférico. Para el mes de febrero, no se pudo aplicar la metodología Kriging debido a la no estacionariedad de los datos espaciales, asociada a la presencia de valores atípicos, relacionados al bajo flujo vehicular en un sitio de monitoreo, asociado a obras viales.

Al analizar los buffers seleccionados para el desarrollo de esta metodología, es evidenciable que, para la zona de estudio, el buffer de 50 m no presentó el mejor rendimiento en ninguno de los meses, debido a la ubicación de las estaciones de monitoreo de PM_{2.5}, dificultando la intersección con alguna variable espacial. Por otra parte, el buffer más común entre los doce modelos LUR finales fue el de 200 m de radio, indicando que los niveles del contaminante se encuentran influenciados por la presencia de vías a menos de 200 m a la redonda, debido a que el flujo vehicular a la vía más cercana fue la variable espacial más común para estos meses.

En total, una variable de movilidad fue significativa en el modelo de regresión final para siete meses, indicando que la variabilidad del PM_{2.5} para la mayoría del tiempo, es explicada por el parque automotor, de forma directa o indirecta. Este comportamiento se corrobora con diferentes estudios en donde las variables de tráfico o movilidad suelen ser significativas para tamaños de buffer pequeños, corroborando la principal suposición de los modelos LUR, la cual indica que los contaminantes en la zona de estudio, son emitidos principalmente por el parque automotor (Bertazzon et al., 2015; Briggs et al., 1997; Habermann & Gouveia, 2012; Londoño & Cañón, 2015).

Generalmente, en diferentes estudios se observa la inclusión de variables de usos del suelo para grandes tamaños de buffer, excluyendo la intersección con los radios más pequeños. Con el fin de no comprometer la influencia de las variables de tráfico, es común combinar los tamaños para obtener diferentes radios de influencia en una sola regresión espacial, ya sea de fuentes de emisión del contaminante, como la distancia a la vía más cercana, o de remoción, como distancia a zonas verdes. En este trabajo, al estar limitada la cantidad de estaciones de monitoreo de PM_{2.5}, las cuales representan el conjunto de entrenamiento con el que se construye el modelo, se decidió no combinar los distintos tamaños de buffer, ya que al hacer esto, implicaría un mayor número de variables en el modelo final, generando un posible sobreajuste (Basagaña et al., 2012). Por esta razón, se optó por evaluar individualmente los diferentes tamaños de buffer.

¹ Fuentes: <https://www.eltiempo.com/colombia/medellin/en-medellin-comienzan-las-obras-de-metroplus-en-la-calle-12-sur-guayabal-243446> y https://caracol.com.co/emisora/2018/12/21/medellin/1545395363_887268.html

Otra variable común en los modelos seleccionados fue la altura a nivel de piso de las estaciones de monitoreo de $PM_{2.5}$. Esta variable, al igual que las variables meteorológicas, son comúnmente evaluadas en la construcción de los modelos LUR, para tener de manera implícita la influencia de la componente vertical en los modelos de regresión espaciales. Los modelos seleccionados que incluyeron la variable PISO, explicaban entre el 46 y 79% de la variabilidad del $PM_{2.5}$, resultados similares a los encontrados por Briggs et al. (1997), donde un modelo construido para NO_2 en Huddersfield, Reino Unido, incluyó esta variable y explicó el 61% de la variabilidad del contaminante.

Como se mencionó en la sección 4.1.1, las variables meteorológicas influyen en gran medida el transporte de los contaminantes. Para incorporar esta influencia en los modelos espaciales, se incluyeron seis variables ambientales: temperatura, velocidad del viento, precipitación y gradientes espaciales de las mismas. Las variables meteorológicas representaron un papel fundamental en la construcción de los modelos, donde la más común fue la velocidad del viento, cuyo efecto negativo era congruente con la física del sistema, ya que el viento es uno de los principales mecanismos que permite la dispersión de partículas en la atmósfera. Para once meses, las variables velocidad del viento promedio mensual y los tres gradientes fueron significativas, de manera variable, en los modelos de regresión.

En el trabajo realizado por Londoño (2018), se caracterizó espacialmente el $PM_{2.5}$ por medio de una regresión espacial multivariada general, es decir, sin utilizar las zonas de influencia características del modelo LUR. Se utilizaron como variables explicativas la temperatura, velocidad del viento y los gradientes de las mismas variables promedio mensual, para los años 2013 y 2014 de estaciones ubicadas en el Valle de Aburrá. A diferencia de los resultados obtenidos en este trabajo, estos modelos mensuales seleccionados explicaban entre el 11 y el 25% de la variabilidad del $PM_{2.5}$, en ocho estaciones de monitoreo, indicando la importancia de la adición de otras posibles variables explicativas en el modelo de regresión espacial. Sin embargo, al agregar una componente temporal en la metodología, se mejoró el rendimiento del modelo y se pudo explicar hasta el 86% de la variabilidad del contaminante, sugiriendo una dependencia de los niveles de $PM_{2.5}$ con los valores del mes anterior (Londoño, 2018a). Esto indica la importancia de la caracterización conjunta del espacio y tiempo para las series de contaminantes atmosféricos, con el fin de mejorar la comprensión de la evolución de los mismos.

El mismo autor, desarrolló un modelo LUR para PM_{10} en el Valle de Aburrá, a partir de 10 estaciones de monitoreo del contaminante para el mes de enero del año 2007. El modelo escogido, explicaba la variabilidad del PM_{10} en un 79% y estaba definido por una única variable, la distancia a la vía principal más cercana desde la estación de monitoreo (Londoño & Cañón, 2015); a diferencia de los modelos construidos en este trabajo, en donde no se refleja esta variable en ninguno de los modelos finales. Aunque esta variable es común en los modelos LUR construidos por diferentes autores, es importante mencionar la importancia de esta variable cuando se tratan de estaciones influenciadas directamente por el tráfico vehicular, que no necesariamente es el caso para las estaciones utilizadas en este trabajo, debido a la disponibilidad de información en las estaciones definidas como estrictamente de tráfico.

La mayoría de los modelos construidos en los meses de enero, febrero y marzo, presentaron signos en las variables explicativas contrarios a la dirección del efecto que se espera de las mismas. Por ejemplo, la distancia a la vía más cercana era positiva, es decir que, a mayor distancia a la vía, mayores los niveles de $PM_{2.5}$, contrario a lo reportado en la literatura (Allen et al., 2011; Bertazzon et al., 2015; Eeftens et al., 2012; Londoño & Cañón, 2015). Debido a esto, para los meses febrero y marzo, sólo se pudo encontrar un modelo en función del gradiente de temperatura promedio mensual y el flujo vehicular promedio mensual en la vía más cercana, el cual explicó 33% y 31% de la varianza del contaminante, respectivamente, indicando que, para la zona de estudio, hace falta incluir variables que puedan influenciar los niveles del $PM_{2.5}$ en la atmósfera. Sin embargo, en el estudio desarrollado por Habermann & Gouveia (2012), solamente el flujo vehicular llegó a explicar el 64% del contaminante evaluado (PM_{10}) en la ciudad de Sao Paulo, Brasil. Esto indica que esta metodología permite encontrar un modelo que es exclusivo de la zona de estudio donde se desarrolla, dado que, como se ha evaluado previamente, transferir un modelo LUR de una ciudad a otra con características similares, puede reducir en gran medida el rendimiento del modelo inicial (Allen et al., 2011).

Como se mencionó anteriormente, un modelo LUR construido en una zona específica es característico de la misma, sin embargo, el modelo puede ser extrapolado a otras zonas o ciudades, teniendo en cuenta que el rendimiento del mismo podrá verse afectado, como consecuencia de la naturaleza del problema que depende de condiciones locales. Siguiendo la metodología propuesta por Allen et al. (2011) para extrapolar un modelo LUR a otras ciudades, primero, se debe evaluar la disponibilidad de información, tanto para la variable dependiente, como para las variables explicativas seleccionadas en el modelo final. Además, se ha sugerido que es importante considerar cómo difieren las condiciones de la nueva zona de estudio, en comparación a la usada en la construcción del modelo inicial, debido a que se puede presentar la necesidad de excluir algunas variables seleccionadas, como, por ejemplo, la altura sobre el nivel del mar debido a una baja variabilidad en las elevaciones, en la nueva zona de estudio. A pesar de la limitación que implica extrapolar o transferir un modelo LUR, los resultados pueden presentar un mejor ajuste que otro tipo de metodologías, sumado, a la ventaja relacionada a la relativa simplicidad y a los bajos costos de construcción del modelo LUR (Allen et al., 2011; Jerrett et al., 2005).

En la caracterización espacial multivariada, una de las principales limitaciones está relacionada con la cantidad de sitios de monitoreo, debido a que, como han sugerido Basagaña et al. (2012), son preferibles los modelos LUR con un gran tamaño de muestra, dado que permiten mejorar el rendimiento y reducir el riesgo de sobreajuste en el modelo. Además, como se observó en los modelos LUR mensuales seleccionados en este trabajo, en muestras pequeñas se pueden presentar grandes variaciones en las posibles variables explicativas seleccionadas para cada modelo, algo que se va restringiendo a medida que se aumenta el número de estaciones de monitoreo del contaminante; indicando así que ante un aumento en la cantidad de estaciones, es posible encontrar un conjunto de variables explicativas que caracterice el contaminante en toda la ventana de tiempo.

A pesar del desempeño general de los modelos, se mantiene la estacionalidad característica de los niveles de $PM_{2.5}$ en el Valle de Aburrá, donde el pico de concentraciones se presenta en marzo y el momento de menos contaminación se presenta a mediados del año, como se ha demostrado en diferentes trabajos (Área Metropolitana del Valle de Aburrá, 2018; Londoño, 2018a). Durante la temporada lluviosa en la Región Andina Colombiana, se presentan condiciones de estabilidad atmosférica en el Valle de Aburrá, limitando la dispersión del material particulado en la atmósfera y generando un incremento en los niveles, que pueden llegar a superar el umbral establecido en la norma nacional.

Por otra parte, para todo el periodo de análisis, las zonas más contaminadas se encontraron en el centro-oriente y sur-occidente de la zona de estudio, sin importar las variables incluidas en el modelo, con variaciones en la magnitud de las concentraciones del $PM_{2.5}$ a lo largo del año; siendo congruente con lo encontrado por Londoño (2018). Sin embargo, al determinar las zonas menos contaminadas, a partir de los modelos mensuales, se encontró que los menores niveles del contaminante en promedio, se presentaron al noroccidente de la zona de estudio, resultados contrarios a lo sugerido previamente, donde estas zonas se encuentran en el sur occidente del municipio de Medellín (Londoño, 2018a). Lo anterior puede estar relacionado con las estaciones utilizadas para la caracterización espacial del contaminante, de las cuales, sólo las estaciones MED.UNNV, ITA.CONC e ITA.CJUS fueron seleccionadas en ambos trabajos. Además, los estudios fueron desarrollados con datos medidos en diferentes periodos de tiempo, aportando esto a la variabilidad de los resultados.

Debido a la naturaleza del $PM_{2.5}$, se dificulta encontrar un enfoque que permita representar espacialmente la variabilidad del contaminante, la cual se encuentra influenciada por diversos factores, como la meteorología local y regional, las fuentes de emisión antropogénicas y naturales y las fuentes de remoción del mismo, que tienen asociado un comportamiento no lineal. Sin embargo, con esta metodología se pueden encontrar modelos bien especificados que permiten explicar parte de la variabilidad de contaminante, de forma económica y con información escasa, siendo una de las principales ventajas de su aplicación.

8. Conclusiones

En este trabajo se presenta una metodología para caracterizar espacialmente y estimar las concentraciones de $PM_{2.5}$ en la ciudad de Medellín, mediante modelos de regresión de usos del suelo, basándose en una metodología utilizada ampliamente como insumo para estimar la exposición a contaminantes atmosféricos. En general, el método consiste en la construcción de una regresión espacial multivariada que permite explicar la variabilidad local de un contaminante atmosférico, evaluando distintas zonas de influencia alrededor de las estaciones de monitoreo del mismo. Por tanto, los modelos de regresión de usos del suelo son útiles para estimar concentraciones de $PM_{2.5}$ en la zona de estudio.

En el proceso de construcción de los modelos LUR, es importante realizar una exploración inicial de la cobertura temporal de la información disponible, sobre todo, cuando se desea aplicar la metodología para distintos momentos en el periodo de análisis (por ejemplo, cada mes). Al ser información secundaria, el análisis exploratorio debe ser riguroso para descartar información poco plausible y de baja calidad ocasionada por la presencia de valores perdidos, teniendo en cuenta que, estas limitaciones se presentan sin importar el tipo de variable a analizar.

A pesar del tamaño de muestra limitado para las estaciones de monitoreo de $PM_{2.5}$ utilizadas en este estudio, la construcción y selección de los mejores modelos basados en el método OLS, permitió seleccionar las regresiones mejor especificadas; además, al evaluar la correlación de todas las posibles variables explicativas con los datos espaciales del contaminante, se pudo evitar la adición de variables que no presentaran una relación de dependencia con el mismo y así evitar una inadecuada especificación en el modelo de regresión final.

En la construcción de los modelos, es importante definir previamente, basados en la naturaleza del fenómeno evaluado, el efecto esperado de las variables explicativas, debido a que es muy probable que se presenten modelos de regresión en donde una variable que aporta al incremento de los niveles de $PM_{2.5}$, como lo es el flujo vehicular, presente un coeficiente negativo en la ecuación, indicando una relación inversa con la variable dependiente.

Para el 50% de los meses estudiados, los mejores modelos LUR fueron seleccionados en el buffer de 200 m, debido a que presentaban el mejor rendimiento entre todos los modelos construidos, de acuerdo a los criterios de especificación. Con la finalidad de definir un buffer específico para la zona de estudio, se requiere tener una ventana de información más amplia y corroborar los resultados obtenidos en este trabajo, que para el año 2018, dan la idea de que los niveles de $PM_{2.5}$ en las estaciones de monitoreo se ven influenciadas, principalmente, por la existencia de fuentes y sumideros a 200 m a la redonda. Sin embargo, es necesario profundizar en la aplicación de los modelos LUR en la zona de estudio para reforzar la construcción y evaluación de esta hipótesis.

Sin importar las variables que definieron el modelo LUR mensual, para todo el periodo de análisis se observa que las zonas más contaminadas se encuentran en el centro-oriente y sur-occidente de la zona de estudio, con variaciones en la magnitud de las concentraciones del $PM_{2.5}$ a lo largo del año. Además, fue posible evidenciar la estacionalidad característica del contaminante, donde entre

febrero y marzo, se presentaron las concentraciones más altas, y entre junio y julio, los niveles más bajos. El entendimiento de la distribución espacial de la concentración del $PM_{2.5}$ en una escala intraurbana, permite aportar información sobre la exposición, soportando acciones que se orienten a la minimización de riesgos, teniendo en cuenta la población susceptible.

Los resultados obtenidos en este trabajo, al igual que en estudios previos, indican que parte de la variabilidad espacial de los niveles de $PM_{2.5}$ en la ciudad de Medellín, estaría explicada tanto por variables relacionadas al tráfico, como por variables meteorológicas. Mediante las distintas ecuaciones de regresión obtenidas, se estimó el valor del contaminante en otros puntos, permitiendo obtener zonas de altas y bajas concentraciones, que se podrían tener en cuenta en los procesos de gestión y ampliación de la Red de Monitoreo de Calidad del Aire del Valle de Aburrá y del Sistema de Vigilancia en Salud Ambiental, debido a que los resultados permiten obtener niveles de exposición a una escala individual de los habitantes en la zona de estudio definida, necesaria para mejorar la estructuración de estudios epidemiológicos.

9. Recomendaciones

El flujo vehicular a 200 m fue una de las variables predominantes en los modelos LUR finales para la zona de estudio, razón por la cual, para el diseño e implementación de políticas de movilidad sostenible, es importante ahondar en el estudio de la posible influencia de este buffer en los niveles de inmisión del $PM_{2.5}$, con el fin de establecer medidas en pro de la mitigación de los impactos que este contaminante pueda generar. Asimismo, los resultados muestran que las zonas verdes tienen un efecto sumidero, disminuyendo las concentraciones del contaminante, sobre todo a distancias menores a 200 m, hallazgo que debe ser estudiado con mayor detalle, debido al potencial que propone la adecuación de esta categoría de uso del suelo en los planes de ordenamiento territorial, para mejorar la calidad del aire.

Ante la limitación que representa la cantidad de estaciones de monitoreo del contaminante de interés, se propone evaluar en un trabajo futuro el desempeño de la metodología con niveles obtenidos a partir de sensores de bajo costo, como la Red de Ciudadanos Científicos administrada por el Sistema de Alertas Tempranas del Valle de Aburrá (SIATA); además de evaluar la inclusión de mediciones móviles para mejorar el rendimiento del modelo resultante. Por otro lado, otra alternativa para mejorar el rendimiento del modelo y la variabilidad espacial explicada del $PM_{2.5}$, es incrementar el acceso a estaciones que estén influenciadas directamente por el tráfico vehicular, además de densificar la información relacionada a los aforos vehiculares.

Aunque el uso de una escala de análisis mensual, congruente con estudios epidemiológicos, permitió evidenciar la estacionalidad en los niveles del contaminante, sería interesante evaluar a futuro, el desempeño de esta metodología en diferentes escalas temporales en la zona de estudio, ya sea aplicando promedios anuales, multianuales, promedios horarios, o valores horarios de un día aleatorio. Siendo la escala horaria de especial interés si se desea evaluar el desempeño del modelo en los periodos de mayor flujo vehicular en el ciclo diurno, con el fin de influenciar la toma de decisiones en términos de movilidad sostenible para la reducción de los niveles de contaminación del aire en el Valle de Aburrá. Además, puede evaluarse la inclusión de una componente temporal en el modelo, que como se ha sugerido en estudios anteriores, permite mejorar en gran medida el desempeño del mismo.

En los últimos años, se han realizado en la región, grandes aportes para el análisis de la contaminación del aire, especialmente la generada por altos niveles de $PM_{2.5}$, encaminando esfuerzos por parte de las autoridades ambientales, la academia y la ciudadanía, para mejorar la calidad del aire. Aunque en este trabajo se exhibe en cierta medida la influencia de las fuentes móviles en las concentraciones del $PM_{2.5}$ en el Valle de Aburrá, actualmente, se ha encontrado la existencia de otras fuentes externas, como los incendios forestales, que alteran los niveles del contaminante en diferentes zonas del país.

Por esta razón, es de vital importancia mejorar el entendimiento, no sólo de fuentes locales de emisión de material particulado y otros contaminantes precursores, y de la influencia de la meteorología local, sino también de las posibles fuentes regionales en conjunto con patrones

atmosféricos, que permitan el transporte de las partículas y su influencia en los niveles del contaminante en zonas específicas. Para tal fin, en un trabajo futuro, se recomienda evaluar la influencia de fuentes regionales, como la presencia de quema de biomasa, definiendo tamaños de buffer más amplios, para ver los efectos del transporte del contaminante de mediano y largo alcance, permitiendo incluir esta información en la construcción del modelo.

10. Referencias

- Alcaldía de Medellín. (2014). *Acuerdo 48 de 2014 - Plan de Ordenamiento Territorial del Municipio de Medellín*. Medellín. Recuperado de https://www.medellin.gov.co/irj/go/km/docs/pccdesign/SubportaldelCiudadano_2/PlandeDesarrollo_0_17/ProgramasyProyectos/Shared_Content/Documentos/2014/POT/ACUERDO POT-19-12-2014.pdf
- Alcaldía de Medellín. (2015). *Proyecciones de Población 2016 a 2020 de Medellín*. Recuperado el 13 de septiembre de 2019, de <https://www.medellin.gov.co/irj/portal/medellin?NavigationTarget=navurl://06bdb4d911e35cb9d3de717115deedc3>
- Allen, R. W., Amram, O., Wheeler, A. J., & Brauer, M. (2011). The transferability of NO and NO₂ land use regression models between cities and pollutants. *Atmospheric Environment*, 45(2), 369–378. <https://doi.org/10.1016/j.atmosenv.2010.10.002>
- Araín, M. A., Blair, R., Finkelstein, N., Brook, J. R., Sahsuvaroglu, T., Beckerman, B., ... Jerrett, M. (2007). The use of wind fields in a land use regression model to predict air pollution concentrations for health exposure studies. *Atmospheric Environment*, 41(16), 3453–3464. <https://doi.org/10.1016/j.atmosenv.2006.11.063>
- Área Metropolitana del Valle de Aburrá. (2013). *Simulación de procesos dispersivos en el Valle de Aburrá*. Medellín. Recuperado de http://www.metropol.gov.co/CalidadAire/Paginas/wz_biblioteca_aire.aspx
- Área Metropolitana del Valle de Aburrá. (2016a). *Acuerdo Metropolitano N° 15. Plan Operacional para Enfrentar Episodios Críticos de Contaminación Atmosférica - POECA*. Recuperado de <http://www.metropol.gov.co/Documents/AcuerdoN15.PDF>
- Área Metropolitana del Valle de Aburrá. (2016b). *Escenario de riesgo por contaminación atmosférica*. Recuperado de http://www.metropol.gov.co/CalidadAire/isdocConvenio243/Escenario_Contaminacion_Atmosferica.pdf
- Área Metropolitana del Valle de Aburrá. (2016c). *Informe de calidad del aire, meteorología y ruido en el municipio de Medellín. Informe acumulado enero - diciembre de 2015*. Medellín. Recuperado de http://www.areadigital.gov.co/CalidadAire/lsdocCalidadAireUnal/Informe_CD280_MEDELLIN_Abril_2014_0037-2_13.0.pdf
- Área Metropolitana del Valle de Aburrá. (2017a). *Inventario de emisiones atmosféricas del Valle de Aburrá, actualización 2015*. Recuperado de http://www.metropol.gov.co/CalidadAire/isdocConvenio243/Informe_Inventario_emisiones_2015.pdf
- Área Metropolitana del Valle de Aburrá. (2017b). *Pronóstico estadístico de calidad del aire*. Medellín. Recuperado de http://www.metropol.gov.co/CalidadAire/Paginas/wz_biblioteca_aire.aspx

- Área Metropolitana del Valle de Aburrá. (2018). *Calidad del aire y sus efectos en la salud de la población de los diez municipios del Valle de Aburrá, 2008-2015*. Medellín. Recuperado de https://www.metropol.gov.co/ambiental/calidad-del-aire/Biblioteca-aire/Re-analisis/ContaminacionAtmosferica_y_sus_Efectos_en_la_Salud-AMVA_2019.pdf
- Basagaña, X., Rivera, M., Aguilera, I., Agis, D., Bouso, L., Elosua, R., ... Künzli, N. (2012). Effect of the number of measurement sites on land use regression models in estimating local air pollution. *Atmospheric Environment*, 54, 634–642. <https://doi.org/10.1016/j.atmosenv.2012.01.064>
- Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., ... de Hoogh, K. (2013). Development of NO₂ and NO_x land use regression models for estimating air pollution exposure in 36 study areas in Europe - The ESCAPE project. *Atmospheric Environment*, 72(2), 10–23. <https://doi.org/10.1016/j.atmosenv.2013.02.037>
- Beelen, R., Voogt, M., Duyzer, J., Zandveld, P., & Hoek, G. (2010). Comparison of the performances of land use regression modelling and dispersion modelling in estimating small-scale variations in long-term air pollution concentrations in a Dutch urban area. *Atmospheric Environment*, 44(36), 4614–4621. <https://doi.org/10.1016/j.atmosenv.2010.08.005>
- Bertazzon, S., Johnson, M., Eccles, K., & Kaplan, G. G. (2015). Accounting for spatial effects in land use regression for urban air pollution modeling. *Spatial and Spatio-temporal Epidemiology*, 14–15, 9–21. <https://doi.org/10.1016/j.sste.2015.06.002>
- Blanes-Vidal, V., Cantuaria, M. L., & Nadimi, E. S. (2017). A novel approach for exposure assessment in air pollution epidemiological studies using neuro-fuzzy inference systems: Comparison of exposure estimates and exposure-health associations. *Environmental Research*, 154(December 2016), 196–203. <https://doi.org/10.1016/j.envres.2016.12.028>
- Briggs, D. (2005). The Role of Gis: Coping With Space (And Time) in Air Pollution Exposure Assessment. *Journal of Toxicology and Environmental Health, Part A*, 68(13–14), 1243–1261. <https://doi.org/10.1080/15287390590936094>
- Briggs, D., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., ... Van Der Veen, A. (1997). Mapping urban air pollution using gis: A regression-based approach. *International Journal of Geographical Information Science*, 11(7), 699–718. <https://doi.org/10.1080/136588197242158>
- Buteau, S., Hatzopoulou, M., Crouse, D. L., Smargiassi, A., Burnett, R. T., Logan, T., ... Goldberg, M. S. (2017). Comparison of spatiotemporal prediction models of daily exposure of individuals to ambient nitrogen dioxide and ozone in Montreal, Canada. *Environmental Research*, 156(January), 201–230. <https://doi.org/10.1016/j.envres.2017.03.017>
- Camilo Blanco-Becerra, L., Miranda-Soberanis, V., Hernandez-Cadena, L., Barraza-Villarreal, A., Junger, W., Hurtado-Diaz, M., & Romieu, I. (2014). Effect of particulate matter less than 10 μ m (PM₁₀) on mortality in Bogota, Colombia: a time-series analysis, 1998-2006. *Salud Publica De Mexico*, 56(4), 363–370.

- Cattani, G., Gaeta, A., Di Menno di Bucchianico, A., De Santis, A., Gaddi, R., Cusano, M., ... Cesaroni, G. (2017). Development of land-use regression models for exposure assessment to ultrafine particles in Rome, Italy. *Atmospheric Environment*, 156, 52–60. <https://doi.org/10.1016/j.atmosenv.2017.02.028>
- Chen, L., Gao, S., Zhang, H., Sun, Y., Ma, Z., Vedal, S., ... Bai, Z. (2018). Spatiotemporal modeling of PM 2.5 concentrations at the national scale combining land use regression and Bayesian maximum entropy in China. *Environment International*, 116(March), 300–307. <https://doi.org/10.1016/j.envint.2018.03.047>
- Chen, L., Shi, M., Li, S., Bai, Z., & Wang, Z. (2017). Combined use of land use regression and BenMAP for estimating public health benefits of reducing PM2.5 in Tianjin, China. *Atmospheric Environment*, 152, 16–23. <https://doi.org/10.1016/j.atmosenv.2016.12.023>
- Clougherty, J. E., Kheirbek, I., Eisl, H. M., Ross, Z., Pezeshki, G., Gorczynski, J. E., ... Matte, T. (2013). Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: The New York City Community Air Survey (NYCCAS). *Journal of Exposure Science and Environmental Epidemiology*, 23(3), 232–240. <https://doi.org/10.1038/jes.2012.125>
- Colville, R. N., Hutchinson, E. J., Mindell, J. S., & Warren, R. F. (2001). The transport sector as a source of air pollution. *Atmospheric Environment*, 35(9), 1537–1565. [https://doi.org/10.1016/S1352-2310\(00\)00551-3](https://doi.org/10.1016/S1352-2310(00)00551-3)
- Cordioli, M., Pironi, C., De Munari, E., Marmiroli, N., Lauriola, P., & Ranzi, A. (2017). Combining land use regression models and fixed site monitoring to reconstruct spatiotemporal variability of NO2 concentrations over a wide geographical area. *Science of the Total Environment*, 574, 1075–1084. <https://doi.org/10.1016/j.scitotenv.2016.09.089>
- Cowie, C. T., Garden, F., Jegasothy, E., Knibbs, L. D., Hanigan, I., Morley, D., ... Marks, G. B. (2019). Comparison of model estimates from an intra-city land use regression model with a national satellite-LUR and a regional Bayesian Maximum Entropy model, in estimating NO 2 for a birth cohort in Sydney, Australia. *Environmental Research*, 174(March), 24–34. <https://doi.org/10.1016/j.envres.2019.03.068>
- Crouse, D. L., Goldberg, M. S., & Ross, N. A. (2009). A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in Montreal, Canada. *Atmospheric Environment*, 43(32), 5075–5084. <https://doi.org/10.1016/j.atmosenv.2009.06.040>
- Dabass, A., Talbott, E. O., Bilonick, R. A., Rager, J. R., Venkat, A., Marsh, G. M., ... Xue, T. (2016). Using spatio-temporal modeling for exposure assessment in an investigation of fine particulate air pollution and cardiovascular mortality. *Environmental Research*, 151, 564–572. <https://doi.org/10.1016/j.envres.2016.08.024>
- de Hoogh, K., Chen, J., Gulliver, J., Hoffmann, B., Hertel, O., Ketznel, M., ... Hoek, G. (2018). Spatial PM2.5 , NO2 , O3 and BC models for Western Europe – Evaluation of spatiotemporal stability. *Environment International*, 120(2), 81–92. <https://doi.org/10.1016/j.envint.2018.07.036>

- de Hoogh, K., Gulliver, J., Donkelaar, A. van, Martin, R. V., Marshall, J. D., Bechle, M. J., ... Hoek, G. (2016). Development of West-European PM_{2.5} and NO₂ land use regression models incorporating satellite-derived and chemical transport modelling data. *Environmental Research*, 151(2), 1–10. <https://doi.org/10.1016/j.envres.2016.07.005>
- de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M. J., ... Bellander, T. (2014). Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environment International*, 73, 382–392. <https://doi.org/10.1016/j.envint.2014.08.011>
- Departamento Administrativo de Planeación. (2018). *Anuario Estadístico Municipio de Itagüí*. Itagüí.
- Departamento Nacional de Estadística. (2019). *Resultados Censo Nacional de Población y Vivienda: tercera entrega*. Recuperado de <https://www.dane.gov.co/files/censo2018/informacion-tecnica/cnpv-2018-presentacion-3ra-entrega.pdf>
- Dijkema, M. B., Gehring, U., van Strien, R. T., van der Zee, S. C., Fischer, P., Hoek, G., & Brunekreef, B. (2011). A comparison of different approaches to estimate small-scale spatial variation in outdoor NO₂ concentrations. *Environmental Health Perspectives*, 119(5), 670–675. <https://doi.org/10.1289/ehp.0901818>
- Dirgawati, M., Heyworth, J. S., Wheeler, A. J., McCaul, K. A., Blake, D., Boeyen, J., ... Hinwood, A. (2016). Development of Land Use Regression models for particulate matter and associated components in a low air pollutant concentration airshed. *Atmospheric Environment*, 144, 69–78. <https://doi.org/10.1016/j.atmosenv.2016.08.013>
- Donahue, N. M. (2017). Air Pollution and Air Quality. En *Green Chemistry: An Inclusive Approach* (pp. 151–176). <https://doi.org/10.1016/B978-0-12-809270-5.00007-8>
- Dons, E., Van Poppel, M., Int Panis, L., De Prins, S., Berghmans, P., Koppen, G., & Matheussen, C. (2014). Land use regression models as a tool for short, medium and long term exposure to traffic related air pollution. *Science of the Total Environment*, 476–477, 378–386. <https://doi.org/10.1016/j.scitotenv.2014.01.025>
- EAFIT, U. (2019). *Informe Anual de Calidad del Aire 2018. Contrato CCT 504 de 2019*. Recuperado de [https://www.metropol.gov.co/ambiental/calidad-del-aire/informes_red_calidaddeaire/Informe Anual Aire 2018.pdf](https://www.metropol.gov.co/ambiental/calidad-del-aire/informes_red_calidaddeaire/Informe%20Anual%20Aire%202018.pdf)
- Eeftens, M., Beelen, R., De Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., ... Hoek, G. (2012). Development of land use regression models for PM_{2.5}, PM_{2.5} absorbance, PM₁₀ and PM coarse in 20 European study areas; results of the ESCAPE Project. *Environmental Science and Technology*, 46(20). <https://doi.org/10.1021/es301948k>
- Eom, J. K., Park, M. S., Heo, T. Y., & Huntsinger, L. F. (2006). Improving the prediction of annual average daily traffic for nonfreeway facilities by applying a spatial statistical method. *Transportation Research Record: Journal of the Transportation Research Board*, (1968), 20–29. <https://doi.org/https://doi.org/10.1177/0361198106196800103>

- ESRI. (2019a). Cómo funciona la densidad kernel—Ayuda | ArcGIS for Desktop. Recuperado el 10 de noviembre de 2019, de <http://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-analyst-toolbox/how-kernel-density-works.htm>
- ESRI. (2019b). Interpretar los resultados de OLS. Recuperado el 10 de octubre de 2019, de <https://desktop.arcgis.com/es/arcmap/latest/tools/spatial-statistics-toolbox/interpreting-ols-results.htm>
- Fischer, M. M., & Getis, A. (2010). *Handbook of Applied Spatial Analysis*. Springer. <https://doi.org/10.1007/978-3-642-03647-7>
- Golub, E., Klytchnikova, I., Sanchez-Martinez, G., Belausteguigoitia, J. C., & Molina, C. M. (2014). *Environmental Health Costs in Colombia*. Washington DC. Recuperado de <http://documents.worldbank.org/curated/en/657631468047104545/Environmental-health-costs-in-Colombia-the-changes-from-2002-to-2010>
- Guerrero Hoyos, L. Á., & Aristizábal Giraldo, E. (2019). Estimación y análisis de umbrales críticos de lluvia para la ocurrencia de avenidas torrenciales en el Valle de Aburrá (Antioquia) Luis. *Revista EIA*, 16(32), 97–111. <https://doi.org/https://doi.org/10.24050/reia.v16i32.1281>
- Habermann, M., & Gouveia, N. (2012). Aplicação de regressão baseada no uso do solo para prever a concentração de material particulado inalável no município de São Paulo, Brasil. *Engenharia Sanitaria e Ambiental*, 17(2), 155–162. <https://doi.org/10.1590/S1413-41522012000200004>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. del P. (2010). *Metodología de la investigación* (Quinta Edición). Mexico DF: McGraw-Hill.
- Hoek, G., Beelen, R., de Hoogh, K., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, 42(33), 7561–7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>
- Huang, H. Bin, Lai, C. H., Chen, G. W., Lin, Y. Y., Jaakkola, J. J. K., Liou, S. H., & Wang, S. L. (2012). Traffic-related air pollution and DNA damage: A longitudinal study in Taiwanese traffic conductors. *PLoS ONE*, 7(5), 1–8. <https://doi.org/10.1371/journal.pone.0037412>
- IDEAM. (2018). *Informe del estado de la calidad del aire en Colombia 2017* (Primera Ed). Bogotá D.C. Recuperado de http://documentacion.ideam.gov.co/openbiblio/bvirtual/023844/Informe_ECalidadl_Aire_2017.pdf
- Janssen, N. A. H., Fischer, P., Marra, M., Ameling, C., & Cassee, F. R. (2013). Short-term effects of PM_{2.5}, PM₁₀ and PM_{2.5-10} on daily mortality in the Netherlands. *Science of the Total Environment*, 463–464, 20–26. <https://doi.org/10.1016/j.scitotenv.2013.05.062>
- Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahuvaroglu, T., ... Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of Exposure Analysis and Environmental Epidemiology*, 15(2), 185–204.

<https://doi.org/10.1038/sj.jea.7500388>

- Khreis, H., & Nieuwenhuijsen, M. J. (2017). Traffic-related air pollution and childhood asthma: Recent advances and remaining gaps in the exposure assessment methods. *International Journal of Environmental Research and Public Health*, 14(3), 1–19. <https://doi.org/10.3390/ijerph14030312>
- Kim Oanh, N. T., Chutimon, P., Ekbordin, W., & Supat, W. (2005). Meteorological pattern classification and application for forecasting air pollution episode potential in a mountain-valley area. *Atmospheric Environment*, 39(7), 1211–1225. <https://doi.org/10.1016/j.atmosenv.2004.10.015>
- Kim, Y., Myung, W., Won, H. H., Shim, S., Jeon, H. J., Choi, J., ... Kim, D. K. (2015). Association between air pollution and suicide in South Korea: A nationwide study. *PLoS ONE*, 10(2), 1–10. <https://doi.org/10.1371/journal.pone.0117929>
- Largerón, Y., & Staquet, C. (2016). Persistent inversion dynamics and wintertime PM10 air pollution in Alpine valleys. *Atmospheric Environment*, 135, 92–108. <https://doi.org/10.1016/j.atmosenv.2016.03.045>
- Larkin, A., Geddes, J. A., Martin, R. V., Xiao, Q., Liu, Y., Marshall, J. D., ... Hystad, P. (2017). Global Land Use Regression Model for Nitrogen Dioxide Air Pollution. *Environmental Science and Technology*, 51(12), 6957–6964. <https://doi.org/10.1021/acs.est.7b01148>
- Li, J., & Heap, A. D. (2011). A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecological Informatics*, 6(3–4), 228–241. <https://doi.org/10.1016/j.ecoinf.2010.12.003>
- Li, J., & Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling and Software*, 53, 173–189. <https://doi.org/10.1016/j.envsoft.2013.12.008>
- Li, R., Ma, T., Xu, Q., & Song, X. (2018). Using MAIAC AOD to verify the PM2.5 spatial patterns of a land use regression model. *Environmental Pollution*, 243, 501–509. <https://doi.org/10.1016/j.envpol.2018.09.026>
- Linares, C., & Díaz, J. (2010). Short-term effect of concentrations of fine particulate matter on hospital admissions due to cardiovascular and respiratory causes among the over-75 age group in Madrid, Spain. *Public Health*, 124(1), 28–36. <https://doi.org/10.1016/j.puhe.2009.11.007>
- Londoño Ciro, L. A., & Valdés Quintero, J. C. (2012). *Geoestadística aplicada: Generación de mapas de interpolación para el estudio de fenómenos distribuidos espacialmente*. Editorial Académica Española. <https://doi.org/10.1017/CBO9781107415324.004>
- Londoño, L. (2018a). *Metodología para caracterizar espacio-temporalmente la concentración de material particulado en valles intramontanos con información escasa*. Universidad de Antioquia.
- Londoño, L. (2018b). Metodología para la caracterización espacio-temporal de PM2.5 en el área

- urbana de Medellín-Colombia. *Revista EIA*, (574), 113–132.
- Londoño, L., & Cañón, J. (2015). Metodología para la aplicación de modelos de regresión de usos del suelo en la estimación local de la concentración mensual de PM10 en Medellín – Colombia. *Revista Politécnica*, 11(21), 29–40.
- Lu, B., Charlton, M., Harris, P., & Fotheringham, A. S. (2014). Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *International Journal of Geographical Information Science*, 28(4), 660–681. <https://doi.org/10.1080/13658816.2013.865739>
- Luong, L. M. T., Phung, D., Sly, P. D., Morawska, L., & Thai, P. K. (2017). The association between particulate air pollution and respiratory admissions among young children in Hanoi, Vietnam. *Science of the Total Environment*, 578, 249–255. <https://doi.org/10.1016/j.scitotenv.2016.08.012>
- Maantay, J. A., & McLafferty, S. (2011). *Geospatial Analysis of Environmental Health*. Springer. <https://doi.org/10.1007/978-94-007-0329-2>
- Marshall, J. D., Nethery, E., & Brauer, M. (2008). Within-urban variability in ambient air pollution: Comparison of estimation methods. *Atmospheric Environment*, 42(5), 1359–1369. <https://doi.org/10.1016/j.atmosenv.2007.08.012>
- Mercer, L. D., Szpiro, A. A., Sheppard, L., Lindström, J., Adar, S. D., Allen, R. W., ... Kaufman, J. D. (2011). Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NOx) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Atmospheric Environment*, 45(26), 4412–4420. <https://doi.org/10.1016/j.atmosenv.2011.05.043>
- Michanowicz, D. R., Shmool, J. L. C., Cambal, L., Tunno, B. J., Gillooly, S., Olson Hunt, M. J., ... Clougherty, J. E. (2016). A hybrid land use regression/line-source dispersion model for predicting intra-urban NO2. *Transportation Research Part D: Transport and Environment*, 43, 181–191. <https://doi.org/10.1016/j.trd.2015.12.007>
- Michanowicz, D. R., Shmool, J. L. C., Tunno, B. J., Tripathy, S., Gillooly, S., Kinnee, E., & Clougherty, J. E. (2016). A hybrid land use regression/AERMOD model for predicting intra-urban variation in PM2.5. *Atmospheric Environment*, 131, 307–315. <https://doi.org/10.1016/j.atmosenv.2016.01.045>
- Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia. (2010a). Manual de diseño de sistemas de vigilancia de la calidad del aire. *Protocolo para el Monitoreo y Seguimiento de la Calidad del Aire*. Bogotá D.C. Recuperado de http://www.minambiente.gov.co/images/AsuntosambientalesySectorialyUrbana/pdf/contaminacion_atmosferica/Protocolo_Calidad_del_Aire_-_Manual_Diseño.pdf
- Ministerio de Ambiente Vivienda y Desarrollo Territorial. República de Colombia. (2010b). Resolución No 610 de 2010: Por la cual se modifica la Resolución 601 del 4 de abril de 2006.

- Ministerio de Ambiente y Desarrollo Sostenible. República de Colombia. (2017). Resolución 2254 de 2017. Por la cual se adopta la norma de calidad del aire ambiente y se dictan otras disposiciones. Bogotá D.C. Recuperado de http://www.minambiente.gov.co/images/normativa/app/resoluciones/96-res_2254_de_2017.pdf
- Mølter, A., Lindley, S., de Vocht, F., Simpson, A., & Agius, R. (2010a). Modelling air pollution for epidemiologic research - Part I: A novel approach combining land use regression and air dispersion. *Science of the Total Environment*, 408(23), 5862–5869. <https://doi.org/10.1016/j.scitotenv.2010.08.027>
- Mølter, A., Lindley, S., de Vocht, F., Simpson, A., & Agius, R. (2010b). Modelling air pollution for epidemiologic research - Part II: Predicting temporal variation through land use regression. *Science of the Total Environment*, 409(1), 211–217. <https://doi.org/10.1016/j.scitotenv.2010.10.005>
- Montagne, D. R., Hoek, G., Klompmaker, J. O., Wang, M., Meliefste, K., & Brunekreef, B. (2015). Land Use Regression Models for Ultrafine Particles and Black Carbon Based on Short-Term Monitoring Predict Past Spatial Variation. *Environmental Science and Technology*, 49(14), 8712–8720. <https://doi.org/10.1021/es505791g>
- Musunuru, A., Wei, R., & Porter, R. J. (2017). Predicting Day and Night Traffic Volumes on Rural Roads for Statistical Road Safety Modeling. *Transportation Research Record: Journal of the Transportation Research Board*, 2659, 192–203. <https://doi.org/http://dx.doi.org/10.3141/2659-21>
- Naughton, O., Donnelly, A., Nolan, P., Pilla, F., Misstear, B. D., & Broderick, B. (2018). A land use regression model for explaining spatial variation in air pollution levels using a wind sector based approach. *Science of the Total Environment*, 630, 1324–1334. <https://doi.org/10.1016/j.scitotenv.2018.02.317>
- Oliver, M. A., & Webster, R. (2014). A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena*, 113, 56–69. <https://doi.org/10.1016/j.catena.2013.09.006>
- Papanastasiou, D. K., Melas, D., & Kioutsioukis, I. (2007). Development and assessment of neural network and multiple regression models in order to predict PM10 levels in a medium-sized Mediterranean city. *Water, Air, and Soil Pollution*, 182(1–4), 325–334. <https://doi.org/10.1007/s11270-007-9341-0>
- Pateraki, S., Asimakopoulos, D. N., Flocas, H. A., Maggos, T., & Vasilakos, C. (2012). The role of meteorology on different sized aerosol fractions (PM10, PM2.5, PM2.5-10). *Science of the Total Environment*, 419, 124–135. <https://doi.org/10.1016/j.scitotenv.2011.12.064>
- Rodopoulou, S., Chalbot, M. C., Samoli, E., DuBois, D. W., San Filippo, B. D., & Kavouras, I. G. (2014). Air pollution and hospital emergency room and admissions for cardiovascular and respiratory diseases in Doña Ana County, New Mexico. *Environmental Research*, 129, 39–46. <https://doi.org/10.1016/j.envres.2013.12.006>

- Sarlas, G., & Axhausen, K. W. (2016). Exploring spatial methods for prediction of traffic volumes. En *16 th Swiss Transport Research Conference* (pp. 12–19). Ascona. Recuperado de <https://doi.org/10.3929/ethz-b-000116988>
- Sava, F., & Carlsten, C. (2012). Respiratory Health Effects of Ambient Air Pollution: An Update. *Clinics in Chest Medicine*, 33(4), 759–769. <https://doi.org/10.1016/j.ccm.2012.07.003>
- Setton, E. M., Hystad, P. W., & Keller, C. P. (2005). Opportunities for using spatial property assessment data in air pollution exposure assessments. *International journal of health geographics*, 4, 26. <https://doi.org/10.1186/1476-072X-4-26>
- Shekarrizfard, M., Valois, M. F., Goldberg, M. S., Crouse, D., Ross, N., Parent, M. E., ... Hatzopoulou, M. (2015). Investigating the role of transportation models in epidemiologic studies of traffic related air pollution and health effects. *Environmental Research*, 140, 282–291. <https://doi.org/10.1016/j.envres.2015.04.002>
- Shi, Y., Ho, H. C., Xu, Y., & Ng, E. (2018). Improving satellite aerosol optical Depth-PM2.5 correlations using land use regression with microscale geographic predictors in a high-density urban context. *Atmospheric Environment*, 190(July), 23–34. <https://doi.org/10.1016/j.atmosenv.2018.07.021>
- Shi, Y., Lau, K. K. L., & Ng, E. (2017). Incorporating wind availability into land use regression modelling of air quality in mountainous high-density urban environment. *Environmental Research*, 157(April), 17–29. <https://doi.org/10.1016/j.envres.2017.05.007>
- Son, Y., Osornio-Vargas, Á. R., O'Neill, M. S., Hystad, P., Texcalac-Sangrador, J. L., Ohman-Strickland, P., ... Schwander, S. (2018). Land use regression models to assess air pollution exposure in Mexico City using finer spatial and temporal input parameters. *Science of the Total Environment*, 639, 40–48. <https://doi.org/10.1016/j.scitotenv.2018.05.144>
- Squizzato, S., Cazzaro, M., Innocente, E., Visin, F., Hopke, P. K., & Rampazzo, G. (2017). Urban air quality in a mid-size city — PM2.5 composition, sources and identification of impact areas: From local to long range contributions. *Atmospheric Research*, 186, 51–62. <https://doi.org/10.1016/j.atmosres.2016.11.011>
- Tam, W. W. S., Wong, T. W., Ng, L., Wong, S. Y. S., Kung, K. K. L., & Wong, A. H. S. (2014). Association between air pollution and general outpatient clinic consultations for upper respiratory tract infections in Hong Kong. *PLoS ONE*, 9(1), 1–6. <https://doi.org/10.1371/journal.pone.0086913>
- Vallero, D. (2008). *Fundamentals of Air Pollution*. (Elsevier Inc, Ed.) (4th Editio). San Diego.
- Vallero, D. (2019). *Air Pollution Calculations*. Elsevier. <https://doi.org/10.1016/b978-0-12-814934-8.09989-0>
- Van Nunen, E., Vermeulen, R., Tsai, M. Y., Probst-Hensch, N., Ineichen, A., Davey, M., ... Hoek, G. (2017). Land Use Regression Models for Ultrafine Particles in Six European Areas. *Environmental Science and Technology*, 51(6), 3336–3345. <https://doi.org/10.1021/acs.est.6b05920>

- Wallace, J. M., & Hobbs, P. V. (2006). *Atmospheric Science. An Introductory Survey*. Elsevier (Second edi). <https://doi.org/10.1007/s007690000247>
- Wang, M., Beelen, R., Eeftens, M., Meliefste, K., Hoek, G., & Brunekreef, B. (2012). Systematic evaluation of land use regression models for NO₂. *Environmental Science and Technology*, 46(8), 4481–4489. <https://doi.org/10.1021/es204183v>
- Wang, X., & Kockelman, K. M. (2009). Forecasting network data: Spatial interpolation of traffic counts using texas data. *Transportation Research Record: Journal of the Transportation Research Board*, 100–108. <https://doi.org/https://doi.org/10.3141/2105-13>
- Webster, R., & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists* (Second Edi). John Wiley & Sons, Ltd. <https://doi.org/10.2136/vzj2002.0321>
- World Health Organization. (2016). *Ambient Air Pollution: A global assessment of exposure and burden of disease*. <https://doi.org/9789241511353>
- Wu, C. Da, Chen, Y. C., Pan, W. C., Zeng, Y. T., Chen, M. J., Guo, Y. L., & Lung, S. C. C. (2017). Land-use regression with long-term satellite-based greenness index and culture-specific sources to model PM_{2.5} spatial-temporal variability. *Environmental Pollution*, 224, 148–157. <https://doi.org/10.1016/j.envpol.2017.01.074>
- Yorifuji, T., Kashima, S., Tsuda, T., Ishikawa-Takata, K., Ohta, T., Tsuruta, K. I., & Doi, H. (2013). Long-term exposure to traffic-related air pollution and the risk of death from hemorrhagic stroke and lung cancer in Shizuoka, Japan. *Science of the Total Environment*, 443, 397–402. <https://doi.org/10.1016/j.scitotenv.2012.10.088>
- Zapata, C. E., Cano, N. A., Ramírez, M., Rubiano, C., & Jiménez, J. (2015). Influence of the extreme phases of the ENSO phenomenon (El Niño and La Niña) on air quality in the Metropolitan Area of the Aburrá Valley (Colombia). *Sustainable Development*, 168(2), 663–675. <https://doi.org/10.2495/SD150582>
- Zhang, B., Liang, S., Zhao, J., Qian, Z., Bassig, B. A., Yang, R., ... Yang, S. (2016). Maternal exposure to air pollutant PM_{2.5} and PM₁₀ during pregnancy and risk of congenital heart defects. *Journal of Exposure Science and Environmental Epidemiology*, 26(4), 422–427. <https://doi.org/10.1038/jes.2016.1>
- Zhang, D., & Wang, X. (2013). *Traffic Volume Estimation using Network Interpolation Techniques: An Application on Transit Ridership in NYC Subway System*. New York. Recuperado de http://www.utrc2.org/sites/default/files/pubs/Final-Traffic-Volume-Interpolation_0.pdf
- Zhao, R., Zhan, L., Yao, M., & Yang, L. (2020). A geographically weighted regression model augmented by Geodetector analysis and principal component analysis for the spatial distribution of PM_{2.5}. *Sustainable Cities and Society*, 56(September 2019), 102106. <https://doi.org/10.1016/j.scs.2020.102106>

II. Anexos

Anexo A. Porcentaje de valores perdidos mensuales para para las estaciones de PM_{2.5}

Estación	Porcentaje de valores perdidos mensuales para PM _{2.5} (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
BEL.FEVE	3.2	0	0	0	0	0	0	3.2	0	0	0	0
COP.CVID	0	0	0	0	0	0	0	3.2	13.3	0	3.3	9.7
ENV.HOSP	0	0	3.2	6.7	0	0	0	3.2	0	3.2	0	0
EST.HOSP	0	0	0	3.3	0	0	3.2	9.7	0	0	0	0
SUR.TRAF	19.4	0	100	100	12.9	0	0	0	0	3.2	0	0
ITA.CJUS	16.1	0	0	0	6.5	0	0	0	0	0	0	19.4
ITA.CONC	12.9	0	0	0	0	13.3	0	0	6.7	3.2	0	6.5
MED.ALTA	0	0	6.5	0	0	0	0	0	0	0	3.3	12.9
MED.AРАН	0	0	0	0	0	0	0	0	0	12.9	0	9.7
MED.BEME	3.2	0	0	0	6.5	0	0	0	10.0	0	0	19.4
MED.LAYE	12.9	0	0	3.3	22.6	100	0	0	0	0	0	0
CEN.TRAF	54.8	7.1	3.2	100	100	0	0	0	13.3	6.5	3.3	0
MED.PJIC	100.0	3.6	25.8	100	19.4	3.3	6.5	9.7	0	54.8	76.7	64.5
MED.SCRI	3.2	10.7	0	0	0	3.3	0	0	0	0	0	12.9
MED.TESO	6.5	7.1	0	0	3.2	6.7	0	0	0	12.9	0	0
MED.UNNV	16.1	0	3.2	0	12.9	26.7	0	0	6.7	0	0	3.2
MED.VILL	0	0	3.2	0	6.5	0	0	3.2	13.3	0	3.3	0
SAB.RAME	100.0	100	58.1	0	6.5	3.3	0	0	0	3.2	6.7	3.2

Anexo B. Porcentaje de valores perdidos mensuales para las estaciones de temperatura

Estación	Porcentaje de valores perdidos mensuales para temperatura (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
12	0	0	0	43.3	100	100	100	100	100	100	100	100
25	0	0	0	0	0	0	0	0	26.7	83.9	53.3	0
28	0	0	0	0	0	0	0	0	3.3	0	0	12.9
38	0	0	0	0	0	0	0	0	0	3.2	36.7	6.5
48	0	0	16.1	16.7	0	0	96.8	100	100	100	100	100
59	32.3	42.9	25.8	50.0	54.8	43.3	45.2	6.5	0	0	0	0
68	0	0	0	0	0	0	0	0	0	0	0	6.5
73	0	0	0	0	0	0	0	0	0	0	0	0
83	0	0	0	0	0	0	0	0	0	0	0	9.7
197	3.2	7.1	0	0	12.9	0	0	6.5	13.3	19.4	0	0
198	16.1	0	0	0	0	0	0	0	0	0	0	0
201	0	0	0	3.3	0	0	0	19.4	3.3	9.7	0	0
202	0	0	0	0	0	0	12.9	0	0	0	0	0
203	0	0	0	0	12.9	10.0	3.2	38.7	0	0	0	0
205	0	0	3.2	0	0	0	0	0	0	0	0	0

Estación	Porcentaje de valores perdidos mensuales para temperatura (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
206	0	0	12.9	0	0	0	0	0	0	0	36.7	9.7
229	0	0	0	0	0	0	0	0	0	0	0	0
249	0	0	0	0	0	0	0	0	0	0	0	0
252	0	0	0	6.7	22.6	0	0	0	6.7	0	0	0
271	9.7	0	3.2	0	0	0	0	0	0	0	0	0
313	100	100	100	100	100	16.7	0	0	0	3.2	0	0
318	100	100	100	100	100	100	0	0	0	0	0	9.7

Anexo C. Porcentaje de valores perdidos mensuales para las estaciones de velocidad del viento

Estación	Porcentaje de valores perdidos mensuales para velocidad del viento (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
12	0	0	0	43.3	100	100	100	100	100	100	100	100
25	0	0	0	0	0	0	0	0	26.7	83.9	53.3	0
28	0	0	0	0	0	0	0	0	3.3	0	0	12.9
38	0	0	0	0	48.4	0	0	0	0	3.2	0	6.5
48	0	0	16.1	100	100	100	100	100	100	100	100	100
59	32.3	42.9	25.8	50.0	54.8	43.3	45.2	6.5	0	0	0	0
68	0	0	0	0	0	0	0	0	0	0	0	6.5
73	0	0	0	0	0	0	0	0	0	0	0	0
83	0	0	0	0	0	0	0	0	0	0	0	9.7
197	3.2	7.1	0	0	12.9	0	0	6.5	13.3	19.4	0	0
198	16.1	0	0	0	0	0	0	0	0	0	0	0
201	0	0	0	3.3	0	0	0	19.4	3.3	9.7	0	0
202	0	0	0	0	0	0	12.9	0	0	0	0	0
203	0	0	0	0	12.9	10.0	3.2	38.7	0	0	0	0
205	0	0	3.2	0	0	0	0	0	0	0	0	0
206	0	0	12.9	0	0	0	0	0	0	0	36.7	9.7
229	0	0	0	0	0	0	0	0	0	0	0	0
249	0	0	0	0	0	0	0	0	0	0	0	0
252	0	0	0	6.7	22.6	0	0	0	6.7	0	0	0
271	9.7	0	3.2	0	0	0	0	0	0	0	0	0
313	100	100	100	100	90.3	16.7	0	0	0	3.2	0	0
318	100	100	100	100	100	43.3	0	0	0	0	0	9.7

Anexo D. Porcentaje de valores perdidos mensuales para las estaciones de precipitación

Estación	Porcentaje de valores perdidos mensuales para precipitación (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
68	0	0	0	0	0	0	0	0	0	0	0	6.5
73	0	0	0	0	0	0	0	0	0	0	0	0
77	19.4	39.3	0	3.3	3.2	0	0	19.4	6.7	9.7	0	0
83	0	0	0	0	0	0	0	0	0	0	0	9.7
121	9.7	14.3	6.5	0	6.5	0	22.6	19.4	0	0	0	0
129	0	7.1	16.1	13.3	0	0	3.2	3.2	0	0	0	0
146	0	7.1	0	0	0	0	0	0	0	0	0	0
154	16.1	0	0	0	0	3.3	0	0	0	6.5	6.7	0
184	0	7.1	0	0	0	0	0	0	0	16.1	0	0
189	0	0	3.2	0	6.5	0	0	0	0	0	6.7	83.9
193	0	0	0	0	3.2	20.0	0	0	0	0	0	6.5
197	3.2	7.1	0	0	12.9	0	0	6.5	13.3	19.4	0	0
198	16.1	0	0	0	0	0	0	0	0	0	0	0
201	0	0	0	3.3	0	0	0	19.4	3.3	9.7	0	0
202	0	0	0	0	0	0	12.9	0	0	0	0	0
205	0	0	3.2	0	0	0	0	0	0	0	0	0
206	0	0	12.9	0	0	0	0	0	0	0	36.7	9.7
211	0	0	3.2	0	0	0	0	0	0	0	0	0
230	0	0	0	0	0	0	0	0	0	0	0	0
241	0	0	16.1	13.3	0	0	0	0	0	0	0	0
242	0	0	0	0	0	0	0	0	0	0	0	0
244	0	3.6	0	0	0	0	0	0	0	0	0	0
249	0	0	0	0	0	0	0	0	0	0	0	0
252	0	0	0	6.7	22.6	0	0	0	6.7	0	0	0
261	0	0	0	0	0	0	100	0	0	0	0	0
266	0	0	0	100	0	0	0	0	0	0	0	22.6
267	0	0	0	0	0	0	0	0	0	0	0	38.7
271	9.7	0	3.2	0	0	0	0	0	0	0	0	0
282	100	50.0	0	0	0	0	0	0	0	0	0	9.7

Anexo E. Porcentaje de valores perdidos mensuales para los aforos vehiculares

Estación	Porcentaje de valores perdidos mensuales para aforos vehiculares (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
58	12.90	0	6.45	13.33	3.23	0	0	0	0	35.48	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	10.00	3.23	0	0	0	0	0	0	0
56	0	0	0	3.33	0	0	0	0	0	0	0	0
74	0	0	0	0	0	0	0	0	0	0	0	0
54	0	0	0	0	0	0	0	0	0	0	0	9.68
40	0	0	0	0	0	0	0	19.36	0	0	0	0

Estación	Porcentaje de valores perdidos mensuales para aforos vehiculares (%)											
	1	2	3	4	5	6	7	8	9	10	11	12
53	0	0	0	13.33	3.23	0	0	0	0	0	0	90.32
77	0	0	0	0	0	0	0	0	50.00	100	100	100
65	0	0	0	0	0	0	0	0	0	0	0	0
51	0	0	0	3.33	0	0	0	0	0	0	0	0
18	0	0	0	13.33	3.23	0	0	0	0	0	0	0
38	0	0	3.23	13.33	3.23	3.33	0	6.45	0	0	0	0
69	0	0	0	6.67	0	0	0	0	0	0	0	0
15	0	0	0	0	38.71	0	0	0	0	0	0	0
5	0	0	0	73.33	22.58	0	0	0	0	0	0	0
27	0	0	0	6.67	3.23	0	0	0	0	80.65	100.0	100.0
72	0	0	0	10.00	9.68	0	0	0	0	0	0	0
71	0	0	0	6.67	3.23	0	0	0	0	0	0	0
73	0	0	0	6.67	3.23	0	3.23	0	0	0	0	0
70	0	0	0	13.33	3.23	0	0	0	0	0	0	0
24	0	0	0	10.00	3.23	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	100	100
31	0	0	3.23	3.33	0	0	6.45	0	0	0	0	0
48	0	0	0	0	0	0	0	0	3.33	0	0	0
9	32.26	78.57	0	0	0	0	0	0	0	3.23	0	0
35	0	0	0	6.67	0	60.00	29.03	0	0	0	33.33	0
59	0	0	0	0	3.23	93.33	58.07	0	0	0	0	0
1	0	0	0	6.67	1.61	0	0	0	10.0	0	0	0
66	0	0	0	0	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	10.0	0	0	0
12	0	3.57	3.23	0	3.23	0	0	0	0	0	0	0
19	0	0	0	13.33	3.23	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0
67	0	0	0	13.33	3.23	0	0	0	0	3.23	0	0
50	0	0	0	0	0	0	12.90	100	100	100	70	0
61	0	0	6.45	0	0	0	0	0	0	0	0	0
2	0	0	0	10.00	3.23	0	0	0	0	0	0	0
76	0	0	0	3.33	0	0	0	0	0	0	0	0
75	0	0	0	0	0	0	0	0	0	0	0	0

Anexo F. Criterios de especificación de los 72 modelos LUR construidos

#	Mes	Buffer	AdjR ²	AICc	JB	K(BP)	VIF	SA	Variables
1	1	50	0.26	71.65	0.79	0.13	1.30	0.39	-GRADT* GRADV**

#	Mes	Buffer	AdjR ²	AICc	JB	K(BP)	VIF	SA	Variables
2	2	---	---	---	---	---	---	---	---
3	3	50	0.56	-22.34	0.75	0.33	3.35	0.89	-P** -GRADP*** -DIST_VERD** -DIST_RESID** -AREA_RESID**
4	4	50	0.61	81.00	0.75	0.40	3.56	0.53	+ALTURA** +P*** +T** -GRADT** -AREA_VERD***
5	5	50	0.56	-12.94	0.68	0.42	2.24	0.64	+P*** +T** +GRADP** -AREA_VERD**
6	6	50	0.42	61.63	0.6	0.82	1	0.53	+T* -Vv**
7	7	50	0.19	62.42	0.46	0.25	1	0.62	-Vv*
8	8	50	0.19	-25.48	0.69	0.24	1	0.63	-DIST_VERD*
9	9	50	0.23	9.71	0.84	0.48	1	0.36	-Vv*
10	10	50	0.3	65.72	0.6	0.14	1	0.24	-PISO**
11	11	50	0.34	70.44	0.59	0.31	1.04	0.46	-GRADP* -DIST_VERD**
12	12	50	0.3	70.79	0.81	0.05	2.33	0.4	+P*** +T** -DIST_VERD**
13	1	0.56	69.13	0.83	0.16	1.15	0.90	0.56	-AREA_VERD** -DEN_POB* -AREA_VIA***
14	2	100	0.34	74.54	0.3	0.54	1.13	0.91	GRAD_V* DIST_VERD**
15	3	100	0.59	-31.72	0.66	0.33	1.17	0.6	-AREA_VERD* -DIST_RESID* -DEN_POB* -LONG_VIA***
16	4	100	0.53	68.88	0.42	0.18	1.19	0.8	+T** +DIST_VERD** -DIST_RESID**
17	5	100	0.49	-16.83	0.56	0.37	1.15	0.55	-V* GRADV* DIST_VERD***
18	6	100	0.59	61.46	0.62	0.65	1.41	0.41	-PISO*** -Vv** +AREA_RESID*
19	7	100	0.65	64.91	0.83	0.04	1.74	0.58	-PISO* -Vv*** +AREA_RESID* -LONG_VIA**
20	8	100	0.44	-27	0.22	0.49	1.44	0.3	-AREA_VIA*** +FLUJO_VEHI**

#	Mes	Buffer	AdjR ²	AICc	JB	K(BP)	VIF	SA	Variables
21	9	100	0.58	9.2	0.77	0.07	1.42	0.59	-PISO** -Vv*** +AREA_RESID**
22	10	100	0.34	67.92	0.93	0.44	1.91	0.38	-AREA_VIA** +FLUJO_VEHI**
23	11	100	0.66	66.01	0.41	0.13	1.46	0.32	-PISO*** -Vv*** +AREA_RESID**
24	12	100	0.61	63.15	0.87	0.08	1.42	0.56	-PISO*** -Vv*** AREA_RESID**
25	1	150	0.38	69.33	0.68	0.91	1.14	0.29	-DIST_RESID** -AREA_RESID**
26	2	150	0.3	79.53	0.28	0.63	2.04	0.97	GRADV* DIST_VERD** -DEN_POB*
27	3	150	0.5	-29.21	0.73	0.21	1.84	0.97	-ALTURA* +DIST_VERD** -DEN_POB** +FLUJO_VEHI**
28	4	150	0.66	70.41	0.52	0.68	1.44	0.69	+P*** -GRADT*** -AREA_VERD*** +FLUJO_VEHI***
29	5	150	0.54	-18.04	0.68	0.89	3.12	0.65	-GRADT*** +GRADV*** +FLUJO_VEHI***
30	6	150	0.69	57.59	0.81	0.45	1.07	0.8	-PISO** -Vv*** -DIST_RESID**
31	7	150	0.47	60.08	0.8	0.11	1.05	0.83	-Vv** -DIST_RESID**
32	8	150	0.40	-26.21	0.62	0.42	1.01	0.25	-PISO** +FLUJO_VEHI*
33	9	150	0.59	8.65	0.74	0.24	1.75	0.91	-PISO*** -Vv*** +AREA_RESID**
34	10	150	0.63	70.41	0.77	0.44	1.51	0.71	-PISO*** -GRADP* -DIST_RESID** +NUM_VIA*
35	11	150	0.62	67.60	0.37	0.17	1.80	0.98	-PISO*** -Vv** +AREA_RESID**
36	12	150	0.64	62.39	0.53	0.18	1.74	0.99	-PISO*** -Vv*** +AREA_RESID**
37	1	200	0.26	71.65	0.79	0.13	1.30	0.39	-GRADT* GRADV**
38	2	200	0.33	74.76	0.69	0.33	1.25	0.55	-GRADT* FLUJO_VEHI**

#	Mes	Buffer	AdjR ²	AICc	JB	K(BP)	VIF	SA	Variables
39	3	200	0.33	74.76	0.69	0.33	1.25	0.55	-GRADT* +FLUJO_VEHI**
40	4	200	0.42	71.44	0.86	0.75	1.30	0.64	-GRADT** -AREA_VERD*** +FLUJO_VEHI***
41	5	200	0.41	-14.85	0.55	0.54	2.17	0.88	+GRADP** -GRADT** +DIST_VERD**
42	6	200	0.43	61.53	0.74	0.59	1.03	0.55	-PISO* -Vv*
43	7	200	0.65	64.62	0.73	0.05	1.49	0.98	-Vv*** +AREA_RESID** -DEN_POB*** +DIST_VIA***
44	8	200	0.65	-23.03	0.82	0.22	1.7	0.62	-PISO*** -Vv** +AREA_RESID** +FLUJO_VEHI*
45	9	200	0.58	9.16	0.89	0.16	1.74	0.53	-PISO** -Vv** +AREA_RESID**
46	10	200	0.54	67.58	0.31	0.31	1.98	0.14	-PISO*** -GRADP** +AREA_VIA**
47	11	200	0.46	67.92	0.74	0.10	1.09	0.56	-PISO* -V**
48	12	200	0.62	62.93	0.85	0.15	1.70	0.62	-PISO*** -Vv** +AREA_RESID**
49	1	300	0.54	84.22	0.62	0.58	1.93	0.28	+GRADP** -GRADT*** DIST_VERD*** -DEN_POB** +AREA_VIA*
50	2	300	0.73	73.18	0.55	0.22	3.31	0.51	+P* -GRADT** -AREA_VERD*** +FLUJO_VEHI***
51	3	300	0.43	-27.47	0.77	0.07	4.89	0.33	-Vv** +DIST_RESID*** +LONG_VIA* -NUM_VIA***
52	4	300	0.43	77.10	0.74	0.81	2.68	0.37	+P** -GRADT** -AREA_VERD** +FLUJO_VEHI**
53	5	300	0.41	-14.85	0.55	0.54	2.17	0.88	+GRADP** -GRADT** +DIST_VERD**
54	6	300	0.46	64.89	0.53	0.44	1.26	0.6	-PISO*** -DIST_RESID*

#	Mes	Buffer	AdjR ²	AICc	JB	K(BP)	VIF	SA	Variables
									+AREA_VIA*
55	7	300	0.6	60.67	0.62	0.14	1.38	0.21	-PISO** -Vv** -DIST_RESID**
56	8	300	0.79	-29.78	0.3	0.08	1.41	0.76	-PISO*** -Vv*** -DIST_RESID*** +FLUJO_VEHI***
57	9	300	0.53	10.65	0.75	0.34	1.37	1	-PISO** -Vv** -DIST_RESID*
58	10	300	0.53	67.92	0.55	0.46	1.58	0.04	-PISO*** -GRADP* +AREA_VIA**
59	11	300	0.57	69.01	0.65	0.22	1.38	0.76	-PISO** -Vv** -DIST_RESID*
60	12	300	0.58	64.22	0.77	0.09	1.32	0.65	-PISO*** -Vv*** +AREA_RESID**
61	1	500	0.56	75.12	0.80	0.12	2.60	0.86	-ALTURA** -V*** GRADP** AREA_VERD***
62	2	500	0.32	75.01	0.73	0.21	1.17	0.73	-GRADT* FLUJO_VEHI**
63	3	500	0.5	-29.21	0.92	0.16	1.29	0.64	+DIST_IND** -AREA_IND** -NUM_VIA* +FLUJO_VEHI**
64	4	500	0.54	74.41	0.81	0.39	3.44	0.77	+P*** -Vv*** +GRADP** -AREA_IND***
65	5	500	0.34	-13.38	0.7	0.13	2.5	0.84	+GRADP* -GRADT** +GRADV*
66	6	500	0.54	62.9	0.97	0.31	1.05	0.14	-PISO** -Vv* +AREA_VIA*
67	7	500	0.44	64.89	0.58	0.35	1.66	0.66	-PISO** -DIST_RESID** +AREA_VIA*
68	8	500	0.42	-26.66	0.8	0.71	1.01	0.15	-PISO** +FLUJO_VEHI*
69	9	500	0.49	11.57	0.54	0.99	1.67	0.45	-PISO*** +AREA_RESID** +AREA_VIA**
70	10	500	0.5	64.34	0.74	0.52	1	0.18	-PISO** +LONG_VIA**
71	11	500	0.54	70.01	0.38	0.25	1.35	0.38	-PISO***

#	Mes	Buffer	AdjR ²	AICc	JB	K(BP)	VIF	SA	Variables
									-DIST_RESID* +LONG_VIA*
72	12	500	0.54	65.36	0.71	0.19	1.25	0.80	-PISO*** -Vv** AREA_RESID***

Significancia de las variables en el modelo (*0.10, **0.05, ***0.01)

Anexo G. Estadístico de error ARE para cada sitio de monitoreo de PM_{2.5}

Mes	Datos	ID del sitio de monitoreo												
		0	1	2	3	4	5	6	7	8	9	10	11	12
1	Oi	17.71	15.65	18.22	20.67	22.78	18.60	25.84	21.00	21.04	22.21	17.69	16.15	23.78
	Ei	19.70	19.39	20.93	19.95	23.61	19.24	23.85	17.88	19.49	21.59	18.26	18.27	19.19
	ARE	0.11	0.24	0.15	0.04	0.04	0.03	0.08	0.15	0.07	0.03	0.03	0.13	0.19
2	Oi	25.17	27.11	28.50	34.71	33.29	27.26	35.14	28.10	28.93	28.00	28.22	23.30	32.66
	Ei	29.58	25.37	29.57	31.77	32.83	31.55	30.95	29.72	26.99	28.41	26.52	26.07	31.06
	ARE	0.18	0.06	0.04	0.08	0.01	0.16	0.12	0.06	0.07	0.01	0.06	0.12	0.05
3	Oi	1.43	1.43	1.46	1.58	1.55	1.46	1.50	1.45	1.48	1.47	1.45	1.39	1.53
	Ei	1.44	1.43	1.51	1.50	1.53	1.51	1.48	1.48	1.49	1.46	1.44	1.43	1.47
	ARE	0.01	0.00	0.04	0.05	0.01	0.04	0.02	0.02	0.01	0.01	0.01	0.03	0.04
4	Oi	15.92	14.84	17.83	23.22	22.81	17.17	21.87	19.29	20.16	20.81	17.22	17.10	23.60
	Ei	18.53	14.84	20.89	20.54	23.23	17.82	20.91	20.94	20.38	19.80	16.81	17.47	19.68
	ARE	0.16	0.00	0.17	0.12	0.02	0.04	0.04	0.09	0.01	0.05	0.02	0.02	0.17
5	Oi	1.16	1.05	1.18	1.44	1.34	1.19	1.34	1.23	1.22	1.27	1.18	1.13	1.32
	Ei	1.12	1.16	1.19	1.42	1.29	1.14	1.24	1.22	1.19	1.26	1.30	1.20	1.30
	ARE	0.03	0.10	0.01	0.01	0.04	0.04	0.07	0.00	0.03	0.01	0.11	0.06	0.01
6	Oi	11.57	10.20	12.44	12.79	18.30	13.84	14.51	14.41	12.12	16.47	12.29	10.52	15.61
	Ei	10.95	11.15	12.82	13.31	18.05	11.43	14.80	13.53	13.61	15.50	12.85	13.29	13.77
	ARE	0.05	0.09	0.03	0.04	0.01	0.17	0.02	0.06	0.12	0.06	0.05	0.26	0.12
7	Oi	13.04	11.16	13.27	14.57	19.11	15.39	14.53	16.01	12.62	17.11	13.37	11.33	16.38
	Ei	12.95	13.44	14.14	12.06	18.59	14.90	13.65	16.56	14.54	15.63	14.12	12.84	14.54
	ARE	0.01	0.20	0.07	0.17	0.03	0.03	0.06	0.03	0.15	0.09	0.06	0.13	0.11
8	Oi	1.16	1.08	1.14	1.19	1.35	1.19	1.18	1.21	1.10	1.25	1.14	1.05	1.25
	Ei	1.15	1.12	1.19	1.18	1.34	1.11	1.19	1.21	1.10	1.25	1.17	1.06	1.24
	ARE	0.01	0.04	0.04	0.01	0.01	0.06	0.01	0.00	0.00	0.01	0.03	0.00	0.01
9	Oi	3.88	3.74	4.05	4.16	4.77	4.25	4.11	4.19	4.01	4.58	4.06	3.74	4.53
	Ei	4.06	3.77	4.30	4.14	4.72	3.91	4.02	4.33	4.12	4.39	3.95	3.96	4.38
	ARE	0.05	0.01	0.06	0.00	0.01	0.08	0.02	0.03	0.03	0.04	0.03	0.06	0.03
10	Oi	14.25	12.75	15.42	14.63	21.47	15.80	16.76	18.51	16.67	19.58	14.70	13.13	20.73
	Ei	14.50	13.11	16.25	16.10	19.39	13.24	18.81	15.80	17.64	18.29	15.46	16.21	19.57
	ARE	0.02	0.03	0.05	0.10	0.10	0.16	0.12	0.15	0.06	0.07	0.05	0.23	0.06
11	Oi	15.63	13.68	16.96	18.06	23.71	18.64	17.93	20.72	18.67	22.79	16.01	14.83	21.69
	Ei	16.75	15.86	16.39	18.91	23.54	15.00	18.90	19.62	19.74	20.66	17.26	18.04	18.74

Mes	Datos	ID del sitio de monitoreo												
		0	1	2	3	4	5	6	7	8	9	10	11	12
	ARE	0.07	0.16	0.03	0.05	0.01	0.19	0.05	0.05	0.06	0.09	0.08	0.22	0.14
12	Oi	17.57	15.58	18.09	21.47	23.84	20.25	17.88	21.55	18.76	20.55	17.46	15.71	22.63
	Ei	19.29	17.34	18.94	20.66	23.65	17.54	17.20	20.66	18.42	22.13	17.59	16.63	21.32
	ARE	0.10	0.11	0.05	0.04	0.01	0.13	0.04	0.04	0.02	0.08	0.01	0.06	0.06