



UNIVERSIDAD SANTO TOMÁS
PRIMER CLAUSTRO UNIVERSITARIO DE COLOMBIA

Relación entre la contaminación por
material particulado (PM_{10} y $PM_{2,5}$) y
la tasa de mortalidad por
enfermedades respiratorias en Bogotá
D.C (2018-2022): Un análisis
considerando modelos
espacio-temporales

María Isabel Guzmán Murcia
Wilmer Alfonso González Rivera

Universidad Santo Tomás
Facultad de Estadística
División de Ciencias Económicas y Administrativas
Bogotá, D.C., Colombia
2024

Relación entre la contaminación por material particulado (PM_{10} y $PM_{2,5}$) y la tasa de mortalidad por enfermedades respiratorias en Bogotá D.C (2018-2022): Un análisis considerando modelos espacio-temporales

María Isabel Guzmán Murcia
Wilmer Alfonso González Rivera

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Magister en Estadística Aplicada

Director:
Andrés Felipe Ortiz Rico
Estadístico
Magister en Ciencias – Estadística
Candidato a Doctor en Ciencias – Estadística

Línea de Investigación:
Modelamiento y econometría
Grupo de Investigación:
USTAdística

Universidad Santo Tomás
Facultad de Estadística
División de Ciencias Económicas y Administrativas
Bogotá, D.C., Colombia
2024

Agradecimientos

En primer lugar, agradecemos a nuestro director de tesis Andrés Felipe Ortiz Rico, por su constante orientación, paciencia y valiosos comentarios que nos guiaron durante cada etapa del proceso. Su experiencia y dedicación fueron fundamentales para el desarrollo de esta investigación. A la Universidad Santo Tomás, en especial a la Facultad de Estadística, por brindarnos los recursos académicos y el espacio necesario para realizar esta investigación.

Agradecemos también a la Red de Monitoreo de Calidad del Aire de Bogotá (RMCAB) de la Secretaría Distrital de Ambiente y al Observatorio de Salud de Bogotá (SaluData) de la Secretaría Distrital de Salud, por facilitar los datos que fueron esenciales para llevar a cabo el análisis de este estudio.

Finalmente, a nuestros familiares y amigos quienes con su apoyo incondicional nos brindaron la motivación y la fortaleza para culminar este proyecto. A nuestros padres, quienes nos enseñaron el valor del esfuerzo y la dedicación.

Resumen

La contaminación del aire representa un riesgo significativo para la salud humana, pues este tipo de externalidades afectan directamente la calidad de vida de millones de seres humanos alrededor del mundo, es así como la comunidad científica respalda ampliamente la relación entre la contaminación y las enfermedades respiratorias. Según la Organización Mundial de la Salud (OMS), en 2019, “aproximadamente el 37 % de las muertes prematuras relacionadas con la contaminación del aire exterior se debieron a cardiopatías isquémicas y accidentes cerebrovasculares, el 18 % y el 23 % a enfermedades pulmonares obstructivas crónicas e infecciones respiratorias agudas, respectivamente, y el 11 % a cáncer de las vías respiratorias” [Organización Mundial de la Salud, 2022]. Por tanto, resulta relevante analizar la relación existente entre la contaminación del aire por materia particulada (PM_{10} y $PM_{2,5}$) y la tasa de mortalidad por enfermedades respiratorias agudas en la ciudad de Bogotá D.C para el periodo 2018-2022, generando un punto de referencia que permita visibilizar y mitigar esta problemática.

Para ello, se utilizaron modelos espaciales y espacio-temporales, que combinan la dimensión espacial y temporal para analizar datos geográficamente dispersos y registrados en diferentes momentos del tiempo, estos modelos permiten capturar la variación espacial y temporal de la contaminación del aire y la mortalidad, considerando la autocorrelación espacial y temporal, las tendencias a lo largo del tiempo, la predicción de valores futuros y la interacción entre variables en el espacio y el tiempo. Los hallazgos, sugieren la necesidad de implementar medidas efectivas para controlar la contaminación del aire y reducir la mortalidad por enfermedades respiratorias, destacando la relevancia de considerar la ubicación geográfica y los periodos de tiempo en la planificación de intervenciones para mitigar los impactos negativos de la contaminación atmosférica en la salud de los habitantes de Bogotá.

Palabras clave: Contaminación del aire, material particulado, enfermedades respiratorias agudas, modelos espacio-temporales.

Abstract

Air pollution represents a significant risk to human health, as such externalities directly affect the quality of life of millions of human beings around the world, thus, the scientific community widely supports the link between pollution and respiratory diseases. According to the World Health Organization (WHO), in 2019, “approximately 37 % of premature deaths related to outdoor air pollution were due to ischaemic heart disease and stroke, 18 % and 23 % to chronic obstructive pulmonary diseases and acute respiratory infections, respectively, and 11 % to respiratory tract cancer” [World Health Organization, 2022]. Therefore, it is relevant to analyse the relationship between air pollution by particulate matter (PM_{10} and $PM_{2.5}$) and the mortality rate due to acute respiratory diseases in the city of Bogotá D.C. for the period 2018-2022, generating a reference point that allows us to visualise and mitigate this problem.

For this purpose, spatial and spatio-temporal models, which combine the spatial and temporal dimension to analyse geographically dispersed data recorded at different points in time, were used, these models allow capturing the spatial and temporal variation of air pollution and mortality, considering spatial and temporal autocorrelation, trends over time, prediction of future values and the interaction between variables in space and time. The findings suggest the need to implement effective measures to control air pollution and reduce mortality from respiratory diseases, highlighting the relevance of considering geographic location and time periods in planning interventions to mitigate the negative impacts of air pollution on the health of Bogotá’s inhabitants.

Keywords: Air pollution, particulate matter, acute respiratory diseases, spatiotemporal models.

Índice general

Agradecimientos	v
Resumen	vii
Lista de figuras	xiii
Lista de tablas	1
Introducción	1
1. Objetivos	7
1.1. Objetivo General	7
1.2. Objetivos Específicos	7
2. Marco Teórico	8
2.1. Material Particulado (<i>PM</i>)	8
2.1.1. Red de Monitoreo de Calidad de Aire de Bogotá – RMCAB	9
2.2. Enfermedades-Infecciones Respiratorias Agudas (IRA)	10
2.2.1. Observatorio de Salud de Bogotá - SaluData	10
2.2.2. Clasificación Internacional de Enfermedades (CIE)	11
2.3. Series de tiempo	11
2.3.1. Proceso Autorregresivo (AR)	12
2.3.2. Proceso de Medias Móviles (MA)	13
2.3.3. Proceso Autorregresivo de Medias Móviles (ARMA)	14
2.4. Estadística Espacial	14
2.5. Interpolación Espacial - Predicción Espacial Óptima	19
2.6. Econometría Espacial	22
2.7. Modelos Espaciales y Espacio-Temporales	23
3. Metodología	29
3.1. Descripción del área de estudio	29
3.2. Fuentes de Información	29
3.2.1. Concentración de Material Particulado (<i>PM</i>)	29
3.2.2. Mortalidad por Infecciones Respiratorias Agudas (IRA)	32
3.2.3. Tasa de Mortalidad por IRA	33

3.3. Metodología Desarrollada	33
3.3.1. Fase 1. Depurar y estructurar las bases de datos	33
3.3.2. Fase 2. Realizar un análisis exploratorio y espacial	34
3.3.3. Fase 3. Desarrollar métodos de interpolación espacial	35
3.3.4. Fase 4. Identificar y estandarizar los datos de mortalidad	36
3.3.5. Fase 5. Estimar un modelo espacio-temporal	36
3.3.6. Fase 6. Documentar la relación entre las variables	37
4. Análisis exploratorio de los datos	38
4.1. Concentración por Material Particulado (PM)	38
4.1.1. Descripción de datos por hora de Material Particulado (PM)	38
4.1.2. Descripción de datos mensuales de Material Particulado (PM)	43
4.2. Mortalidad por Enfermedades-Infecciones Respiratorias Agudas (IRA)	49
4.2.1. Exploración inicial de los datos	49
4.3. Relación entre Material Particulado (PM) y Tasa de Mortalidad por IRA	53
5. Análisis de tendencias temporales y patrones de comportamiento de los niveles de materia particulada (PM_{10} y $PM_{2,5}$) en Bogotá D.C	54
5.1. Análisis series de tiempo	54
5.1.1. Función de Autocorrelación:	56
5.1.2. Estimación de los coeficientes:	58
5.1.3. Backtesting:	62
5.1.4. Pronósticos y Raíz del Error Cuadrático Medio:	63
6. Interpolación espacial de la concentración de material particulado (PM_{10} y $PM_{2,5}$) en las localidades de Bogotá D.C	65
6.1. Análisis descriptivo de los datos	65
6.2. Análisis geostadístico de los datos	68
6.2.1. Análisis de la tendencia	69
6.2.2. Ajuste de modelo para remover la tendencia	70
6.3. Interpolación espacial	73
6.3.1. Construcción de la grilla a interpolar	73
6.3.2. Interpolación espacial de la Materia Particulada PM_{10} - Cortes 2022	76
6.3.3. Interpolación espacial de la Materia Particulada $PM_{2,5}$ - Cortes 2022	79
7. Estimación de modelos espaciales y espacio temporales	84
7.1. Visualización de la Capa Geográfica de la Tasa de Mortalidad	84
7.2. Modelos Espaciales SAR y SEM	86
7.3. Modelo Espacio - Temporal STFDF	91
7.3.1. Resultados Modelo Espacio - Temporal STFDF	93

8. Conclusiones y recomendaciones	97
A. Anexo A: Promedio mensual de material particulado PM_{10}	100
B. Anexo B: Promedio mensual de material particulado $PM_{2,5}$	101
C. Anexo C: Promedio de medición de material particulado PM_{10}-Interpolada	102
D. Anexo D: Promedio de medición de material Particulado $PM_{2,5}$-Interpolada	103
E. Anexo E: Clasificación Códigos CIE - Relacionados con IRA	104
Bibliografía	107

Lista de Figuras

3-1. Localidades de Bogotá D.C	30
4-1. Diagrama de caja concentración Materia Particulada PM_{10} por hora por localidad en Bogotá (2018-2022)	40
4-2. Diagrama de caja concentración Materia Particulada $PM_{2,5}$ por hora por localidad en Bogotá (2018-2022)	41
4-3. Mapa de calor PM_{10} por trimestre	42
4-4. Mapa de calor $PM_{2,5}$ por trimestre	43
4-5. Diagrama de caja concentración Materia Particulada PM_{10} mensual por localidad en Bogotá (2018-2022)	44
4-6. Diagrama de caja concentración Materia Particulada $PM_{2,5}$ mensual por localidad en Bogotá (2018-2022)	45
4-7. Histograma de Frecuencia PM_{10}	46
4-8. Histograma de Frecuencia $PM_{2,5}$	46
4-9. Mapa de calor PM_{10} por trimestre	47
4-10. Mapa de calor $PM_{2,5}$ por trimestre	47
4-11. Diagramas de caja tasa de mortalidad por IRA por localidades en Bogotá (2018-2022)	50
4-12. Mapa de Calor Régimen vs Rango Etario	50
4-13. Mapa de Calor Trimestres vs Rango Etario	51
4-14. Mapa de Calor Trimestres vs Localidad	52
4-15. Relación Material Particulado (PM_{10}) y Tasa de Mortalidad por IRA	53
4-16. Relación Material Particulado ($PM_{2,5}$) y Tasa de Mortalidad por IRA	53
5-1. Serie de tiempo de la materia particulada PM_{10} y $PM_{2,5}$	55
5-2. Month Plot PM_{10} y $PM_{2,5}$	55
5-3. Diagramas de caja Month Plot PM_{10} y $PM_{2,5}$	56
5-4. Función de Autocorrelación Parcial	57
5-5. Backtesting PM_{10}	62
5-6. Backtesting $PM_{2,5}$	63
5-7. Pronósticos de los modelos PM_{10}	64
5-8. Pronósticos de los modelos $PM_{2,5}$	64
6-1. Descripción del Logaritmo de la Materia Particulada PM_{10} y $PM_{2,5}$	66

6-2. Histograma del Logaritmo de la Materia Particulada PM_{10} y $PM_{2,5}$	67
6-3. Gráficos del Logaritmo de la Materia Particulada PM_{10}	68
6-4. Gráficos del Logaritmo de la Materia Particulada $PM_{2,5}$	69
6-5. Gráficos de dispersión contra la longitud y latitud del Logaritmo de PM_{10} .	70
6-6. Gráficos de dispersión contra la longitud y latitud del Logaritmo de $PM_{2,5}$.	70
6-7. Gráfico de dispersión contra la longitud y latitud de los residuales del modelo para el logaritmo de PM_{10}	71
6-8. Gráficos de dispersión contra la longitud y latitud de los residuales del modelo para el logaritmo de $PM_{2,5}$	71
6-9. Visualización grilla a interpolar localidades de Bogotá D.C	73
6-10. Concentración del logaritmo de PM_{10} y $PM_{2,5}$ año 2022 en Bogotá D.C . . .	74
6-11. Concentración del logaritmo de PM_{10} cortes año 2022 en Bogotá D.C	75
6-12. Concentración del logaritmo de $PM_{2,5}$ cortes año 2022 en Bogotá D.C	76
6-13. Interpolación para el Logaritmo de la Materia Particulada PM_{10} cortes año 2022	78
6-14. Interpolación para el Logaritmo de la Materia Particulada PM_{10} cortes año 2022 por localidades	79
6-15. Interpolación para la Materia Particulada PM_{10} cortes año 2022 por localidades	79
6-16. Interpolación para el Logaritmo de la Materia Particulada $PM_{2,5}$ cortes año 2022	81
6-17. Interpolación para el Logaritmo de la Materia Particulada $PM_{2,5}$ cortes año 2022 por localidades	82
6-18. Interpolación para la Materia Particulada $PM_{2,5}$ cortes año 2022 por localidades	83
7-1. Tasa de Mortalidad en las localidades de Bogotá D.C año 2022	85
7-2. Tasa de Mortalidades en las localidades de Bogotá D.C cortes año 2022 . . .	85
7-3. Modelo Espacio - Temporal Material Particulado PM_{10}	93
7-4. Modelo Espacio - Temporal Material Particulado $PM_{2,5}$	93
7-5. Modelo Espacio Temporal Tasa de Mortalidad por IRA	94

Lista de Tablas

2-1. Niveles Máximos Permisibles de Materia Particulada (PM) a nivel mundial .	9
2-2. Niveles Máximos Permisibles de Materia Particulada (PM) a nivel nacional .	9
3-1. Estaciones de la RMCAB por localidades	31
3-2. Porcentaje de datos faltantes PM_{10} y $PM_{2,5}$	35
4-1. Base de datos concentración de Materia Particulada PM_{10} y $PM_{2,5}$ en Bogotá D.C (2018-2022)	39
4-2. Resumen de los datos de Materia Particulada (PM) a nivel distrital	39
4-3. Resumen de los datos mensuales de Materia Particulada (PM) a nivel distrital	44
4-4. Resumen datos promedios de Materia Particulada (PM) a nivel distrital . .	48
4-5. Base de datos consolidada por localidades	48
4-6. Registro Mensual de mortalidad por IRA 2018-2022 en Bogotá	49
5-1. Salida de Materia particulada PM_{10} ARIMA	58
5-2. Salida de Materia particulada $PM_{2,5}$ ARIMA	60
6-1. Resumen concentración de PM_{10} y $PM_{2,5}$ en Bogotá D.C (2018-2022)	66
6-2. Anova modelo para los residuales del modelo de primer orden con la longitud del logaritmo de PM_{10}	72
6-3. Anova modelo para los residuales del modelo de primer orden con la longitud del logaritmo de $PM_{2,5}$	72
6-4. Resultados validación cruzada métodos inverso de la distancia y modelos lineales de primer y segundo orden PM_{10}	77
6-5. Resultados validación cruzada métodos inverso de la distancia y modelos lineales de primer y segundo orden $PM_{2,5}$	80
7-1. Pruebas de significancia modelo SAR	87
7-2. Pruebas de significancia modelo SEM	88
7-3. Coeficientes Modelo Cortes (Enero-Abri-Julio-Octubre)	89
7-4. Coeficientes Modelo STFDF	95

Introducción

La contaminación del aire es un tema de creciente relevancia en la salud pública y en las políticas gubernamentales, debido a su significativo riesgo para el medio ambiente y la salud humana. En septiembre del 2021, la Organización Mundial de la Salud (OMS) publicó las nuevas directrices mundiales sobre la calidad del aire, evidenciando el aumento de los niveles de contaminación y sus efectos negativos en la salud humana [Organización Mundial de la Salud, 2022].

En la actualidad, se evalúan a nivel global seis contaminantes clave para medir la calidad del aire: partículas en suspensión (PM_{10} y $PM_{2,5}$), ozono (O_3), dióxido de nitrógeno (NO_2), dióxido de azufre (SO_2) y monóxido de carbono (CO). Según declaraciones de la OMS, se estima que la exposición a estos contaminantes causa alrededor de 7 millones de muertes prematuras anuales, afectando especialmente a los niños con reducciones en el crecimiento, la función pulmonar e infecciones respiratorias, particularmente en países de ingresos medios y bajos, donde el crecimiento de la contaminación atmosférica ha sido más pronunciado debido al proceso de industrialización [Organización Mundial de la Salud, 2022].

En consecuencia, Colombia y en particular la ciudad de Bogotá, no escapa a esta problemática. En septiembre de 2021, la alcaldesa Claudia López firmó el Decreto 332 de 2021, adoptando el Plan Estratégico para la Gestión Integral de la Calidad del Aire de Bogotá 2030. Este plan busca reducir los niveles de contaminación atmosférica, reconociendo al transporte y la industria como los principales emisores de contaminantes, destacándose las localidades de Fontibón, Puente Aranda, Kennedy, Engativá, Tunjuelito y Usaquén como aquellas que cuentan con mayor presencia de fuentes fijas de combustión [Alcaldía de Bogotá, 2021].

Es esencial destacar que, la concentración de material particulado (PM), considerado uno de los contaminantes más perjudiciales para la salud humana, debido a que puede ser inhalado profundamente en los pulmones, provocando inflamación y deterioro de la función pulmonar, intensificando las condiciones respiratorias preexistentes y aumentando el riesgo de infecciones graves, ha experimentado un aumento significativo en la ciudad de Bogotá en los últimos años. Este tipo de contaminación, clasificada como carcinógena por la OMS, está estrechamente relacionada con el aumento de enfermedades respiratorias agudas en la ciudad; “aunque las infecciones respiratorias superiores son más frecuentes, las infecciones respiratorias bajas [inferiores] son responsables de los cuadros más graves de influenza, neumonía y bronquiolitis, y son las que más contribuyen a la mortalidad por Infección Respiratoria

Aguda (IRA), [estas infecciones se definen como cualquier enfermedad infecciosa del tracto respiratorio superior o inferior; para el primer caso, se afectan la nariz, senos paranasales, faringe y laringe, e incluyen el resfriado común, laringitis, faringitis/tonsilitis, rinitis aguda, rinosinusitis aguda y otitis media aguda; mientras que para el segundo caso, se afectan la tráquea, bronquios y pulmones incluyendo la bronquitis aguda, bronquiolitis, neumonía y traqueítis]. En los países en desarrollo se estima que, en los niños menores de dos años, el 2 % al 3 % tuvo neumonía grave que requirió hospitalización y las tasas de mortalidad por IRA oscilan entre 60 a 100 casos por 1.000 niños menores de 5 años” [Alcaldía de Bogotá, 2019].

En consecuencia, se destaca en los últimos años una tendencia creciente en el desarrollo de investigaciones y estudios relacionados con la contaminación por material particulado (PM) y su relación con las enfermedades y morbilidad respiratoria, como respuesta a la creciente necesidad de diseñar y desarrollar políticas públicas dirigidas a la reducción de emisiones de estos contaminantes que contribuya a la protección de la salud pública. A nivel distrital (Bogotá), se pueden destacar investigaciones como:

- **Análisis de series de tiempo de contaminantes atmosféricos en la ciudad de Bogotá a partir del desarrollo de modelos estadísticos ARIMA:** esta investigación pretendía por medio del análisis de series de tiempo de los contaminantes atmosféricos PM_{10} , O_3 , SO_2 , NO_2 y $PM_{2,5}$, en la ciudad de Bogotá durante los años 2007 y 2017, pronosticar el estado de la calidad del aire y con ello poder determinar periodos de tiempo de prevención, alerta y emergencia. Entre las principales conclusiones se obtuvo que todas las estaciones analizadas sobrepasaron los límites anuales establecidos por la Organización Mundial de la Salud (OMS), respecto al material particulado, de igual forma las partículas $PM_{2,5}$ y PM_{10} presentan una correlación fuerte, principalmente debido a que a pesar de su tamaño, son generadas por la misma fuente de emisión [Pinzón Hassan and Tique Ortiz, 2019].
- **Relación entre $PM_{2,5}$ y PM_{10} en la ciudad de Bogotá:** este artículo presenta los principales resultados de un análisis realizado a las mediciones de estos dos contaminantes durante dos meses en varias estaciones de la Red de Monitoreo de Calidad del Aire de la Ciudad de Bogotá. Los resultados muestran la existencia de una correlación positiva y lineal entre los dos contaminantes, por tanto se concluye que si se tienen en cuenta los registros de PM_{10} se puede mejorar la predicción de la concentración de $PM_{2,5}$. Por otro lado, se resalta la importancia de seguir ahondado en la relación de estos dos contaminantes que pueda ayudar a entender y predecir mejor su comportamiento en el tiempo [Rojas and Galvis, 2005].
- **Contaminación del aire y enfermedad respiratoria en menores de cinco años de Bogotá, 2007:** este artículo presenta los resultados del estudio el cual pretendía determinar la relación existente entre la contaminación del aire por material particulado

y los síntomas respiratorios en niños menores de 5 años en tres localidades de Bogotá. Finalmente se encontró que un niño menor de 5 años expuesto a materia particulada gruesa tiene 1,70 veces más riesgo de presentar ausentismo escolar por enfermedad respiratoria aguda, por consiguiente, una disminución de la exposición a este contaminante podría reducir en un 41,1 % el ausentismo escolar a causa de dicha enfermedad [Hernández Florez et al., 2013].

- **¿Cuál es la relación existente entre las concentraciones de material particulado y las admisiones a centros médicos por enfermedades respiratorias agudas en Bogotá?:** este artículo presenta los resultados de una investigación realizada como tesis de maestría, en donde se buscaba encontrar la relación entre las emisiones diarias del material particulado y las admisiones en los hospitales por enfermedades respiratorias agudas, en la ciudad de Bogotá en el periodo 2016 y 2017. En este sentido, se concluyó que por cada aumento de $10 \mu g/m^3$ en la medición diaria de materia particulada fina, las asistencias a los centros médicos de niños menores de 5 años por enfermedades respiratorias aumentaban en promedio 0,0434 [Quintero Forero, 2020].

Así mismo, a nivel nacional se pueden destacar las siguientes investigaciones:

- **Relación entre las partículas finas ($PM_{2,5}$) y respirables (PM_{10}) en la ciudad de Medellín:** este artículo presenta los resultados del proyecto denominado “Patologías respiratorias en niños preescolares y su relación con la contaminación atmosférica de Medellín”, el cual fue realizado por la Universidad de Medellín y la Universidad CES para la Secretaria de Salud de Medellín en el año 2007. Para este estudio, se realizaron mediciones de las partículas finas y gruesas en el período de febrero a octubre de 2007, obteniendo que existe una correlación positiva y lineal entre las dos partículas [Echeverri Londoño and Maya Vasco, 2008].
- **Contaminación por material particulado ($PM_{2,5}$ y PM_{10}) y consultas por enfermedades respiratorias en Medellín (2008-2009):** en este artículo se utilizaron modelos de series de tiempo con el fin de analizar la relación positiva entre la contaminación por material particulado y las consultas externas y por urgencia debidas a enfermedades respiratorias como asma, bronquitis, infecciones y rinitis, para ello se estimaron diez modelos por mínimos cuadrados ordinarios, en donde se concluyó que la contaminación del aire por material particulado es un problema de salud pública en Medellín, ya que afecta directamente en la cantidad de consultas por estas enfermedades. Finalmente, los autores sugieren la necesidad de analizar otros efectos que estos contaminantes tienen en el ser humano, como las tasas de morbilidad y de mortalidad, para con ello generar mayor evidencia de la relación que existe entre la contaminación y la salud humana [Gaviria G et al., 2011].

- **Evaluación de la concentración de material particulado ($PM_{2,5}$) y su relación con enfermedades respiratorias en el municipio de Sogamoso, Boyacá-Colombia:** esta investigación es resultado de una tesis de la Universidad de Barcelona realizada en el año 2021, la cual presenta el estudio de la relación entre las concentraciones de materia particulada fina y la presencia de enfermedades respiratorias en el municipio de Sogamoso, Boyacá. Para esta investigación se realizó un modelo geoestadístico en donde se concluyó que la concentración diaria de material particulado fino presenta un índice alto, lo que demuestra que estas concentraciones podrían ser dañinas a la salud para grupos sensibles del municipio [Sandoval Meneses, 2021].
- **Morbilidad respiratoria asociada con la exposición a material particulado en el ambiente:** éste artículo presenta los resultados de un análisis en donde se estudió la exposición prolongada a material particulado en el aire y su relación con la morbilidad respiratoria en los habitantes de las zonas urbanas de Medellín. En este sentido, los autores encontraron que efectivamente los altos niveles de este contaminante contribuyen a aumentar el riesgo en la presencia de enfermedades respiratorias [Martínez L et al., 2011].
- **Relación entre la calidad del aire y la incidencia de enfermedades respiratorias en las comunas 4, 6, 7 y 8 del municipio de Cúcuta, Norte de Santander:** éste artículo presenta los resultados del estudio realizado para analizar la relación entre la calidad del aire y la incidencia de enfermedades respiratorias en las comunas 4, 6, 7 y 8 del municipio de Cúcuta, Colombia, en donde se encontró que hay una asociación dependiente entre estas variables de estudio, resaltándose una relación significativa en la presencia de rinofaringitis aguda en pacientes vulnerables y el estado de calidad del aire [Becerra Moreno et al., 2021].

Finalmente a nivel internacional se resaltan las siguientes investigaciones:

- **Carga de enfermedad por contaminación atmosférica debida a materia particulada (PM_{10} y $PM_{2,5}$) y ozono (O_3) en la ciudad de Valladolid durante el periodo 1999-2008:** esta tesis doctoral tuvo como objetivo realizar una estimación de la carga ambiental de enfermedad causada por la contaminación atmosférica por materia particulada y ozono durante los años 1999 a 2008 en la ciudad de Valladolid, España. En dicho estudio se concluye que, se habría podido disminuir considerablemente los fallecimientos por enfermedades respiratorias si se hubieran aplicado los valores de referencia de calidad del aire señalados por la OMS. Expresan que, con respecto a la exposición a corto plazo de PM_{10} se podría haber prevenido una proporción anual del 2,5 % en la mortalidad por enfermedad respiratoria, de igual forma, en cuanto a la exposición de $PM_{2,5}$ se habría prevenido una proporción anual del 25 % en la mortalidad [Cárdaba Arranz, 2012].

- **Impacto de la contaminación del aire por PM_{10} sobre la mortalidad diaria en Temuco, Chile:** en este artículo se presentan los resultados de un estudio realizado el cual buscó estudiar los efectos a corto plazo del PM_{10} sobre la mortalidad diaria en Temuco, en el sur de Chile, por causas respiratorias y cardiovasculares. En consecuencia, como resultados se presenta que una vez estimados modelos de regresión multivariante poisson a las variables mortalidad, meteorología y contaminación del aire, se encontró que hay una asociación significativa entre el incremento de materia particulada gruesa y el aumento en la mortalidad diaria en Temuco por enfermedad respiratoria total, cardiovascular total y cardiorrespiratorias [Sanhueza H et al., 2006].
- **Efectos de la contaminación atmosférica por material particulado en las enfermedades respiratorias agudas en menores de 5 años:** en este artículo se presenta un análisis realizado para determinar la correlación entre los fenómenos críticos de contaminación atmosférica y las consultas por Infecciones Respiratorias Agudas (IRA) en la ciudad de Temuco, Chile. Se encontró que, en los períodos de contaminación atmosférica aumentaban las consultas por IRA [Barrios Casas et al., 2004].

En este contexto, este estudio tiene como objetivo analizar la relación entre la contaminación por material particulado (PM_{10} y $PM_{2,5}$) y la tasa de mortalidad por infecciones respiratorias agudas en el distrito de Bogotá durante los años 2018 a 2022, mediante un enfoque espacio-temporal, proporcionando información relevante para el diseño y la implementación de políticas públicas efectivas que contribuyan a mitigar los efectos negativos de la contaminación del aire en la salud de la población.

Por tanto, este documento presenta en una primera sección los objetivos general y específicos del estudio, en una segunda sección se encuentra el marco teórico con los fundamentos teóricos relacionados con el material particulado, la mortalidad por enfermedades respiratorias y los principales conceptos de los aspectos metodológicos empleados, seguido de una tercera sección con la metodología de la investigación, en donde se describe el área de estudio, las fuentes de información, y el desarrollo metodológico que comprendió seis fases. En la siguiente sección se presentan los resultados obtenidos distribuidos en cuatro capítulos de acuerdo con los objetivos específicos planteados, en el primer capítulo se aborda el análisis exploratorio de los datos, en un segundo capítulo se presenta un análisis de tendencias temporales y patrones de comportamiento de los niveles de material particulado, seguido de un tercer capítulo con la interpolación espacial de la concentración de material particulado en la región de estudio; y un último capítulo con la estimación de modelos espaciales y espacio temporales que examinan la relación entre la contaminación por material particulado y la tasa de mortalidad por IRA. Finalmente, se presenta en una última sección las conclusiones y recomendaciones de la investigación.

1. Objetivos

1.1. Objetivo General

Analizar la relación entre la contaminación por material particulado (PM_{10} y $PM_{2,5}$) y la tasa de mortalidad por enfermedades respiratorias en Bogotá D.C, durante el periodo 2018-2022 considerando modelos espacio temporales.

1.2. Objetivos Específicos

1. Analizar y cuantificar tendencias temporales y patrones de comportamiento de los niveles de material particulado (PM_{10} y $PM_{2,5}$) en Bogotá D.C. durante el periodo 2018-2022.
2. Realizar un análisis de factores de riesgo asociados para identificar y examinar aquellos elementos que puedan influir en las tasas de mortalidad por enfermedades respiratorias en Bogotá D.C. durante el periodo 2018-2022, considerando el rango etario y el régimen afiliación.
3. Determinar la concentración de material particulado (PM_{10} y $PM_{2,5}$) en las localidades de Bogotá D.C para el periodo 2018 a 2022, a partir de los registros en las estaciones de la Red de Monitoreo de Calidad de Aire de Bogotá RMCAB.
4. Examinar patrones espaciales para investigar variaciones geográficas en la concentración de material particulado (PM_{10} y $PM_{2,5}$) en Bogotá D.C. durante el periodo 2018-2022, y cómo estas se correlacionan con la distribución espacial de las tasas de mortalidad por enfermedades respiratorias.

2. Marco Teórico

2.1. Material Particulado (PM)

El material particulado es considerado uno de los contaminantes más importantes y comunes del aire, debido a su gran efecto negativo en la salud humana. Es una “mezcla compleja de materiales sólidos y líquidos suspendidos en el aire, que pueden variar significativamente en tamaño, forma y composición, dependiendo fundamentalmente de su origen” [Alcaldía de Bogotá, 2008], compuesta principalmente por sulfatos, nitratos, amoníaco, cloruro de sodio, hollín, polvos minerales y agua. Así mismo, son generadas por diferentes sectores de la economía distribuidos entre fuentes fijas y móviles. Las primeras hacen referencia a aquellas fuentes que se encuentran en un lugar o sitio determinado e inmóvil, desde donde emiten contaminantes; por su parte las fuentes móviles se definen como aquellas que debido a su uso pueden desplazarse por diferentes lugares.

Dichas partículas son generadas principalmente por la quema de combustibles en diferentes sectores (transporte, energía, industria, agricultura, entre otros), destacándose entre sus efectos más nocivos la relación directa con la presencia de cáncer de pulmón en la población vulnerable, debido principalmente a que al ser una partícula suspendida en el aire, según el tamaño de su diámetro, puede entrar profundamente en los pulmones y en algunos casos ingresar en el torrente sanguíneo, lo que afecta principalmente el sistema cardiovascular y respiratorio del ser humano. Según la Organización Mundial de la Salud - OMS la contaminación del aire es uno de los principales riesgos ambientales para la salud; los efectos combinados de la contaminación del aire exterior e interior se vinculan a 6,7 millones de muertes prematuras anualmente, así mismo, se estima que en 2019 la contaminación del aire exterior causó 4,2 millones de muertes prematuras a nivel mundial debido a la exposición a partículas finas, que provocan enfermedades cardiovasculares, respiratorias y cánceres [Organización Mundial de la Salud, 2022].

En consecuencia, estas partículas según su tamaño se pueden clasificar así:

1. **Materia particulada gruesa (PM_{10}):** partícula con diámetro de 10 micrones o menos, la cual puede ingresar profundamente en los pulmones.
2. **Materia particulada fina ($PM_{2,5}$):** partícula con diámetro de 2,5 micrones o menos,

la cual tiene un mayor efecto en la salud humana, ya que, debido a su pequeño tamaño puede ingresar a los pulmones y al torrente sanguíneo.

Teniendo en cuenta lo anterior, la Organización Mundial de la Salud - OMS establece la necesidad de adoptar políticas que permitan reducir algunas fuentes de contaminación del aire, es así como la OMS, por medio de las Directrices Mundiales sobre la Calidad del Aire ofrece orientaciones a escala mundial sobre los umbrales y límites de los principales contaminantes atmosféricos que conllevan riesgos para la salud [Organización Mundial de la Salud, 2022]. Para el año 2021 la OMS actualizó las directrices con respecto a los valores máximos permisibles mundiales para estas dos partículas (Tabla 2-1). Así mismo, para el territorio nacional (Colombia), el Ministerio de Ambiente y Desarrollo Sostenible adoptó la norma de calidad del aire mediante Resolución No. 2254 de 2017, en donde se establecen los niveles máximos permisibles para contaminantes criterio que rigen en todo el país desde el año 2018 hasta 2030 [Ministerio de Ambiente y Desarrollo Sostenible, 2017] (Tabla 2-2).

Tabla 2-1.: Niveles Máximos Permisibles de Materia Particulada (PM) a nivel mundial

Contaminante	Nivel Máximo Permisible($\mu g/m^3$)	Tiempo de Exposición
PM_{10}	15	Media Anual
	45	Media Diaria
$PM_{2,5}$	5	Media Anual
	15	Media Diaria

Tabla 2-2.: Niveles Máximos Permisibles de Materia Particulada (PM) a nivel nacional

Contaminante	Nivel Máximo Permisible($\mu g/m^3$)	Tiempo de Exposición
PM_{10}	30	Anual
$PM_{2,5}$	15	Anual

2.1.1. Red de Monitoreo de Calidad de Aire de Bogotá – RMCAB

La Secretaría Distrital de Ambiente de Bogotá cuenta con la Red de Monitoreo de Calidad del Aire de Bogotá - RMCAB, la cual opera desde 1997, encargándose de vigilar las concentraciones de contaminantes criterio (material particulado de diámetro menor a 10 y 2.5

micras - PM_{10} y $PM_{2.5}$, Ozono - O_3 , dióxido de azufre - SO_2 y dióxido de nitrógeno - NO_2 y monóxido de carbono - CO), de black carbón (o carbono negro) y variables meteorológicas (precipitación, temperatura, presión atmosférica, radiación solar, velocidad y dirección del viento), por medio de veinte (20) estaciones instaladas en diferentes lugares de la ciudad [Secretaría Distrital de Ambiente de Bogotá, sf].

Dichas estaciones están conformadas por monitores, analizadores y sensores automáticos que recolectan datos cada hora sobre el estado de la calidad del aire en Bogotá. Esta información es almacenada y enviada vía internet hacia el servidor central de la Secretaría de Ambiente. En este se prevalece para posteriormente ser publicada en tiempo real en la página web de la entidad permitiendo la armonización, intercambio y uso de los datos de forma libre a nivel mundial [Secretaría Distrital de Ambiente de Bogotá, sf].

2.2. Enfermedades-Infecciones Respiratorias Agudas (IRA)

Las Infecciones Respiratorias Agudas (IRA), son un grupo de enfermedades producidas específicamente en el aparato respiratorio o vías respiratorias, causadas en su mayoría por microorganismos como virus y bacterias, los cuales ingresan al cuerpo humano por medio del aire o el contacto directo con personas enfermas; entre estas enfermedades se destacan el resfriado común, la influenza, neumonía y asma. Según el Ministerio de Salud y Protección Social esta es la “infección más frecuente en el mundo y representa un importante tema de salud pública en nuestro país; la mayoría de estas infecciones como el resfriado común son leves, pero dependiendo del estado general de la persona pueden complicarse y llegar a amenazar la vida, como en el caso de las neumonías” [Ministerio de Salud y Protección Social de Colombia, sf].

Según datos del Observatorio de Salud de Bogotá, durante el periodo 2018-2022 se presentaron un total de 260 defunciones (históricamente los niños de sexo masculino menores de 5 años son los más afectados) por IRA y neumonía. De igual forma, se precisa que después de la pandemia por COVID-19 se presenta un aumento considerable de las tasas de mortalidad por dichas enfermedades, debido al desarrollo de comorbilidades y a que en este periodo se pospusieron varios tratamientos cardiorrespiratorios [Secretaría Distrital de Salud de Bogotá, sfb].

2.2.1. Observatorio de Salud de Bogotá - SaluData

La Secretaría Distrital de Salud de Bogotá cuenta con el Observatorio de Salud - SaluData como un espacio web de información pública y abierta sobre la situación de salud de Bogotá

D.C. y sus factores determinantes, que permite a los ciudadanos interactuar y acceder de manera fácil y ordenada a información, recursos y servicios, proporcionando evidencia para fomentar comportamientos saludables en la comunidad [Secretaría Distrital de Salud de Bogotá, sfa].

En tal sentido, este observatorio a través del uso de tecnologías de la información y comunicación contribuye a la generación de conocimiento e información para la ciudadanía sobre la situación de salud y sus determinantes. Por medio de su sitio web recopilan, analizan y divulgan información en salud, así mismo establecen canales de comunicación bidireccionales entre la ciudadanía y la administración pública, fomentando espacios de escucha y participación ciudadana en la promoción y fomento de la salud, promueven la apropiación social del conocimiento como un espacio vital para incentivar prácticas saludables, informando sobre los factores de riesgo que afectan la salud de los ciudadanos y comunidades, así como la identificación y formulación de soluciones a los problemas en los servicios de salud [Secretaría Distrital de Salud de Bogotá, sfa].

2.2.2. Clasificación Internacional de Enfermedades (CIE)

En la actualidad se utiliza internacionalmente la Clasificación Internacional de Enfermedades - CIE, para clasificar enfermedades, esta es utilizada en 117 países de todo el mundo en registros clínicos y certificados de defunción [Gómez Rivadeneira, 2015].

Dicha clasificación corresponde a un sistema creado por la Organización Mundial de la Salud (OMS), el cual se utiliza para clasificar o codificar los diagnósticos en las historias clínicas. Este sistema contiene más de 55.000 códigos que se agrupan en 21 capítulos, cada uno cubre una parte diferente del cuerpo o un sistema de órganos [Gómez Rivadeneira, 2015].

2.3. Series de tiempo

Una serie de tiempo es la realización de un proceso estocástico, el cual se define como una “colección de variables aleatorias ordenadas en el tiempo” [Gujarati and Porter, 2010], es decir una observación numérica de una variable en intervalos de tiempo fijos e igualmente espaciados. En este proceso surge el concepto de estacionariedad, atribuido de cierto modo con la palabra estable el cual se busca en el análisis tradicional o clásico de cualquier serie de tiempo, debido a que si el comportamiento de los datos a través del tiempo no es estable, solo se podrá considerar su análisis bajo un periodo específico, por lo tanto, no se podrá generalizar. De ahí la importancia de encontrar o lograr una serie de tiempo estacionaria.

En consecuencia, se dice que una serie de tiempo puede tener un comportamiento estacionario en sentido débil o fuerte; se considera un proceso estocástico como estacionario en sentido fuerte si la función de probabilidad conjunta de cualquier colección de variables aleatorias es invariante en el tiempo, y se considera estacionario en sentido débil si su media y varianza son constantes y su función de autocovarianza o su función de autocorrelación solo dependen del rezago y no del tiempo. De igual forma, si se encuentra con un proceso estocástico normal, es decir que cumpla con el supuesto de normalidad, bastará con tener estacionariedad en sentido débil, ya que un proceso normal está especificado por su media y varianza [Gujarati and Porter, 2010].

Se dice que Y_t es una serie de tiempo estacionaria en sentido débil si tiene las siguientes propiedades:

$$\text{Media : } E(Y_t) = \mu \quad (1)$$

$$\text{Varianza : } \text{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \quad (2)$$

$$\text{Covarianza : } \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \quad (3)$$

2.3.1. Proceso Autorregresivo (AR)

Son aquellos modelos explicados solamente por su información pasada, es decir el valor de Y en el periodo t solo depende de su valor en sus periodos anteriores y un término aleatorio [Gujarati and Porter, 2010].

Si se modela Y_t como

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + u_t \quad (4)$$

donde δ es la media de Y y u_t es un término de error aleatorio no correlacionado con media cero y varianza constante σ^2 (es decir, *ruido blanco*), se dice que Y_t sigue un proceso estocástico **autorregresivo de primer orden**, o **AR(1)**. Aquí el valor de Y en el tiempo t depende de su valor en el periodo anterior y de un término aleatorio; los valores de Y están expresados como desviaciones de su valor medio. En otras palabras, este modelo dice que el valor de pronóstico de Y en el periodo t es simplemente alguna proporción ($= \alpha_1$) de su valor en el periodo $(t - 1)$ más un “choque” o perturbación aleatoria en el tiempo t ; de nuevo, los valores de Y están expresados alrededor del valor de su media [Gujarati and Porter, 2010].

Pero si se considera este modelo,

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + u_t \quad (5)$$

se dice que Y_t sigue un **proceso autorregresivo de segundo orden** o **AR(2)**. Es decir, el valor de Y en el tiempo t depende de sus valores en los dos periodos anteriores, los valores de Y expresados alrededor del valor de su media δ [Gujarati and Porter, 2010].

En general, se tiene

$$(Y_t - \delta) = \alpha_1(Y_{t-1} - \delta) + \alpha_2(Y_{t-2} - \delta) + \dots + \alpha_p(Y_{t-p} - \delta) + u_t \quad (6)$$

en cuyo caso Y_t es un **proceso autorregresivo de orden p** , o **AR(p)**.

Se observa que en todos los modelos anteriores sólo se consideran los valores actuales y anteriores de Y ; no hay otras regresoras. En este sentido, se dice que “los datos hablan por sí mismos” [Gujarati and Porter, 2010].

2.3.2. Proceso de Medias Móviles (MA)

El proceso AR recién expuesto no es el único mecanismo que pudo generar a Y . Si se supone un modelo de Y de la siguiente manera [Gujarati and Porter, 2010]:

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} \quad (7)$$

donde μ es una constante y u , al igual que antes, es el término de error estocástico de ruido blanco. Aquí, Y en el periodo t es igual a una constante más un promedio móvil de los términos de error presente y pasado. Así, en este caso decimos que Y sigue un proceso de **promedios móviles de primer orden**, o **MA(1)** [Gujarati and Porter, 2010].

Pero si Y sigue la expresión

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} \quad (8)$$

es un proceso **MA(2)**. En forma más general,

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q} \quad (9)$$

es un proceso **MA(q)**. En resumen, un proceso de promedios móviles es tan sólo una combinación lineal de términos de error de ruido blanco [Gujarati and Porter, 2010].

2.3.3. Proceso Autorregresivo de Medias Móviles (ARMA)

Desde luego, es muy probable que Y_t tenga características de AR y de MA a la vez, y, por consiguiente, sea ARMA. Así, Y_t sigue un proceso **ARMA(1, 1)** si se escribe como [Gujarati and Porter, 2010]:

$$Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1} \quad (10)$$

porque hay un término autorregresivo y uno de promedios móviles. En donde, θ representa un término constante.

En general, en un proceso **ARMA**(p, q), habrá p términos autorregresivos y q términos de promedios móviles [Gujarati and Porter, 2010].

2.4. Estadística Espacial

La estadística espacial se enfoca en estudiar procesos que ocurren en el espacio y cuya ubicación se conoce. Esta rama de la estadística se centra en el análisis e interpretación de datos que están distribuidos en el espacio geográfico; en lugar de tratar los datos como una simple colección de observaciones independientes, la estadística espacial considera la ubicación geográfica como una variable importante en sí misma [Gómez Rubio, 2022].

Trata con el análisis de realizaciones de procesos estocásticos donde el espacio de parámetros es un subconjunto de R^d . Definiendo un proceso estocástico como una familia de variables aleatorias $Z(x) : x \in D \subset R^d$ definida sobre un espacio de probabilidad (Ω, F, P) . El conjunto D de índices del proceso se denomina espacio de parámetros. Los valores que toma $Z(x)$ se llaman estados y el conjunto de todos los posibles valores de $Z(x)$ se llama espacio de estados. Los procesos estocásticos son clasificados de acuerdo con el espacio de parámetros y de estados (discreto y continuo) [Cressie, 1993].

Dentro de la estadística espacial se destacan tres áreas principales:

1. Geoestadística:

Se centra en el estudio de variables espaciales que cambian de manera continua dentro de la región de estudio. Se cuenta con observaciones de las variables de interés en unas pocas localizaciones del espacio, a partir de las cuales es posible estimar el valor de la variable de interés en cualquier punto de la región de estudio [Gómez Rubio, 2022].

La geoestadística surgió a principios de la década de 1980 como una disciplina híbrida de ingeniería minera, geología, matemáticas y estadística. Su punto fuerte con respecto a los enfoques más clásicos se centran en que la variabilidad espacial a pequeña y gran escala, modela la tendencia y correlación espacial [Cressie, 1993].

Dentro de la geoestadística se destaca el estudio de la contaminación ambiental como uno de sus principales ejemplos, en donde a partir de una serie de mediciones en unos pocos puntos se encuentra que se puede estimar la concentración de algún contaminante. Mediante técnicas de modelización geoestadística es posible obtener una estimación de la concentración de algún contaminante en la atmosfera en una región de estudio. Adicionalmente, en epidemiología el uso de modelos geoestadísticos es relevante puesto que permite modelizar la variación continua en la región de estudio del riesgo de sufrir una determinada enfermedad [Gómez Rubio, 2022].

2. Patrones puntuales:

En el análisis de procesos puntuales se estudia la distribución de la ubicación de una serie de eventos o un patrón puntual dentro de la región de estudio. Surgen cuando la variable a analizar es la localización de los acontecimientos. En la mayoría de los casos, la primera pregunta que hay que responder es si el patrón muestra una aleatoriedad espacial completa, agrupación o regularidad [Cressie, 1993].

Las ubicaciones pertenecen a un conjunto D que puede ser discreto o continuo y su selección no depende del investigador (D aleatorio) [Giraldo Henao, 2002].

3. Datos de área:

El análisis de datos agregados o de área se refiere al estudio de variables que no están necesariamente relacionadas con una ubicación en el espacio, sino que están asociadas a determinadas divisiones del mismo. Su análisis se basa en distintos modelos de regresión en los que, se modeliza la correlación en la variable respuesta entre regiones vecinas [Gómez Rubio, 2022].

Una de las áreas en las que más se han utilizado los modelos espaciales de datos agregados o datos de área es la econometría. En tal sentido, el desarrollo de modelos de econometría espacial (Anselin, 1988 y LeSage y Pace, 2009) ha ido en paralelo al desarrollo de modelos espaciales dentro del mundo de la estadística, principalmente debido a la necesidad de poder modelizar la dependencia espacial con regiones vecinas. Lo que permite entender que regiones vecinas tengan comportamientos similares y a su vez,

generar diferentes entre cada una de ellas [Gómez Rubio, 2022].

A continuación, se presentan algunos conceptos importantes a tener en cuenta en estadística espacial:

Variable Regionalizada

Es aquella variable medida en el espacio de forma que presente una estructura de correlación. Se puede definir como un proceso estocástico con dominio contenido en un espacio euclidiano d -dimensional R^d , $Z(x) : x \in D \subset R^d$, en donde $Z(x)$ puede verse como una medición de una variable aleatoria en un punto x de una región de estudio [Giraldo Henao, 2002].

Momentos de una Variable Regionalizada

Sea $Z(x) : x \in D \subset R^d$ el proceso estocástico que define la variable regionalizada.

Para cualquier n puntos $x_1, x_2, \dots, x_n$, el vector aleatorio $\vec{Z}(x) = [Z(x_1), Z(x_2), \dots, Z(x_n)]^T$ está definido por su función de distribución conjunta:

$$F[z_1, z_2, \dots, z_n] = P[Z(x_1) \leq z_1, Z(x_2) \leq z_2, \dots, Z(x_n) \leq z_n] \quad (11)$$

Conocidas las densidades marginales univariadas y bivariadas se pueden establecer los siguientes valores esperados (momentos univariados y bivariados) [Giraldo Henao, 2002]:

$$E(Z(x_i)) = m(x_i) \quad (12)$$

$$V(Z(x_i)) = E[Z(x_i) - m(x_i)]^2 = \sigma_i^2 \quad (13)$$

$$\begin{aligned} C(Z(x_i), Z(x_j)) &= E[Z(x_i) - m(x_i)][Z(x_j) - m(x_j)] : \text{Función de autocovarianza} \\ \gamma(Z(x_i), Z(x_j)) &= \frac{1}{2} E[Z(x_i) - Z(x_j)]^2 : \text{Función de semivarianza} \end{aligned}$$

Estacionariedad

La variable regionalizada es estacionaria si su función de distribución conjunta es invariante respecto a cualquier translación del vector h , o lo que es lo mismo, la función de distribución del vector aleatorio [Giraldo Henao, 2002].

$$\begin{aligned} \vec{Z}(x) &= [Z(x_1), Z(x_2), \dots, Z(x_n)]^T \text{ es idéntica a la del vector} \\ \vec{Z}(x) &= [Z(x_1 + h), Z(x_2 + h), \dots, Z(x_n + h)]^T \text{ para cualquier } h \end{aligned}$$

1. Estacionariedad de Segundo Orden:

Se dice que $Z(x)$ es estacionario de segundo orden si cumple:

$$a) E[Z(x)] = m, k \in R, \forall x \in D \subset R^d$$

El valor esperado de la variable aleatoria es finito y constante para todo punto en el dominio.

$$b) COV[Z(x), Z(x+h)] = C(h) < \infty$$

Para toda pareja $Z(x), Z(x+h)$ la covarianza existe y es función única del vector de separación h .

2. Estacionariedad Débil o Intrínseca:

Se dice que $Z(x)$ es estacionario débil si cumple:

a) $Z(x)$ tiene esperanza finita y constante para todo punto en el dominio. Lo que implica que la esperanza de los incrementos es cero.

$$E[Z(x+h) - Z(x)] = 0 \quad (14)$$

b) Para cualquier vector h , la varianza del incremento está definida y es una función única de la distancia.

$$V[Z(x+h) - Z(x)] = E[Z(x+h) - Z(x)]^2 = 2\gamma(h) \quad (15)$$

Si la variable regionalizada es estacionaria fuerte entonces también será estacionaria débil. En el campo espacial existen múltiples direcciones y por lo tanto se debe asumir que en todas el fenómeno es estacionario. Cuando la esperanza de la variable no es la misma en todas las direcciones o cuando la covarianza o correlación dependan del sentido en que se determinan, no habrá estacionariedad. Si la correlación entre los datos no depende de la dirección en la que esta se calcule se dice que el fenómeno es isotrópico, en caso contrario se hablará de anisotropía [Giraldo Henao, 2002].

Funciones de Correlación Espacial

La primera fase para el análisis geoestadístico es la determinación de la dependencia espacial entre los datos medidos de una variable. Para llevarla a cabo, se usan tres funciones: El semivariograma, el covariograma y el correlograma [Giraldo Henao, 2002].

1. Variograma y Semivariograma:

El variograma permite ver cómo es la dependencia espacial de la variable de interés. En un escenario de estacionariedad débil se asume que la varianza de los incrementos

de la variable regionalizada es finita. A esta función denotada por $2\gamma(h)$ se le denomina variograma. Utilizando la definición teórica de la varianza en términos del valor esperado de una variable aleatoria, se tiene [Giraldo Henao, 2002]:

$$2\gamma(h) = V(Z(x+h) - Z(x)) \quad (16)$$

$$\begin{aligned} &= E((Z(x+h) - Z(x))^2) - \underbrace{(E(Z(x+h) - Z(x)))^2}_0 \\ &= E((Z(x+h) - Z(x))^2) \end{aligned}$$

La mitad del variograma $\gamma(h)$, se conoce como la función de semivarianza y caracteriza las propiedades de dependencia espacial del proceso. Dada una realización del fenómeno, la función de semivarianza es estimada, por el método de momentos, a través del semivariograma experimental, el cual para su interpretación parte del criterio de que a menor distancia entre los sitios mayor similitud o correlación espacial entre las observaciones. Por ello en presencia de autocorrelación se espera que para valores de h pequeños el semivariograma experimental tenga magnitudes menores a las que este toma cuando las distancias h se incrementan [Giraldo Henao, 2002].

2. Covariograma y Correlograma:

La función de covarianza muestral entre parejas de observaciones que se encuentran a una distancia h se calcula, empleando la formula clásica de la covarianza muestral, por:

$$\begin{aligned} C(h) &= COV(Z(x+h), Z(x)) = \frac{\sum_{i=1}^n (Z(x+h) - m)(Z(x) - m)}{n} \\ &= \frac{\sum_{i=1}^n (Z(x+h) \cdot Z(x))}{n} - m^2 = C(h) \end{aligned} \quad (17)$$

Donde:

m representa el valor promedio en todo punto de la región de estudio

n representa el número de parejas de puntos que se encuentran a una distancia h

Asumiendo que el fenómeno es estacionario y estimando la varianza de la variable regionalizada a través de la varianza muestral, se tiene que el correlograma muestral está dado por [Giraldo Henao, 2002]:

$$r(h) = \frac{COV(Z(x+h), Z(x))}{S_{x+h} \cdot S_x} = \frac{C(h)}{S_x^2} = \frac{C(h)}{C(0)} \quad (18)$$

Bajo el supuesto de estacionariedad, las tres funciones de dependencia espacial (semivariograma, covariograma o correlograma), pueden ser usadas en la determinación de la relación espacial entre los datos. Sin embargo, la única que no requiere hacer estimación de parámetros es la función de semivarianza. Por esta razón, en la práctica se emplea el semivariograma y no las otras dos funciones [Giraldo Henao, 2002].

Modelos Teóricos de Semivarianza

Teniendo en cuenta que el semivariograma experimental se calcula sólo para algunas distancias promedios particulares, se hace necesario el ajuste de modelos que generalicen lo observado en el semivariograma experimental a cualquier distancia [Giraldo Henao, 2002].

Existen diversos modelos teóricos de semivarianza que pueden ajustarse al semivariograma experimental; dichos modelos pueden dividirse en no acotados (lineal, logarítmico, potencial) y acotados (esférico, exponencial, gaussiano) [Warrick et al., 1986] citado en [Giraldo Henao, 2002].

Todos estos modelos tienen tres parámetros comunes:

1. **Efecto Pepita:**

Se denota por C_0 y representa una discontinuidad puntual del semivariograma en el origen. Puede ser debido a errores de medición en la variable o a la escala de la misma. En algunas ocasiones puede ser indicativo de que parte de la estructura espacial se concentra a distancias inferiores a las observadas [Giraldo Henao, 2002].

2. **Meseta:**

Es la cota superior del semivariograma. También puede definirse como el límite del semivariograma cuando la distancia h tiende a infinito. La meseta puede ser o no finita. Los semivariogramas que tienen meseta finita cumplen con la hipótesis de estacionariedad fuerte; mientras que cuando ocurre lo contrario, el semivariograma define un fenómeno natural que cumple sólo con la hipótesis intrínseca. La meseta se denota por C_1 o por $(C_0 + C_1)$ cuando la pepita es diferente de cero [Giraldo Henao, 2002].

3. **Rango:**

Corresponde a la distancia a partir de la cual dos observaciones son independientes. El rango se interpreta como la zona de influencia. Entre más pequeño sea el rango, más cerca se está del modelo de independencia espacial [Giraldo Henao, 2002].

2.5. Interpolación Espacial - Predicción Espacial Óptima

La interpolación espacial es una técnica que permite estimar valores desconocidos de una variable espacial a partir de valores conocidos. Específicamente en la temática de contami-

nación ambiental, existen diferentes estaciones de monitoreo, las cuales registran los niveles de contaminación en los puntos del espacio donde se encuentran ubicadas. Por tanto, por medio de la interpolación espacial, se logran estimar los niveles de contaminación en puntos del espacio donde no se dispone de datos de medición o en toda el área de estudio, a partir de datos en las estaciones ubicadas [Gómez Rubio, 2022].

Por medio de esta, se logran generar mapas de contaminación que permiten visualizar la distribución de la contaminación en un área determinada [Gómez Rubio, 2022]. De la teoría de la decisión se conoce que si Z_0 es una cantidad aleatoria y Z_0^* es su predictor, entonces $L(Z_0; Z_0^*)$ representa la pérdida en que se incurre cuando se predice Z_0 con Z_0^* y el mejor predictor será el que minimice $E[L(Z_0; Z_0^*)/Z]$ con $Z = [Z_1, Z_2, \dots, Z_n]'$ es decir el predictor óptimo es el que minimice la esperanza condicional de la función de pérdida [Giraldo Henao, 2002].

$$SiL(Z_0; Z_0^*) = [Z_0 - Z_0^*]^2 \Rightarrow Z_0^* = E(Z_0/Z) \quad (19)$$

La expresión anterior indica que para encontrar el predictor óptimo se requiere conocer la distribución conjunta de la $n + 1$ variables aleatorias.

Método Kriging

Este término encierra un conjunto de métodos de predicción espacial que se fundamentan en la minimización del error cuadrático medio de predicción. Esta palabra procede del nombre del geólogo sudafricano D. G. Krige, cuyos trabajos en la predicción de reservas de oro, realizados en la década del cincuenta, suelen considerarse como pioneros en los métodos de interpolación espacial [Giraldo Henao, 2002].

Kriging Ordinario:

Suponiendo que se hacen mediciones de la variable de interés Z en los puntos $x_i, i = 1, 2, \dots, n$, de la región de estudio, es decir se tienen realizaciones de las variables $Z(x_1), \dots, Z(x_n)$ y se desea predecir $Z(x_0)$, en el punto x_0 donde no hubo medición. En esta circunstancia, el método kriging ordinario propone que el valor de la variable puede predecirse como una combinación lineal de las n variables aleatorias así [Giraldo Henao, 2002]:

$$\begin{aligned} Z^*(x_0) &= \lambda_1 Z(X_1) + \lambda_2 Z(X_2) + \lambda_3 Z(X_3) + \lambda_4 Z(X_4) + \lambda_5 Z(X_5) + \dots + \lambda_n Z(X_n) \quad (20) \\ &= \sum_{i=1}^n \lambda_i Z(X_i) \end{aligned}$$

En donde los λ_i representan los pesos o ponderaciones de los valores originales. Dichos pesos se calculan en función de la distancia entre los puntos muestreados y el punto donde se va a

hacer la correspondiente predicción [Giraldo Henao, 2002].

Método Inverso de la Distancia (IDW)

Por medio de este método los puntos de muestra se ponderan durante la interpolación de tal forma que la influencia de un punto en relación con otro disminuye con la distancia. “La ponderación es asignada a los puntos de muestreo mediante la utilización de un coeficiente de ponderación que controla cómo la influencia de la ponderación decae mientras la distancia hacia el punto nuevo se incrementa” [QGIS Documentation, sf]. Por tanto el coeficiente de ponderación incrementa y el valor de los puntos desconocidos se aproxima al valor del punto de observación más cercano [QGIS Documentation, sf].

Para el método del inverso de la distancia la estimación en el punto X_0 se obtiene como:

$$Z^*(X_0) = \frac{\sum_{i=1}^n \lambda_i Z(X_i)}{\sum_{i=1}^n \lambda_i} \quad (21)$$

Donde los pesos λ_i son inversamente proporcionales a la distancia $d_{i0} = ||X_i - X_0||$ entre X_i y X_0 , es decir:

$$\lambda_i = \left(\frac{1}{d_{i0}} \right)^p \quad (22)$$

Método Modelos Lineales

Utilizar modelos lineales para interpolar en el ámbito de la geoestadística puede parecer contraproducente, dado que estos modelos requieren independencia entre las observaciones, mientras que la geoestadística se beneficia de la dependencia espacial. No obstante, hay situaciones en las que al filtrar la tendencia de un conjunto de datos georeferenciados, los residuos resultantes pueden no mostrar dependencia espacial, con un semivariograma constante. En estos casos, la supuesta dependencia espacial en los datos originales se debe realmente a la tendencia presente [Verdoy, 2005], [Melo Martínez, 2012].

Una alternativa para generar una estimación $Z^*(x_0)$ involucra la utilización de modelos lineales, para esto, la variable de interés $Z(\cdot)$ se explica en función de las dos direcciones $X = longitud$, $Y = latitud$, en donde usualmente se ajusta un modelo de primer orden:

$$Z = \beta_0 + \beta_1 X + \beta_2 Y + \varepsilon \quad (23)$$

o un modelo de segundo orden:

$$Z = \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 X^2 + \beta_4 Y^2 + \beta_5 XY + \varepsilon \quad (24)$$

Es importante señalar que al utilizar modelos lineales se requiere información auxiliar disponible en cada punto del espacio x_i , por esta razón se utilizan como variables explicativas del modelo la longitud y la latitud.

2.6. Econometría Espacial

Considerada como una subdisciplina de la econometría la cual se ocupa del análisis de datos que tienen una componente espacial o geográfica, combina métodos estadísticos y económicos para estudiar, modelar y analizar las interacciones espaciales. El desarrollo de modelos de econometría espacial [Anselin, 1998] y [LeSage and Pace, 2009] ha ido en paralelo al desarrollo de modelos espaciales dentro del mundo de la estadística, principalmente por la necesidad de modelizar explícitamente dentro del modelo la dependencia espacial con regiones vecinas [Gómez Rubio, 2022].

Los métodos desarrollados por la econometría espacial permiten atender problemas de violación a los supuestos del modelo de regresión, que no es posible resolver en el marco de los modelos econométricos estándar. Estos problemas son típicos en los datos espaciales y se refieren a [Quintana Romero and Mendoza, 2016]:

1. **Correlación espacial:** Dependencia espacial entre observaciones
2. **Heteroscedasticidad espacial:** Heterogeneidad espacial entre observaciones

Los modelos de econometría espacial pueden incluir un término autorregresivo sobre el término del error o la variable respuesta. En el primer caso, el modelo se escribiría así:

$$Y = X\beta + u; u = \lambda Wu + e \quad (25)$$

En donde:

Y representa la variable respuesta

X es una matriz de covariables que incluye las variables independientes que se utilizan para explicar la variable respuesta Y

β son los coeficientes asociados a las covariables en la matriz X

u es el efecto autocorrelado, que representa el término de error en el modelo

λ es el parámetro que mide la intensidad de la autocorrelación espacial en el término de error u

W es una matriz de pesos espaciales que captura la estructura de dependencia espacial entre las observaciones

e es un efecto aleatorio de media cero y varianza igual a $\sigma_e^2 I$

Este modelo se puede reescribir tal que u es un error normal mutivariante de media cero y varianza $\sigma_e^2 I((I - \lambda W^T)(I - \lambda W))^{-1}$.

Cuando el término autorregresivo es en la variable respuesta, el modelo es:

$$Y = pWY + X\beta + e \quad (26)$$

En donde p representa un parámetro que mide la intensidad de la dependencia espacial en el modelo. De esta manera, el modelo se puede reescribir así:

$$Y = (I - pW)^{-1} X\beta + u \quad (27)$$

En donde I es la matriz identidad y u es un error autorregresivo como el definido arriba.

Este modelo presenta la característica de que el término de las covariables va multiplicado por $(I - pW)^{-1}$, de manera que el efecto de las covariables no solo afectaría al área en la que se han medido, sino también a los vecinos [Gómez Rubio, 2022].

Los modelos clásicos de econometría espacial permiten estimar el efecto que tendría el cambio en una covariable en la variable respuesta en las regiones vecinas, “una aplicación muy interesante del análisis espacial de datos agregados es el desarrollo de atlas de mortalidad dentro de la epidemiología espacial [Martinez Beneito and Botella Rocamora, 2019]. A partir de la información de población y mortalidad, es posible estimar el riesgo relativo en cada división administrativa [o espacial] utilizando modelos espaciales” [Gómez Rubio, 2022].

2.7. Modelos Espaciales y Espacio-Temporales

Son herramientas estadísticas y matemáticas utilizadas para analizar las variaciones en el espacio y en el tiempo; estos modelos buscan comprender las relaciones que existen entre ciertas variables y su ubicación geográfica, describiendo cómo cambian los fenómenos en el espacio y cómo están relacionados con otros factores. En particular, estos modelos consideran la autocorrelación espacial y temporal.

Para lograrlo, se describen los atributos específicos de cada objeto, basados en sus propiedades espaciales como tipo, longitud, forma, pendiente y orientación, así como en su ubicación geográfica en un sistema de coordenadas y sus relaciones espaciales como conectividad, contigüidad y proximidad [Xin, 2010]. Además, se centran en analizar los eventos que ocurren en un objeto situado en un espacio específico, cuya forma, ubicación o características cambian

a lo largo del tiempo [Díaz, 2007].

A partir de esta especificación se pueden mencionar algunos modelos [Quintana Romero and Mendoza, 2016]:

1. Modelo de Regresión Espacial

$$y_i = \sum_{q=1}^Q x_{iq}\beta_q + \varepsilon_i \quad (28)$$

$$i = 1, \dots, n$$

Donde:

y_i es la observación de la variable dependiente

X_i es la observación de una variable explicativa

β_q es el coeficiente de regresión

ε_i es el error aleatorio

2. SDM (Modelo Durbin Espacial) y Modelo SAR (Spatial Autoregressive Model)

El modelo SDM se muestra en (29) junto con su correspondiente *proceso generador de datos* en (30) [LeSage and Pace, 2009],

$$y = \rho W y + \alpha \iota_n + X\beta + W X\theta + \varepsilon \quad (29)$$

$$y = (I_n - \rho W)^{-1}(\alpha \iota_n + X\beta + W X\theta + \varepsilon) \quad (30)$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

donde 0 representa un vector de ceros de $n \times 1$ y ι_n representa un vector de unos de $n \times 1$ asociado con el término constante α . Este modelo puede escribirse como un modelo SAR definiendo: $Z = [\iota_n \ X \ WX]$ y $\delta = [\alpha \ \beta \ \theta]'$, lo que conduce a (31). Esto significa que la función de verosimilitud para los modelos SAR y SDM puede escribirse de la misma forma donde: $Z = [\iota_n \ X]$ para el modelo SAR y $Z = [\iota_n \ X \ WX]$ para el modelo SDM.

$$y = \rho W y + Z\delta + \varepsilon \quad (31)$$

$$y = (I_n - \rho W)^{-1} Z\delta + (I_n - \rho W)^{-1} \varepsilon \quad (32)$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

A partir de la expresión del modelo en (31), si el valor verdadero del parámetro ρ fuera conocido, digamos ρ^* , podríamos reorganizar la ecuación del modelo en (31) como se muestra en (33):

$$y - \rho^* W y = Z\delta + \varepsilon \quad (33)$$

Esto sugiere una estimación para δ de $\hat{\delta} = (Z'Z)^{-1}Z'(I_n - \rho^* W)y$. En este caso, también podríamos encontrar una estimación para el parámetro de varianza del ruido $\hat{\sigma}^2 = n^{-1}e(\rho^*)'e(\rho^*)$, donde $e(\rho^*) = y - \rho^* W y - Z\hat{\delta}$.

Maximizar la función de log-verosimilitud completa para el caso del modelo SAR implica establecer las primeras derivadas respecto a los parámetros β , σ^2 y ρ igual a cero y resolver simultáneamente estas condiciones de primer orden para todos los parámetros.

$$\ln L = -(n/2) \ln(\pi\sigma^2) + \ln |I_n - \rho W| - \frac{e'e}{2\sigma^2} \quad (34)$$

$$e = y - \rho W y - Z\delta \quad (35)$$

$$\rho \in (\min(\omega)^{-1}, \max(\omega)^{-1})$$

donde ω es el vector propio de W de dimensión $n \times 1$. Si ω contiene solo valores propios reales, se asegura una matriz de varianza-covarianza definida positiva bajo la condición: $\rho \in (\min(\omega)^{-1}, \max(\omega)^{-1})$, como se muestra en Ord (1975). La matriz W puede construirse para tener un valor propio máximo de 1, escalando la matriz de pesos por su valor propio máximo, como se menciona en Barry y Pace (1999); Kelejian y Prucha (2007). De esta manera, el intervalo para ρ se convierte en $(\min(\omega)^{-1}, 1)$, y en la práctica se suele emplear $\rho \in [0, 1]$.

3. Modelo SEM (Spatial Error Model):

Estos modelos abordan la dependencia espacial en el error. Dicha situación es causada por la omisión de variables cruciales que se hallen espacialmente relacionados, errores de medición o unos límites de área que no reflejan con precisión la vecindad [Moreno Serrano and Vayá Valcarce, 2000].

En la forma matricial el modelo se puede escribir como:

$$\begin{aligned} y &= X\beta + \epsilon^* \\ \epsilon^* &= \rho W\epsilon + u^* \end{aligned} \tag{36}$$

donde $\mu^* \sim N(0, \sigma^2 I)$.

Generalidades. El modelo evidencia la dependencia espacial en el término de error. Este puede expresarse como:

$$y = X\beta + (I - \lambda W)^{-1} \mu^* \tag{37}$$

Teniendo en cuenta que el término de error se puede factorizar de la siguiente forma:

$$\begin{aligned} \epsilon^* - \lambda W\epsilon^* &= u^* \\ (I - \lambda W)\epsilon^* &= u^* \\ \epsilon^* &= (I - \lambda W)^{-1} u^*, \text{ con } |\lambda| < 1 \end{aligned} \tag{38}$$

4. Modelos espacio - temporales STFDF

Un modelo espacio-temporal *STFDF* es una extensión de los modelos lineales generalizados (GLM) o modelos lineales mixtos (GLMM) que permite la inclusión de estructuras espaciales y temporales en el análisis estadístico. Estos modelos tienen en cuenta la autocorrelación espacial y temporal presente en los datos y permiten estimar cómo esta autocorrelación afecta a la variable de interés. Un STFDF organiza las observaciones espacio temporales en un marco que combina puntos espaciales (ubicaciones geográficas) y tiempo (instantes temporales) [Pebesma, 2012], [Blangiardo and Cameletti, 2015]. Los datos se pueden escribir en forma de una estructura de tres dimensiones:

$$\{Z(s_i, t_j) \mid s_i \in S, t_j \in T\} \tag{39}$$

Donde:

$Z(s_i, t_j)$ es la observación de la variable en el lugar.

S Es el conjunto de ubicaciones espaciales.

T es el conjunto de instantes de tiempo.

El modelo puede describirse usando funciones de covarianza espacio temporal que cuantifican la relación entre puntos en el espacio y el tiempo. La covarianza entre dos puntos espacio temporales (s_i, t_j) Y (s_k, t_l) viene dada por:

$$COV(Z(s_i, t_j), Z(s_k, t_l)) = C(s_i - s_k, t_j - t_l) \quad (40)$$

Donde:

$C(s_i - s_k, t_j - t_l)$ es una función de covarianza espacio-temporal que depende de la distancia espacial $s_i - s_k$ y $t_j - t_l$.

Este término de covarianza indica cómo están correlacionadas las observaciones en diferentes ubicaciones y tiempos. Para modelar la covarianza, se pueden utilizar funciones específicas como:

$$C(h, u) = \sigma^2 \rho(h) \gamma(u) \quad (41)$$

Donde:

$h = \| s_i - s_k \|$ es la distancia espacial.

$u = | t_j - t_l |$ es la diferencia temporal.

σ^2 es la varianza.

$\rho(h)$ es una función de correlación espacial.

$\gamma(u)$ es una función de correlación temporal.

Componente espacial: Considera la estructura espacial de los datos, es decir, cómo la ubicación geográfica influye en la variable de interés. Puede incluir efectos de vecindad espacial y la interpolación de datos entre ubicaciones conocidas.

$$\rho(h) = \exp\left(-\frac{h}{\phi}\right) \quad (42)$$

Donde:

h es la distancia espacial entre $s_i - s_k$.

ϕ es un parámetro de escala espacial que controla cómo decae la correlación espacial con la distancia.

Componente temporal: Toma en cuenta la estructura temporal de los datos, es decir, cómo la variable de interés cambia a lo largo del tiempo. Puede modelar tendencias temporales, estacionalidad y autocorrelación temporal.

$$\gamma(u) = \exp\left(-\frac{u}{\theta}\right) \quad (43)$$

Donde:

u es la distancia temporal entre $t_j - t_j$.

θ es un parámetro que mide la velocidad de decaimiento de la correlación temporal.

Interacción espacial-temporal: Permite modelar cómo la relación entre las variables de interés varía tanto en el espacio como en el tiempo. Proporciona estimaciones de la incertidumbre asociada con los parámetros del modelo y las predicciones espaciales y temporales.

$$C(h, u) = \sigma^2 \exp\left(-\frac{h}{\phi}\right) \left(\exp -\frac{u}{\theta}\right) \quad (44)$$

3. Metodología

3.1. Descripción del área de estudio

Bogotá es la capital de Colombia, ubicada a una altitud de aproximadamente 2.640 metros sobre el nivel del mar. La ciudad cuenta con 7.743.955 habitantes (según registros del 2020) y con una participación del 25,5 % del PIB del país; es uno de los principales centros de comercio, servicios financieros y empresariales de América Latina.

[Centro Iberoamericano de Desarrollo Estratégico Urbano CIDEU, sf].

Está dividida administrativa y geográficamente en veinte (20) unidades territoriales conocidas como localidades, cada una con su propio alcalde local y Junta Administradora Local (JAL). Estas localidades son: Usaquén, Chapinero, Santa Fé, San Cristóbal, Usme, Tunjuelito, Bosa, Kennedy, Fontibón, Engativá, Suba, Barrios Unidos, Teusaquillo, Los Mártires, Antonio Nariño, Puente Aranda, La Candelaria, Rafael Uribe Uribe, Ciudad Bolívar y Sumapaz.

Para este estudio, se tuvieron en cuenta las localidades de Bogotá a excepción de la localidad 20 conocida como Sumapaz debido a sus condiciones atípicas; es considerada como la localidad rural de Bogotá “se sitúa en el extremo sur de Bogotá... geográficamente dividida por las cuencas del Río Blanco y del Río Sumapaz; se destaca por su belleza natural... reconocida por su productividad agrícola y lechera” [Alcaldía de Bogotá, sf].

3.2. Fuentes de Información

3.2.1. Concentración de Material Particulado (PM)

Para el desarrollo de esta investigación se tomó como fuente de información los registros disponibles en la Red de Monitoreo de Calidad del Aire de Bogotá - RMCAB, la cual opera desde 1997 con datos de acceso libre, encargándose de vigilar las concentraciones de contaminantes criterio (material particulado de diámetro menor a 10 y 2.5 micras - PM_{10} y $PM_{2,5}$, Ozono - O_3 , dióxido de azufre - SO_2 y dióxido de nitrógeno - NO_2 y monóxido de carbono - CO), de black carbón (o carbono negro), y variables meteorológicas (precipitación, temperatura, presión atmosférica, radiación solar, velocidad y dirección del viento), por medio de



Figura 3-1.: Localidades de Bogotá D.C

veinte (20) estaciones instaladas en diferentes lugares de la ciudad.

En consecuencia, se descargaron del sitio web de la RMCAB los registros horarios en microgramo por metro cúbico de aire ($\mu g/m^3$) de las concentraciones de PM_{10} y $PM_{2,5}$ del año 2018 a 2022 en las 20 estaciones instaladas en diferentes lugares de la ciudad. En la Tabla **3-1** se observa la ubicación de estas 20 estaciones distribuidas en 12 localidades de Bogotá.

Como se mencionó anteriormente, la ciudad de Bogotá está distribuida en 20 localidades mientras que la RMCAB cuenta con 20 estaciones distribuidas en 12 localidades, por tanto no se dispone de data de 8 localidades. Para abordar esta situación en el desarrollo de este estudio se realizaron diferentes métodos de interpolación espacial con el fin de estimar el valor promedio del contaminante dentro del distrito.

Definición y clasificación de las variables

De esta forma, se cuenta con una base de datos con los registros horarios de material particulado en el período de estudio comprendido entre el año 2018 hasta el 2022, con un total de

Tabla 3-1.: Estaciones de la RMCAB por localidades

#	Nombre estación	Latitud	Longitud	Localidad
1	Guaymaral	4°47'01.5" N	74°02'38.9" W	Suba
2	Usaquén	4°42'37.26" N	74°1'49.50" W	Usaquén
3	Suba	4°45'40.49" N	74° 5'36.46" W	Suba
4	Bolivia	4°44'08.9" N	74°07'33.2" W	Engativá
5	Las Ferias	4°41'26.52" N	74°4'56.94" W	Engativá
6	Centro de alto rendimiento	4°39'30.48" N	74°5'2.28" W	Engativá
7	Estación móvil 7ma	4°38'42.7" N	74°03'41.6" W	Chapinero
8	MinAmbiente	4°37'31.75" N	74°4'1.13" W	Santa Fe
9	Fontibón	4°40'41.67" N	74°8'37.75" W	Fontibón
10	Puente Aranda	4°37'54.36" N	74°7'2.94" W	Puente Aranda
11	Kennedy	4°37'30.18" N	74°9'40.80" W	Kennedy
12	Carvajal-Sevillana	4°35'45.0" N	74°08'54.6" W	Kennedy
13	Tunal	4°34'34.41" N	74°7'51.44" W	Tunjuelito
14	San Cristóbal	4°34'21.19" N	74°5'1.73" W	San Cristóbal
15	El Jazmín	4°36'30.6" N	74°06'53.8" W	Puente Aranda
16	Usme	4°31'55.4" N	74°07'01.7" W	Usme
17	Bosa	4°36'20.2" N	74°12'14.6" W	Bosa
18	Ciudad Bolívar	4°34'40.1" N	74°09'58.6" W	Ciudad Bolívar
19	Colina	4°44'13.9" N	74°04'10.1" W	Suba
20	Móvil Fontibón	4°40'04.8" N	74°08'54.6" W	Fontibón

832.605 filas y 7 columnas (Fecha, PM_{10} , $PM_{2.5}$, Estación, Latitud, Longitud y Localidad).

A continuación, se definen las variables a emplear:

1. **Fecha:** Periodo de tiempo de los datos registrados por la RMCAB de la Secretaria de Ambiente. Se expresa en registros horarios (posteriormente se estandarizó a una periodicidad mensual).
2. PM_{10} : Contaminante atmosférico que consiste en partículas sólidas o líquidas de 10 micrones o menos de diámetro. Su concentración en la atmosfera se expresa en microgramo por metro cúbico de aire ($\mu g/m^3$).
3. $PM_{2.5}$: Contaminante atmosférico que consiste en partículas sólidas o líquidas de 2.5 micrones o menos de diámetro. Su concentración en la atmosfera se expresa en microgramo por metro cúbico de aire ($\mu g/m^3$).

4. **Estación:** Nombre de las estaciones instaladas por la Red de Monitoreo de Calidad del Aire de Bogotá - RMCAB.
5. **Latitud:** Ubicación geográfica de las diferentes estaciones. Teóricamente se refiere a la distancia angular entre la línea ecuatorial (el ecuador) y un punto determinado de la Tierra.
6. **Longitud:** Ubicación geográfica de las diferentes estaciones. Teóricamente se refiere a la distancia en grados, minutos y segundos que hay con respecto al meridiano principal, que es el meridiano de Greenwich (0°).
7. **Localidad:** Áreas geográficas de Bogotá en donde están ubicadas las estaciones de la Red de Monitoreo de Calidad del Aire de Bogotá - RMCAB.

3.2.2. Mortalidad por Infecciones Respiratorias Agudas (IRA)

Para el desarrollo de esta investigación se tomó como fuente de información los datos suministrados por el Observatorio de Salud de Bogotá (SaluData). Por medio de su sitio web se accedió a la base de datos con el registro mensual de los casos de mortalidad por localidades en Bogotá D.C durante el periodo de 2018 a 2022, con el fin de analizar la relación entre la concentración por materia particulada (PM_{10} y $PM_{2,5}$) y su incidencia en la mortalidad por enfermedades respiratorias en Bogotá D.C, durante el periodo de estudio.

Es de precisar que, esta base se encuentra clasificada por causas de mortalidad según los códigos CIE, correspondientes a un sistema de clasificación internacional de enfermedades, creado por la Organización Mundial de la Salud (OMS), el cual se utiliza para registrar, analizar y comparar datos sobre la salud de la población. Los códigos CIE contienen más de 55.000 códigos que se agrupan en 21 capítulos, cada capítulo cubre una parte diferente del cuerpo o un sistema de órganos.

En tal sentido, y teniendo en cuenta el objetivo planteado para este estudio, se seleccionaron 161 códigos CIE relacionados con causas de mortalidad por enfermedades o infecciones respiratorias agudas (IRA). En el Anexo E se detallan los códigos CIE relacionados con IRA, que fueron tomados en cuenta para este estudio.

Definición y clasificación de las variables

De esta forma, se cuenta con una base de datos con 578.222 filas y 5 columnas (Fecha, Localidad, Rango etario, Régimen de seguridad social y Cantidad de Fallecidos).

A continuación, se definen las variables a emplear:

1. **Fecha:** Periodo de tiempo de los datos registrados por el Observatorio de Salud de Bogotá. Se expresa en registros mensuales.
2. **Localidad:** Áreas geográficas de Bogotá en donde se registran los fallecidos.
3. **Rango etario:** Intervalos de edades utilizados para clasificar a las personas en diferentes grupos según su edad.
4. **Régimen de seguridad social:** Sistema organizado por el Estado para proporcionar protección y bienestar social a los ciudadanos en diferentes aspectos de sus vidas, como la salud, la jubilación, el desempleo, la discapacidad y otros riesgos sociales (En Colombia se puede dividir entre Contributivo y Subsidiado).
5. **Cantidad de Fallecidos:** Conteo del número de fallecidos por enfermedades o infecciones respiratorias agudas (IRA), según los códigos CIE tomados en cuenta para este estudio.

3.2.3. Tasa de Mortalidad por IRA

La tasa de mortalidad se define como la proporción de defunciones que se producen en un área concreta (país, comunidad, provincia, municipio, etc) respecto a su población total durante un período de tiempo determinado, generalmente a lo largo de un año. Se expresa como el número de muertes por cada 1.000 habitantes y es un indicador clave para evaluar la salud de una población y las condiciones socioeconómicas en una región o país [Datos Macro, sf].

Para calcular la tasa de mortalidad por enfermedades respiratorias agudas en las localidades de Bogotá D.C, se empleó la base de datos del Observatorio de Salud de Bogotá (SaluData), con el registro mensual de los casos de mortalidad en las localidades en Bogotá D.C durante el periodo de 2018 a 2022, junto con la información demográfica del Departamento Administrativo Nacional de Estadística (DANE) relacionada con las proyecciones de población a nivel de localidades de Bogotá para los periodos del año 2018 hasta el 2022.

Dicha tasa se calculó mediante la fórmula:

$$\text{Tasa de mortalidad} = \frac{\text{Número de Fallecidos}}{\text{Población total}} * 1000 \quad (1)$$

3.3. Metodología Desarrollada

3.3.1. Fase 1. Depurar y estructurar las bases de datos

En la primera fase se descargó la información correspondiente a las concentraciones de PM_{10} y $PM_{2,5}$ durante los años 2018 a 2022 de las estaciones de la RMCAB disponibles en su sitio

web, consolidando una base de datos con los registros horarios de este contaminante (posteriormente se estandarizó a una periodicidad mensual con el promedio). Así mismo, para la información de mortalidad se accedió a la base de datos del Observatorio de Salud de Bogotá (SaluData) con el registro mensual de los casos de mortalidad por localidades en Bogotá D.C durante el periodo de 2018 a 2022.

De igual forma, se realiza la verificación de valores faltantes o nulos entendidos como aquellos elementos que no tienen un valor asignado o registrado. Estos valores pueden deberse a diversos motivos, como errores en la entrada de datos, falta de información o problemas en la recopilación de datos. Para este caso, en algunas ocasiones las estaciones de la RMCAB se encuentran fuera de servicio o en mantenimiento, por ende, en ciertas horas del día no se registran mediciones de estos contaminantes, viéndose reflejados como valores o datos faltantes. Estos vacíos en la información pueden ser un problema para el análisis y toma de decisiones, para lo cual es importante identificar y manejar adecuadamente con el fin de evitar distorsiones en el análisis y modelos que se realizan en próximos capítulos.

En los Anexos A y B, se observan los registros originales con presencia de datos faltantes. Para abordar esta situación, se emplea el método de imputación k - Nearest Neighbors kNN (k - vecinos más cercanos), este método estima los valores faltantes al calcular la media de los valores de los k vecinos u observaciones más cercanos. Para llevar a cabo esta imputación, se utilizó la clase `KNNImputer` proporcionada por la librería `scikit-learn`, esta clase facilita la implementación de este modelo para la imputación de valores faltantes. Cabe destacar que, de manera predeterminada la clase utiliza la distancia euclidiana (distancia ordinaria entre dos puntos de un espacio euclídeo, la cual se deduce a partir del teorema de Pitágoras) para calcular la proximidad entre vecinos.

Los valores finales de los promedios mensuales de medición de materia particulada $PM_{2,5}$ y PM_{10} se encuentran detallados en los Anexos C y D, estos datos representan la información con la cual se llevó a cabo los análisis. A partir de estos valores consolidados, se procede con el desarrollo y la interpretación de los análisis estadísticos pertinentes.

3.3.2. Fase 2. Realizar un análisis exploratorio y espacial

Con las bases de datos consolidadas en la segunda fase se busca comprender y caracterizar el comportamiento de los datos georreferenciados mediante un análisis exploratorio y espacial, esto permite identificar patrones, relaciones y tendencias que podrían ser relevantes para el estudio. Para ello se verifica que los datos estén completos, no tengan valores atípicos o errores de registro, y que estén correctamente georreferenciados, así mismo se procede a estandarizar los datos con una temporalidad mensual con el fin de garantizar el objetivo al

Tabla 3-2.: Porcentaje de datos faltantes PM_{10} y $PM_{2,5}$

Localidad	PM_{10}	$PM_{2,5}$
Chapinero	25 %	24 %
Ciudad Bolivar	57 %	57 %
Engativá	25 %	25 %
Fontibón	44 %	44 %
Kennedy	15 %	14 %
Puente Aranda	38 %	34 %
San Cristóbal	5 %	13 %
Santa Fé	12 %	9 %
Suba	26 %	27 %
Tunjuelito	3 %	7 %
Usaquén	15 %	15 %
Usme	57 %	58 %

unificar la periodicidad de las bases de datos.

Posterior, se realiza un análisis exploratorio y descriptivo de las variables de interés y covariables adicionales (rango etario y régimen de afiliación al sistema de salud), con el fin de identificar tendencias temporales o patrones de comportamiento, y su posterior visualización por medio de gráficos como diagramas de caja, mapas de densidad, mapas de calor, mapas de distribución espacial e histogramas para representar el comportamiento de las variables. Así mismo, por medio de series de tiempo se profundiza el análisis del comportamiento del material particulado en las 12 localidades de Bogotá.

3.3.3. Fase 3. Desarrollar métodos de interpolación espacial

Como se mencionó en las fases anteriores se cuenta con registros de la concentración de material particulado en 12 de las 20 localidades del distrito de Bogotá, correspondientes a las 20 estaciones de las RMCAB instaladas en diferentes lugares de la ciudad, por tanto, y con el fin de estimar el valor promedio del contaminante (PM_{10} y $PM_{2,5}$) dentro de cada localidad, se desarrollan métodos de interpolación espacial.

En consecuencia, se realiza un análisis geoestadístico de los datos y la visualización de su capa geográfica, con el fin de entender inicialmente su comportamiento y tendencia, y así determinar los métodos de interpolación - predicción espacial más adecuados. Así mismo,

se establece el área de interés en la que se realizará la interpolación, es decir se define y construye la grilla o malla que cubrirá el área de estudio a interpolar.

Teniendo en cuenta el comportamiento de los datos, se realizan métodos de interpolación espacial a cuatro cortes (enero, abril, julio y octubre) en el último año de observación (2022), con el fin de observar los cambios del material particulado en este periodo de tiempo. Por último se realizan procesos de validación cruzada para evaluar la capacidad de predicción de los modelos. Con dichos resultados se generan mapas de las concentraciones estimadas para PM_{10} y $PM_{2,5}$ en las localidades de Bogotá.

3.3.4. Fase 4. Identificar y estandarizar los datos de mortalidad

Como se mencionó en las fases anteriores, para consolidar los datos de mortalidad se tomaron 161 códigos CIE relacionados con causas de mortalidad por IRA (Anexo E) con una periodicidad mensual, para su posterior análisis y estimación de modelos espacio temporales.

3.3.5. Fase 5. Estimar un modelo espacio-temporal

Una vez realizadas las fases anteriores, se desarrollan y ajustan modelos espaciales y espacio temporales que permiten analizar la relación entre la contaminación por materia particulada (PM_{10} y $PM_{2,5}$) y la tasa de mortalidad por enfermedades respiratorias en 12 localidades de Bogotá D.C, para ello se utiliza la base de datos estandarizada y georreferenciada generada en las fases anteriores, con los promedios mensuales de las concentraciones del contaminante y las tasas de mortalidad por enfermedades respiratorias desagregadas por área geográfica y tiempo.

Se realizan modelos espaciales definiendo como variable dependiente la tasa de mortalidad por enfermedades respiratorias, como variables independientes las concentraciones de PM_{10} y $PM_{2,5}$ en el espacio (localidades) y como covariable adicional el régimen de afiliación al sistema de salud. Así mismo, se evalúa la calidad de ajuste de los modelos, seleccionando el mejor modelo para su análisis.

Finalmente se agrega la componente temporal (mensual), al realizar un modelo espacio temporal, se presenta su análisis y mapas que representan la evolución espacio-temporal de las concentraciones de PM_{10} , $PM_{2,5}$ y la tasa de mortalidad por enfermedades respiratorias.

3.3.6. Fase 6. Documentar la relación entre las variables

A partir de los modelos ajustados en las fases anteriores, se establecen conclusiones claras sobre el impacto de este contaminante en la salud pública, presentando los hallazgos de la relación entre la contaminación por material particulado (PM_{10} , $PM_{2,5}$) y la tasa de mortalidad por enfermedades respiratorias en Bogotá D.C durante el periodo 2018-2022.

4. Análisis exploratorio de los datos

Este capítulo se centra en el análisis exploratorio de los datos de contaminación atmosférica por material particulado y mortalidad por enfermedades respiratorias agudas en Bogotá D.C durante 2018-2022. Se inicia con la exploración inicial de los datos para comprender la distribución y el comportamiento de las variables mediante análisis estadístico descriptivo y gráfico. Esta exploración inicial sienta las bases para el posterior modelado estadístico y la interpretación de resultados, permitiendo una comprensión más profunda de la dinámica de la contaminación por materia particulada y la tasa de mortalidad por IRA en Bogotá.

El análisis exploratorio de datos es una etapa crucial en cualquier estudio estadístico, ya que permite familiarizarse con las características de las variables. En este caso, este análisis permite una comprensión más profunda de la dinámica de la contaminación atmosférica por materia particulada y el comportamiento de la tasa de mortalidad por enfermedades o infecciones respiratorias en Bogotá D.C durante el período de estudio.

4.1. Concentración por Material Particulado (PM)

4.1.1. Descripción de datos por hora de Material Particulado (PM)

Como se mencionó anteriormente, se cuenta con una base de datos con los registros horarios de material particulado en el período de estudio comprendido entre el año 2018 hasta el 2022. La tabla **4-1** muestra los formatos de las variables a analizar y el tipo de las mismas: Fecha (Datetime), PM_{10} (Float), $PM_{2.5}$ (Float), Estación (Object), Latitud (Object), Longitud (Object) y Localidad (Object).

La Tabla **4-2** exhibe las estadísticas descriptivas básicas de la base de datos inicial, que comprende un total de 832.605 filas con registros horarios de la concentración de materia particulada desde 2018 a 2022. En el caso de PM_{10} el valor del percentil 50 (20) es menor que el promedio (24.91), esto indica que el 50 % de los registros de PM_{10} están por debajo de $20\mu g/m^3$, lo que sugiere que hay un sesgo positivo y tiene cola hacia la derecha. Para el caso de $PM_{2.5}$ se observa un comportamiento similar, en donde el valor del percentil 50 (8.9) también es menor que el promedio (12.35), lo que indica que la distribución de los datos podría tener un sesgo hacia la derecha.

Tabla 4-1.: Base de datos concentración de Materia Particulada PM_{10} y $PM_{2,5}$ en Bogotá D.C (2018-2022)

# Registros	Fecha	PM_{10}	$PM_{2,5}$	Estación	Latitud	Longitud	Localidad
1	01/01/2022	90,5	63,8	Ciudad Bolivar	4,577806	-74,166278	Ciudad Bolivar
2	01/01/2022	180,3	133,4	Ciudad Bolivar	4,577806	-74,166278	Ciudad Bolivar
3	01/01/2022	167,0	137,5	Ciudad Bolivar	4,577806	-74,166278	Ciudad Bolivar
4	01/01/2022	162,8	132,3	Ciudad Bolivar	4,577806	-74,166278	Ciudad Bolivar
5	01/01/2022	106,0	85,7	Ciudad Bolivar	4,577806	-74,166278	Ciudad Bolivar
...	...	...	...	...	...	...	...
832601	12/31/2022	64,9	31,8	Usme	4,532056	-74,11714	Usme
832602	12/31/2022	81,8	38,7	Usme	4,532056	-74,117139	Usme
832603	12/31/2022	78,8	46,7	Usme	4,532056	-74,117139	Usme
832604	12/31/2022	83,2	51,9	Usme	4,532056	-74,117139	Usme
832605	12/31/2022	72,6	58,1	Usme	4,532056	-74,117139	Usme

Tabla 4-2.: Resumen de los datos de Materia Particulada (PM) a nivel distrital

index	PM_{10}	$PM_{2,5}$
count	832.605	832.605
mean	24,91	12,35
std	25,63	13,37
min	-	-
25 %	-	-
50 %	20,0	8,9
75 %	38,2	19,1
max	996,0	210

Diagramas de caja

Los diagramas de caja también conocidos como diagramas de bigotes o boxplots, son una forma o herramienta visual o gráfica de representar la distribución de un conjunto de datos, mostrando de manera concisa diversas medidas estadísticas como la mediana, los cuartiles y los valores extremos o atípicos. Su utilidad radica principalmente, en que ofrecen una rápida visualización de la dispersión y los valores atípicos en un conjunto de datos, permitiendo identificar de forma intuitiva la presencia de valores extremos, así como la simetría o asimetría

de la distribución. Esto los convierte en una herramienta para el análisis exploratorio de datos, facilitando la identificación de patrones y la detección de posibles anomalías en los datos.

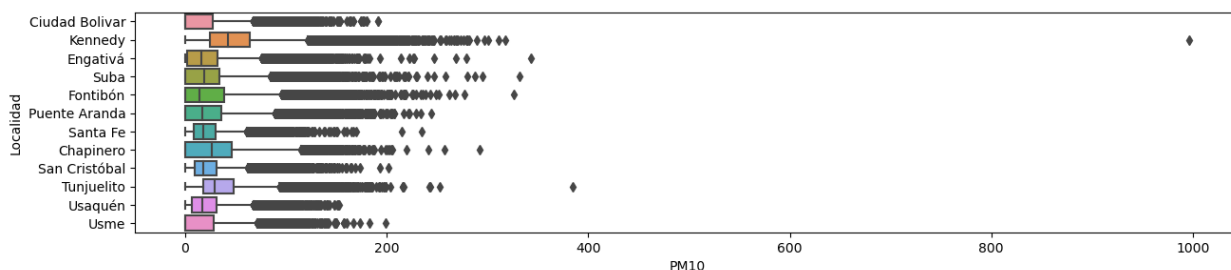


Figura 4-1.: Diagrama de caja concentración Materia Particulada PM_{10} por hora por localidad en Bogotá (2018-2022)

La Figura 4-1 muestra la concentración por hora en microgramos por metro cúbico de Materia Particulada PM_{10} en doce (12) localidades de Bogotá durante el periodo 2018-2022. La escala horizontal del diagrama va de 0 a 1.000, marcando intervalos de 200 en 200.

Inicialmente al observar el comportamiento entre localidades, se evidencia homogeneidad en su distribución, sin embargo, se destaca la presencia de valores atípicos, representados por puntos individuales fuera de la caja, los cuales son indicadores de datos que se encuentran significativamente alejados de la mayoría de las observaciones. En este caso, se observan valores atípicos dispersos hacia la derecha, alcanzando concentraciones cercanas a 1000 en la escala de PM_{10} , lo que muestra una dispersión considerable en los datos.

Con respecto al rango intercuartílico, entendido como la distancia entre el primer y tercer cuartil (longitud de la caja), se evidencia que existe dispersión de los datos alrededor de la mediana. En este caso, se observa que en algunas localidades la mediana o el segundo cuartil no se ubica exactamente en el centro de la caja. Así mismo, se observan mayores concentraciones de material particulado PM_{10} en localidades como Kennedy, Tunjuelito y Chapinero, en donde sus cajas se desplazan hacia la derecha presentando concentraciones del contaminante más elevadas.

En tal sentido, se encuentra que la presencia considerable de valores extremos puede indicar asimetría en los datos de la concentración por PM_{10} . Finalmente, se observa que en las localidades de Ciudad Bolívar y Usme la mediana no se alcanza a visualizar dentro de la caja, sugiriendo que la distribución de los datos es asimétrica, puesto que al no estar la mediana entre el Q_1 y Q_3 podría indicar que hay más valores hacia un extremo de la distribución.

Por otro lado, la Figura 4-2 muestra la concentración por hora en microgramos por metro

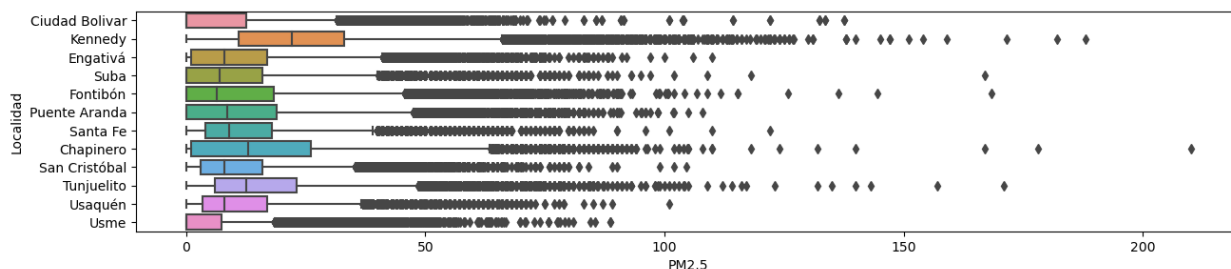


Figura 4-2.: Diagrama de caja concentración Materia Particulada $PM_{2.5}$ por hora por localidad en Bogotá (2018-2022)

cúbico de Materia Particulada $PM_{2.5}$ en doce (12) localidades de Bogotá durante el periodo de estudio. La escala horizontal del diagrama va de 0 a 350, marcando intervalos de 50 en 50.

Se observa un comportamiento muy similar al presentado en la Figura 4-1 correspondiente a la concentración de Materia Particulada PM_{10} , en donde igualmente se evidencia homogeneidad entre localidades, sin embargo, se destaca la presencia de valores atípicos, los cuales indican que al interior de las localidades algunas registros se encuentran significativamente alejados de la mayoría de las observaciones, destacándose valores atípicos dispersos hacia la derecha, los cuales alcanzan concentraciones cercanas a 300 en la escala de $PM_{2.5}$.

Así mismo, se observa que en algunas localidades la mediana no se ubica exactamente en el centro de la caja indicando cierta asimetría de los datos; adicionalmente se evidencia mayor concentración de material particulado $PM_{2.5}$ en localidades como Kennedy, Chapinero y Tunjuelito, en donde sus cajas se desplazan hacia la derecha presentando concentraciones del contaminante más elevadas.

De igual forma se observa que las localidades de Ciudad Bolívar y Usme exhiben una mediana que no está entre el Q_1 y Q_3 , indicando una posible asimetría hacia un extremo de la distribución.

Mapas de calor

A continuación se presenta un resumen gráfico de los Anexos A y B, por medio de mapas de calor bidimensionales, los cuales se definen como representaciones visuales que muestran la distribución de los datos en dos dimensiones mediante el uso de colores y por medio de los cuales se pueden identificar patrones, tendencias o variaciones en conjuntos de datos bidimensionales.

Para este caso se muestran dos mapas de calor, que permiten tener una visión de los periodos

de tiempo y localidades en que se presentan mayores o menores niveles de contaminación producto de las materias particuladas PM_{10} y $PM_{2.5}$.

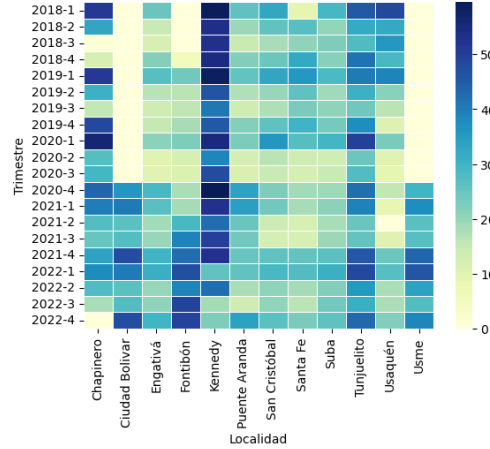


Figura 4-3.: Mapa de calor PM_{10} por trimestre

En la Figura 4-3 se observan que las localidades de Kennedy, Chapinero, Tunjuelito y Fontibón presentan mayor nivel de concentración de la partícula PM_{10} , siendo un indicio de alta contaminación. Así mismo, se evidencia que la localidad de Usaquén tienen principalmente tonos más claros (verde y amarillo), indicando niveles más bajos de PM_{10} a lo largo del tiempo. Con respecto al tiempo, se pueden destacar los periodos de 2019-1, 2020-1, 2020-4, 2021-1, en donde se observan colores más oscuros indicando mayor concentración del contaminante. Finalmente, y teniendo en cuenta la distribución de colores en el gráfico, es posible observar que las emisiones de PM_{10} están concentradas en algunas zonas específicas, probablemente debido a diferentes fuentes de emisión como las fuentes industriales o el tráfico vehicular. Se evidencia en las localidades de Ciudad Bolívar y Usme la presencia de valores faltantes representados por tonos más claros.

De igual forma, en la Figura 4-4, se observa que las localidades de Kennedy, Fontibón, Puente Aranda, Chapinero y Tunjuelito, presentan mayor nivel de concentración de la partícula $PM_{2.5}$, indicando mayores niveles de contaminación. Con respecto al periodo de tiempo, se observan los trimestres 2019-1, 2020-1, 2020-4, 2021-1, 2021-4 y 2022-1 con mayor concentración (representados por colores más oscuros), lo que puede indicar mayores niveles de este contaminante en comparación con otros periodos de tiempo. En general se observa un comportamiento similar al de la partícula PM_{10} , en donde las emisiones se ven concentradas en algunas zonas específicas, y en las localidades de Ciudad Bolívar y Usme se evidencia la presencia de valores faltantes representados por tonos más claros.

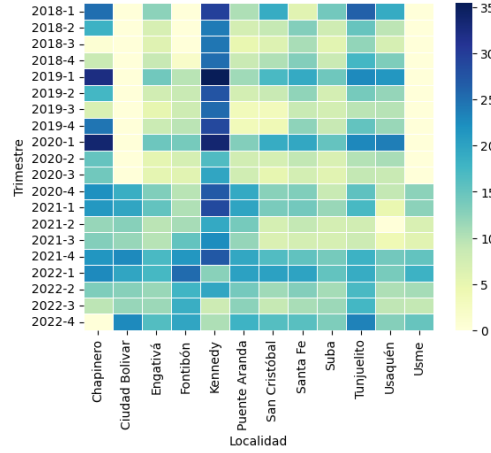


Figura 4-4.: Mapa de calor $PM_{2.5}$ por trimestre

4.1.2. Descripción de datos mensuales de Material Particulado (PM)

Teniendo en cuenta que la investigación busca entender el comportamiento del material particulado PM_{10} y $PM_{2.5}$ con modelos espacio temporales se decide que el tratamiento de los mismos tengan un temporalidad mensual, ya que este enfoque permitirá reducir la cantidad de datos y facilitará la estandarización del análisis.

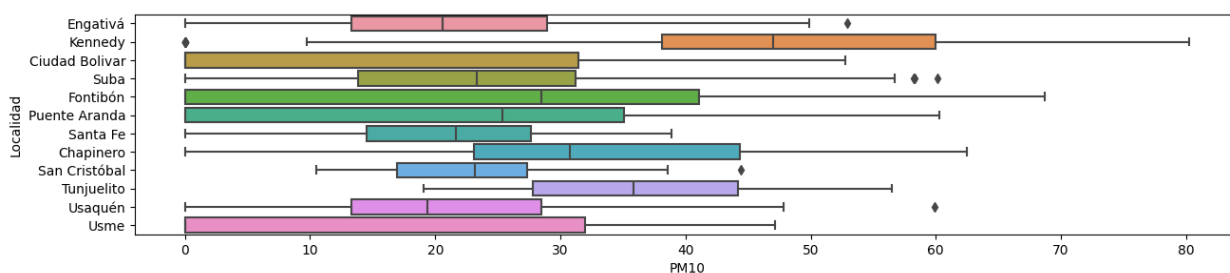
La Tabla 4-3 exhibe las estadísticas descriptivas de los datos mensualizados a nivel de red de monitoreo, que comprende un total de 1.140 filas con registros mensuales de la concentración de materia particulada desde 2018 a 2022. En el caso de PM_{10} el valor del percentil 50 (25.6) es un poco mayor que el promedio (24.95), esto sugiere cierta simetría alrededor de la media con una cola levemente hacia la izquierda. Para el caso de $PM_{2.5}$ se observa un comportamiento similar, en donde el valor del percentil 50 (12.53) también es un poco mayor que el promedio (12.38).

En consecuencia, los Anexos A y B representan el promedio mensual de las partículas PM_{10} y $PM_{2.5}$, respectivamente. Estos registros de datos mensuales consolidados serán las bases para los análisis gráficos, estadísticos y espaciales que se realizarán posteriormente. Así mismo, esta información será fundamental para realizar los modelos espacio temporales que pretende este documento.

Tabla 4-3.: Resumen de los datos mensuales de Materia Particulada (PM) a nivel distrital

index	PM_{10}	$PM_{2,5}$
count	1.140	1.140
mean	24,95	12,38
std	17,92	9,18
min	-	-
25 %	12,14	5,87
50 %	25,60	12,53
75 %	37,18	17,97
max	80,2	49,77

Diagramas de caja

**Figura 4-5.:** Diagrama de caja concentración Materia Particulada PM_{10} mensual por localidad en Bogotá (2018-2022)

La Figura 4-5 muestra la concentración en microgramos mensuales por metro cúbico de Materia Particulada PM_{10} en doce (12) localidades de Bogotá durante el periodo 2018-2022. La escala horizontal del diagrama va de 0 a 80, marcando intervalos de 10 en 10. Se destaca que la localidad de Kennedy presenta los niveles más altos de contaminación por materia particulada PM_{10} , mientras que Ciudad Bolívar y Usme parecieran ser las localidades con niveles más bajos.

Por otro lado, la Figura 4-6 muestra la concentración en microgramos mensual por metro cúbico de Materia Particulada $PM_{2,5}$ en doce (12) localidades de Bogotá durante el periodo 2018-2022. La escala horizontal del diagrama va de 0 a 50, marcando intervalos de 10 en 10. Se destaca que la localidad de Kennedy presenta los niveles más altos de contaminación por materia particulada $PM_{2,5}$, y presenta el dato atípico más alto, mientras que Ciudad Bolívar y Usme parecieran ser las localidades con niveles más bajos; se destaca la localidad de Suba con la mayor cantidad de valores atípicos, consistente con lo observado en PM_{10} .

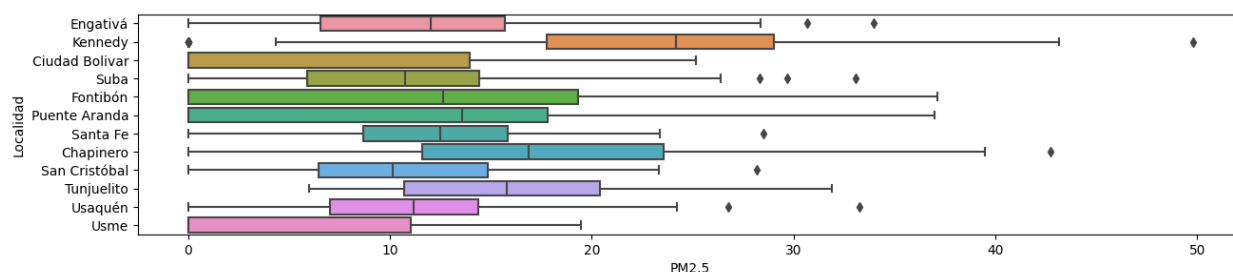


Figura 4-6.: Diagrama de caja concentración Materia Particulada $PM_{2,5}$ mensual por localidad en Bogotá (2018-2022)

Histogramas de frecuencia

Los histogramas de frecuencia se utilizan como herramienta gráfica en el análisis de las propiedades de las variables en un conjunto de datos. Estos gráficos brindan la capacidad de explorar y comprender en detalle la distribución de cada atributo de manera individual.

En el marco de este trabajo de investigación, se elaboran histogramas de frecuencia para cada una de las variables de estudio. Este enfoque, permite detectar patrones y tendencias relacionadas con la dispersión y simetría de los datos, como se evidencia en las Figuras 4-7 y 4-8. Uno de los principales hallazgos obtenidos a partir de este análisis, se centra en la función de distribución de las materias particuladas, encontrando que el histograma de frecuencia de la materia particulada PM_{10} muestra una distribución que parece simétrica, esto sugiere que los datos podrían seguir una distribución normal o algo cercano a ella, mientras que para la materia particulada $PM_{2,5}$ el histograma muestra una distribución asimétrica hacia la derecha, indicando que la mayoría de los datos se agrupan en los valores más bajos (a la izquierda del histograma), mientras que los valores más altos (a la derecha) son menos frecuentes.

El análisis del histograma de frecuencia de la materia particulada PM_{10} revela una concentración significativa en el rango de $15\mu g/m^3$ a $45\mu g/m^3$. Esto indica que en su mayoría, los registros de la concentración de este contaminante se encuentran en los rangos establecidos en Tabla 2-1.

Por otro lado, en la Figura 4-8 se ilustra el comportamiento de la distribución de la partícula $PM_{2,5}$. Este histograma muestra una distribución ligeramente menos simétrica en comparación con la PM_{10} . Se observa una mayor concentración de datos en el intervalo de $5\mu g/m^3$ a $20\mu g/m^3$.

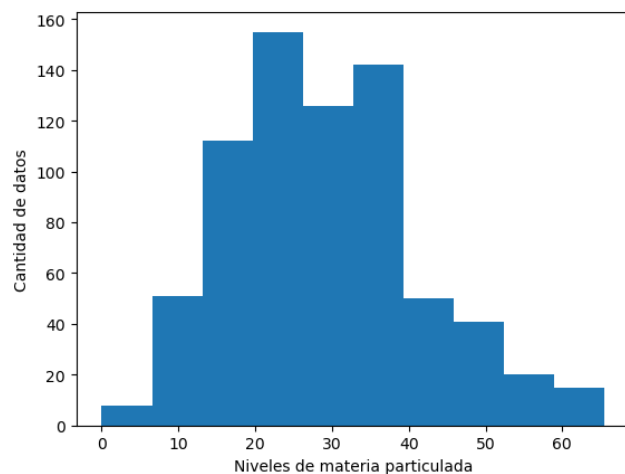


Figura 4-7.: Histograma de Frecuencia PM_{10}

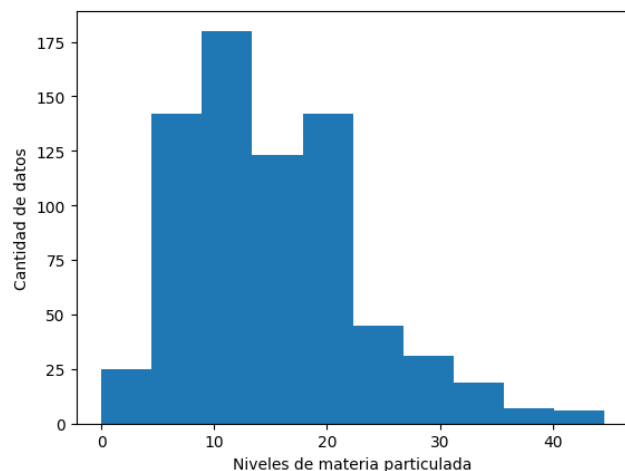


Figura 4-8.: Histograma de Frecuencia $PM_{2,5}$

Mapas de calor

A continuación se muestran dos mapas de calor para los datos mensualizados que permiten tener una visión de los periodos de tiempo y localidades en que se presentan mayores o menores niveles de contaminación producto de las materias particuladas PM_{10} y $PM_{2,5}$.

Las Figuras 4-9 y 4-10 presentan variaciones similares de los contaminantes de las materias particuladas PM_{10} y $PM_{2,5}$ con los presentados en las figuras 4-3 y 4-4. Lo que muestra consistencia en la información y en los análisis

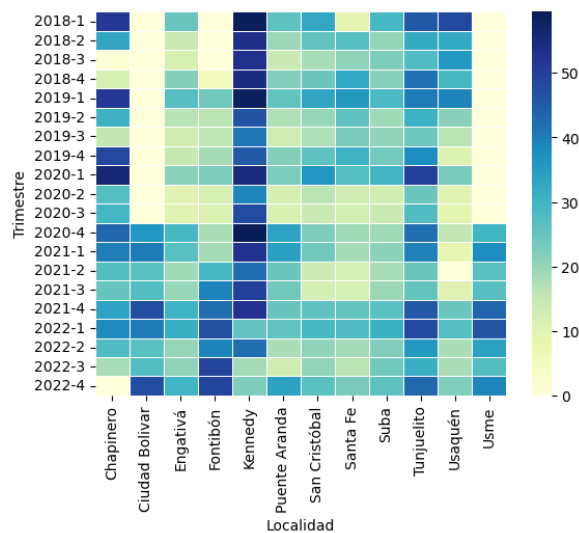


Figura 4-9.: Mapa de calor PM_{10} por trimestre

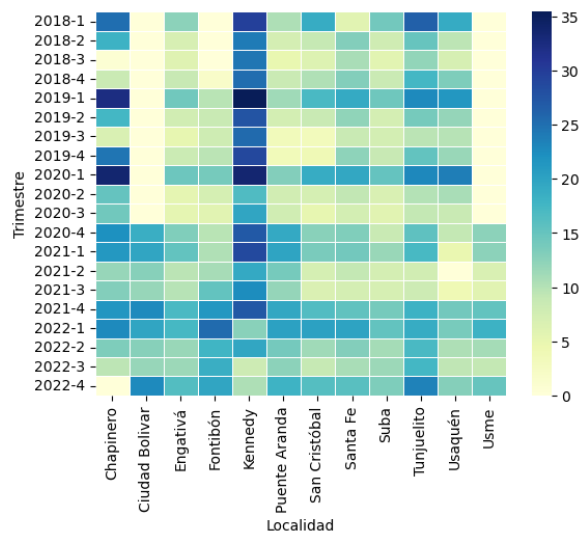


Figura 4-10.: Mapa de calor $PM_{2.5}$ por trimestre

Visualización de los datos sin registros faltantes

Por otro lado, y una vez aplicado el método de imputación de valores faltantes k - Nearest Neighbors kNN (k - vecinos más cercanos) como se mencionó en la Fase 1 de la metodología desarrollada, se presenta en la tabla 4-4 el cambio en la distribución de los datos al tener una periodicidad mensual sin valores faltantes o nulos, resaltándose que aproximadamente el 50 % de los registros de las partículas PM_{10} y $PM_{2,5}$ se encuentran por debajo de la media, lo que indica que hay valores más altos en el extremo derecho de la distribución, lo que puede deberse a la presencia de valores atípicos o a una concentración de observaciones en ese extremo. Sin embargo, se puede observar en general una distribución simétrica, puesto que la media está muy cerca de la mediana.

Tabla 4-4.: Resumen datos promedios de Materia Particulada (PM) a nivel distrital

index	PM_{10}	$PM_{2,5}$
count	1.140	1.140
mean	32,23	15,60
std	13,14	6,86
min	0,008	0,0125
25 %	22,21	11,01
50 %	31,63	14,98
75 %	39,92	19,00
max	80,21	49,77

Tabla 4-5.: Base de datos consolidada por localidades

# Registros	Fecha	Estación	Localidad	Latitud	Longitud	PM_{10}	$PM_{2,5}$
0	201801	Bolivia	Engativá	4,73	-74,12	33,24	14,93
1	201801	Carvajal - Sevillana	Kennedy	4,59	-74,14	59,91	26,45
2	201801	Centro de Alto Rendimiento	Engativá	4,65	-74,08	27,86	15,13
3	201801	Ciudad Bolívar	Ciudad Bolívar	4,57	-74,16	35,58	16,98
4	201801	Colina	Suba	4,73	-74,06	16,16	8,40
...	...	...	...	...	...	...	...
1135	202212	San Cristóbal	San Cristóbal	4,57	-74,08	29,06	18,45
1136	202212	Suba	Suba	4,76	-74,09	36,99	16,37
1137	202212	Tunal	Tunjuelito	4,57	-74,13	47,19	24,74
1138	202212	Usaquén	Usaquén	4,71	-74,03	27,75	14,44
1139	202212	Usme	Usme	4,53	-74,11	37,09	14,19

4.2. Mortalidad por Enfermedades-Infecciones Respiratorias Agudas (IRA)

4.2.1. Exploración inicial de los datos

Para analizar la mortalidad por IRA en Bogotá entre 2018 y 2022, se tomó la información del Observatorio de Salud de Bogotá. La base de datos cuenta con 578.222 filas y 5 columnas (fecha, localidad, rango de edad, régimen de seguridad social y número de muertes en Bogotá), donde se clasificó y se filtró con los 161 códigos CIE seleccionados en el Anexo E.

A continuación se presenta la Tabla 4-6 con los registros mensuales de la tasa de mortalidad por IRA:

Tabla 4-6.: Registro Mensual de mortalidad por IRA 2018-2022 en Bogotá

No.	Fecha	Localidad	Rango etario	Regimen de seguridad social	Tasa de Mortalidad
0	201801	Antonio Nariño	60 a 64 años	SUBSIDIADO	0.056
1	201801	Antonio Nariño	80 a 84 años	CONTRIBUTIVO	0.209
2	201801	Antonio Nariño	85 a 89 años	CONTRIBUTIVO	0.366
3	201801	Barrios Unidos	00 a 04 años	CONTRIBUTIVO	0.050
4	201801	Barrios Unidos	75 a 79 años	SUBSIDIADO	0.080
...	...	...	...	...	...
10502	202212	Usme	60 a 64 años	CONTRIBUTIVO	0.013
10503	202212	Usme	80 a 84 años	SUBSIDIADO	0.128
10504	202212	Usme	85 a 89 años	CONTRIBUTIVO	0.396
10505	202212	Usme	85 a 89 años	SUBSIDIADO	0.561
10506	202212	Usme	90 a 94 años	CONTRIBUTIVO	0.978

Diagramas de caja

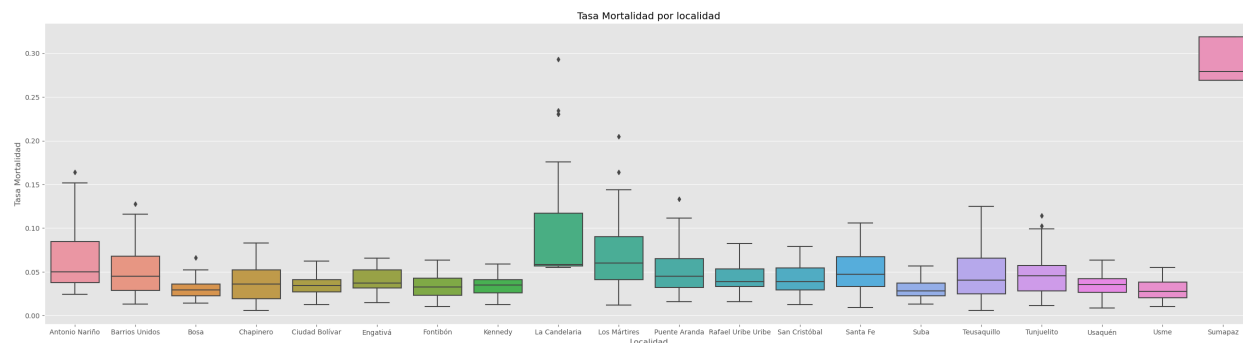


Figura 4-11.: Diagramas de caja tasa de mortalidad por IRA por localidades en Bogotá (2018-2022)

La Figura 4-11 muestra la distribución de la tasa de mortalidad por enfermedades relacionadas con temas respiratorios en Bogotá durante el periodo de tiempo de 2018 a 2022. En tal sentido, se observa que las localidades de Antonio Nariño, La Candelaria, Los Martires y Sumapaz presentan las tasas de mortalidad más altas. Por otro lado, localidades como Suba, Kennedy, Usaquén y Bosa presentan las tasas de mortalidad más bajas de Bogotá.

Se observa un comportamiento atípico de la localidad de Sumapaz, debido a la poca cantidad de registros que hay y sus condiciones particulares, lo que genera que su variabilidad sea mayor que las demás.

Mapas de calor

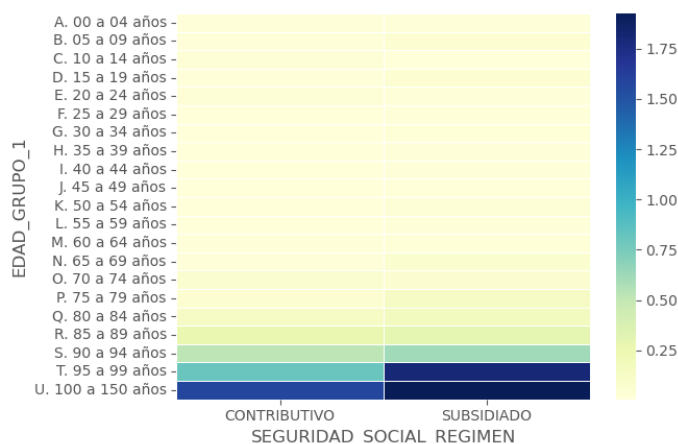


Figura 4-12.: Mapa de Calor Régimen vs Rango Etario

4.2 Mortalidad por Enfermedades-Infecciones Respiratorias Agudas (IRA)

La Figura 4-12 representa visualmente la distribución de la tasa de mortalidad en Bogotá por régimen de afiliación y rango etario, evidenciándose claramente una mayor concentración de defunciones por enfermedades respiratorias en los rangos etarios de 90 años en adelante, tanto en el régimen contributivo como en el subsidiado.

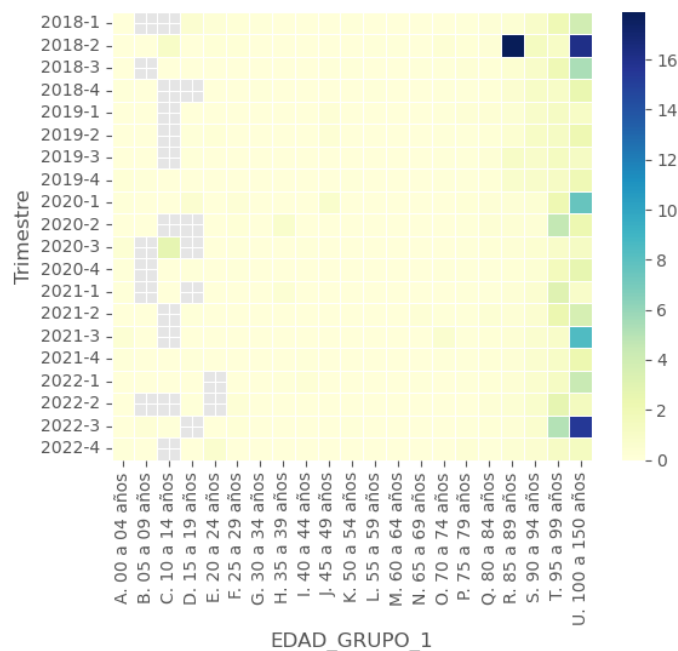


Figura 4-13.: Mapa de Calor Trimestres vs Rango Etario

Por otro lado, la Figura 4-13 representa la tasa de mortalidad por enfermedades respiratorias en Bogotá por trimestres de tiempo y rango etario, donde se observa la misma tendencia que en la figura anterior, la concentración más alta de defunciones se encuentra en los rangos etarios de 90 años en adelante. Así mismo, con respecto al tiempo se evidencia una disminución aparente de la tasa de mortalidad en el año 2020, esto puede deberse a que durante la pandemia por COVID-19, la población estuvo menos expuesta a los virus y la contaminación, o se atribuyeron más muertes al COVID que a IRA.

La Figura 4-14 representa visualmente la concentración de la tasa de mortalidad por enfermedades respiratorias en Bogotá por trimestres de tiempo y localidades, evidenciándose una tendencia similar a la expuesta en figuras anteriores, en donde localidades como Antonio Nariño, La Candelaria, Los Martires y Sumapaz presentan mayores tasas de mortalidad. De igual forma, con respecto al periodo de tiempo, se evidencia mayor concentración en los dos primeros años de estudio (2018 y 2019).

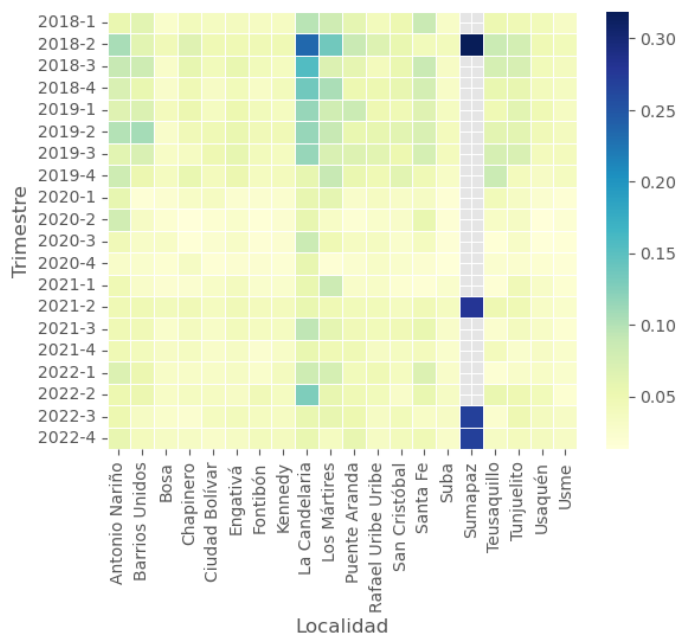


Figura 4-14.: Mapa de Calor Trimestres vs Localidad

Los hallazgos encontrados en estos mapas de calor, indican que hay aspectos como la ubicación geográfica (localidades), periodos de tiempo, rangos etarios y mayor presencia de contaminantes a nivel de Bogotá, que podrían estar afectando la tasa de mortalidad.

Para abordar este aspecto, en capítulos posteriores se realizarán modelos espaciales y espacio temporales para examinar la influencia de factores como el régimen de afiliación en las tasas de mortalidad por enfermedades respiratorias en Bogotá D.C. durante el periodo 2018-2022, así como su influencia según sea el espacio o ubicación geográfica (localidades) y el periodo de tiempo.

4.3. Relación entre Material Particulado (PM) y Tasa de Mortalidad por IRA

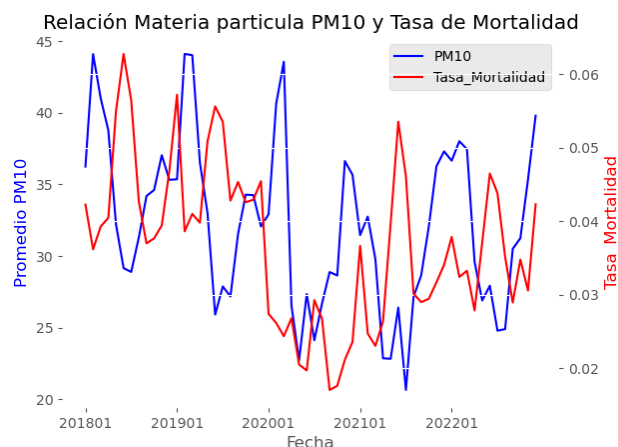


Figura 4-15.: Relación Material Particulado (PM_{10}) y Tasa de Mortalidad por IRA

La Figura 4-15 ilustra la relación entre la tasa de mortalidad y el promedio de contaminación por material particulado PM_{10} en el período comprendido entre enero de 2018 y diciembre de 2022. Si bien la gráfica sugiere una posible asociación entre la contaminación y el aumento de decesos en períodos posteriores, se requieren análisis estadísticos más rigurosos para confirmar esta relación.

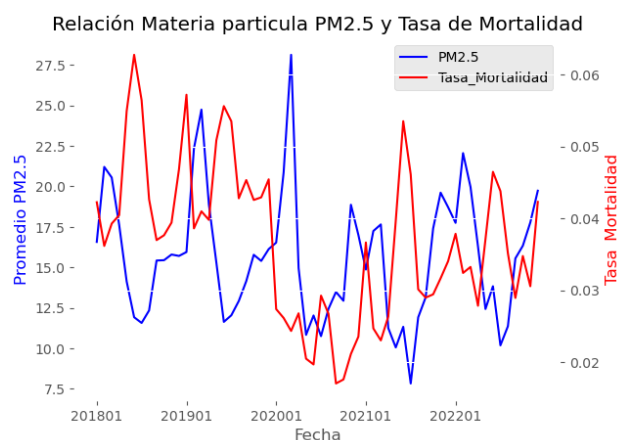


Figura 4-16.: Relación Material Particulado ($PM_{2,5}$) y Tasa de Mortalidad por IRA

La Figura 4-16 ilustra la tendencia de la tasa de mortalidad en relación con el promedio de contaminación por material particulado $PM_{2,5}$ en el período comprendido entre enero de 2018 y diciembre de 2022. A simple vista, no se observa una relación evidente entre ambas variables.

5. Análisis de tendencias temporales y patrones de comportamiento de los niveles de materia particulada (PM_{10} y $PM_{2,5}$) en Bogotá D.C

Este capítulo se centra en analizar el comportamiento de las partículas de materia PM_{10} y $PM_{2,5}$ durante los períodos analizados (2018-2022). Para ello, se utilizan modelos estadísticos de series de tiempo y gráficos auxiliares (PACF y ACF), así como paquetes estadísticos para pronosticar las series temporales. Para este caso, se hará énfasis en los modelos denominados ARMA y ARIMA.

Por medio de este capítulo, se generan predicciones para que las entidades gubernamentales como la Secretaría Distrital de Ambiente y la Alcaldía de Bogotá tengan insumos sobre los posibles niveles de contaminación en la ciudad de Bogotá. Este tipo de resultados contribuyen a mitigar los niveles de contaminación, ya que permite focalizar el tiempo en el cual se presentan, para que el gobierno por medio de distintas políticas públicas incidan en su control.

5.1. Análisis series de tiempo

En esta sección se aplican modelos ARIMA a las series de tiempo correspondiente al promedio de contaminación por mes de las materias particuladas PM_{10} y $PM_{2,5}$ de 2018 a 2022.

Para comenzar con el análisis de las series de tiempo se presenta las Figuras 5-1, 5-2 y 5-3 correspondiente a las partículas PM_{10} y $PM_{2,5}$, con el fin de tener una primera vista sobre el comportamiento de estos datos. Para la serie de tiempo de material particulado PM_{10} se evidencia que el pico ocurre entre los meses de febrero y marzo de todos los periodos, luego de un decaimiento hasta aproximadamente mitad de año. De igual forma, para la serie de tiempo de material particulado $PM_{2,5}$ se evidencia que el pico ocurre entre los meses de febrero y marzo de todos los periodos luego de un decaimiento hasta aproximadamente mitad de año.

Adicionalmente, en la Figura 5-2 se presentan los *Month Plot* de PM_{10} y $PM_{2.5}$, donde pareciera haber presencia de estacionalidad en ambas series, reflejando que los valores más altos ocurren a inicios de año y los valores más bajos alrededor de junio y julio, finalizando el año con una tendencia creciente.

Por otro lado, la Figura 5-3 muestra la distribución mensual de PM_{10} y $PM_{2.5}$ mediante diagramas de caja, lo cual permite identificar la variabilidad y presencia de valores atípicos. Se observa que los meses de enero, febrero y marzo presentan medianas elevadas, lo que refuerza la evidencia de un posible comportamiento estacional.

Estos análisis iniciales permiten comprender mejor la dinámica temporal de las concentraciones de material particulado y justifican la incorporación de componentes estacionales en modelos de predicción como el SARIMA.

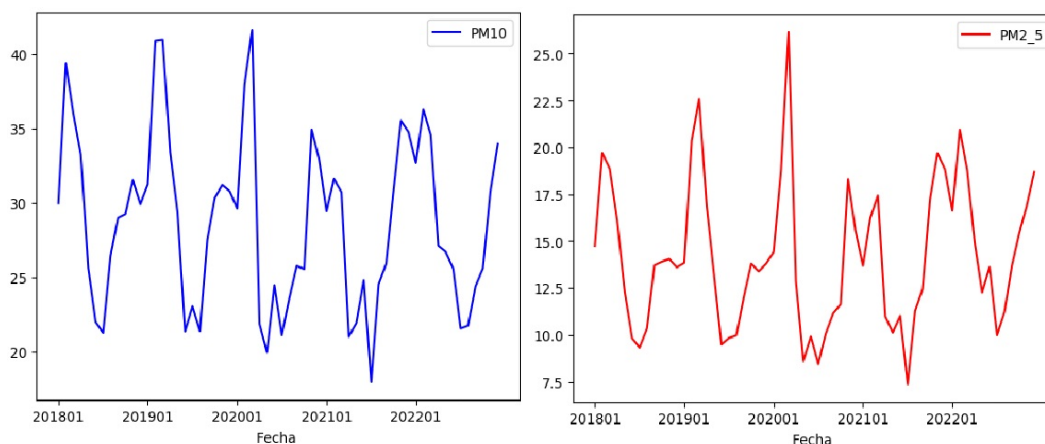


Figura 5-1.: Serie de tiempo de la materia particulada PM_{10} y $PM_{2.5}$

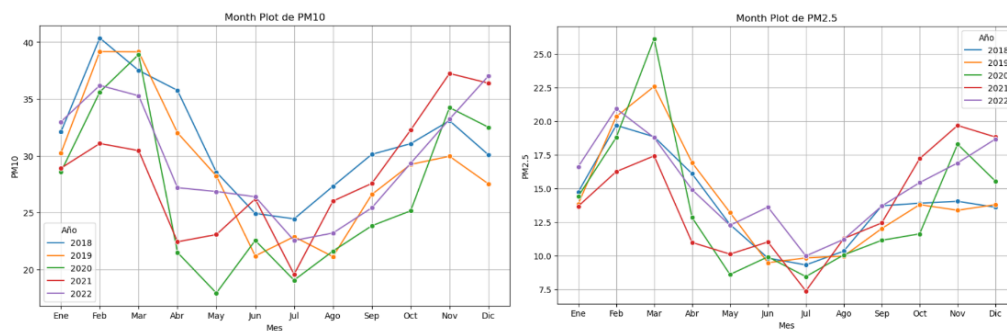


Figura 5-2.: Month Plot PM_{10} y $PM_{2.5}$

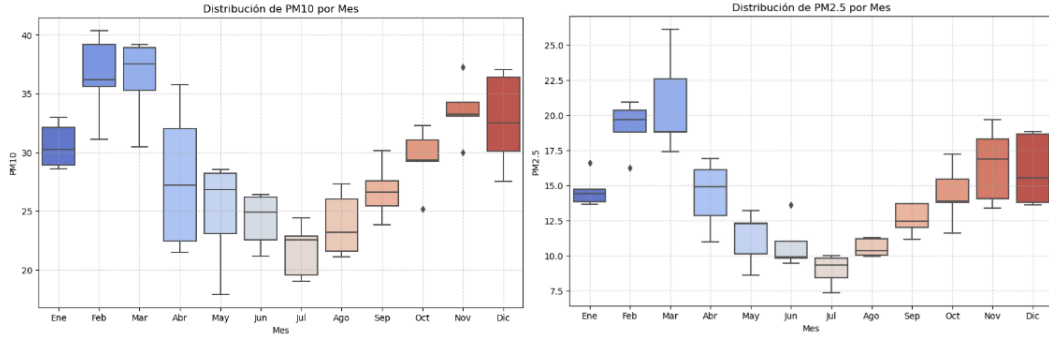


Figura 5-3.: Diagramas de caja Month Plot PM_{10} y $PM_{2,5}$

5.1.1. Función de Autocorrelación:

El término autocorrelación hace referencia a la “correlación entre miembros de series de observaciones ordenadas en el tiempo o en el espacio” [Gujarati and Porter, 2010]. Matemáticamente esta función al rezago k , denotada por p_k , se define como:

$$p_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{covarianza al rezago } k}{\text{varianza}} \quad (1)$$

Por otro lado, se encuentra la Función de Autocorrelación Parcial (FACP), la cual se define como aquella correlación entre observaciones en una determinada cantidad de periodos de tiempo, en donde se elimina el efecto de los rezagos intermedios, “gran parte de la correlación entre la variable Y en el tiempo t y su valor en el tiempo $t - k$ puede estar influenciada por los rezagos intermedios, por tanto, la FACP elimina la influencia de estos valores” [Gujarati and Porter, 2010].

Al graficar esta función, si las autocorrelaciones decaen rápidamente hacia cero, puede ser indicativo de una serie de tiempo estacionaria. En cambio, si las autocorrelaciones permanecen significativas o disminuyen de manera lenta, podría estar relacionado con una serie no estacionaria.

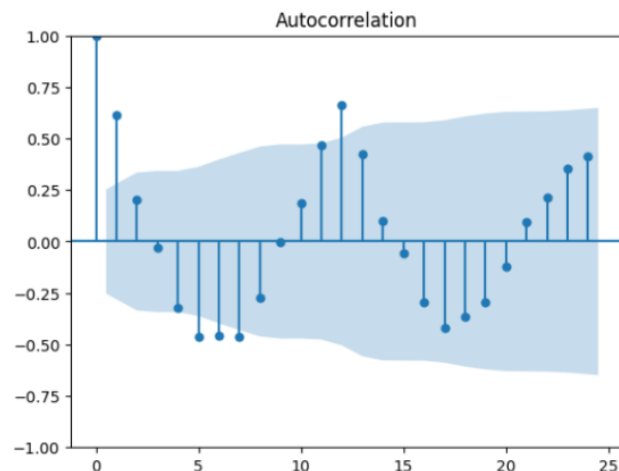
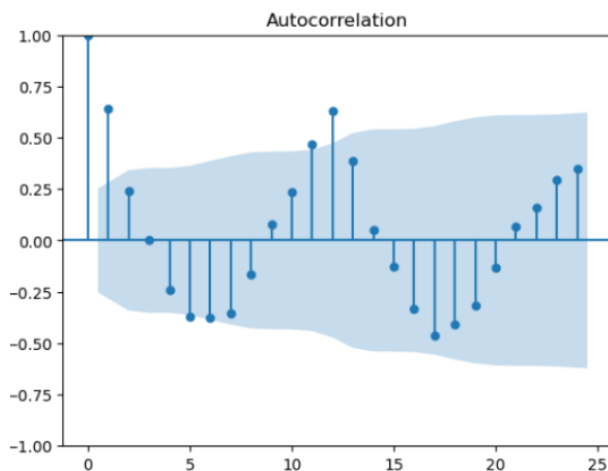
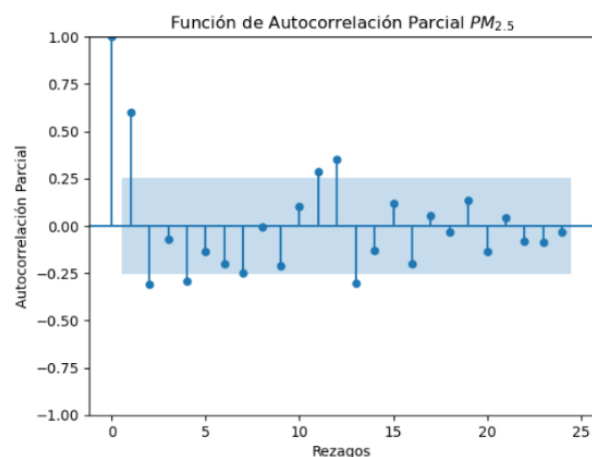
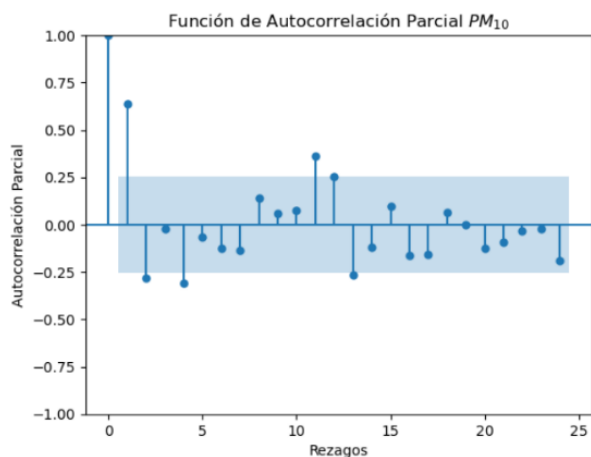
Figura 5-2.: Función de Autocorrelación PM_{10} Figura 5-3.: Función de Autocorrelación $PM_{2.5}$ 

Figura 5-4.: Función de Autocorrelación Parcial

Las figuras anteriores evidencian las funciones de autocorrelación de las series de tiempo para las materias particuladas PM_{10} y $PM_{2.5}$, en donde el valor óptimo para el parámetro p que se va a asignar es 1 para proporcionar un componente autorregresivo al modelo. En cuanto al componente q la función de autocorrelación parcial sugiere un valor de 1.

5.1.2. Estimación de los coeficientes:

Materia particulada PM_{10} :

Tabla 5-1.: Salida de Materia particulada PM_{10} ARIMA

Modelo	AIC	Tiempo
ARIMA(1,1,1)(0,1,1)[12]	AIC=257,571	Time=0,38 sec
ARIMA(0,1,0)(0,1,0)[12]	AIC=267,921	Time=0,05 sec
ARIMA(1,1,0)(1,1,0)[12]	AIC=264,644	Time=0,08 sec
ARIMA(0,1,1)(0,1,1)[12]	AIC=257,743	Time=0,07 sec
ARIMA(1,1,1)(0,1,0)[12]	AIC=259,801	Time=0,06 sec
ARIMA(1,1,1)(1,1,1)[12]	AIC=inf	Time=0,51 sec
ARIMA(1,1,1)(0,1,2)[12]	AIC=inf	Time=0,71 sec
ARIMA(1,1,1)(1,1,0)[12]	AIC=259,386	Time=0,16 sec
ARIMA(1,1,1)(1,1,2)[12]	AIC=inf	Time=1,15 sec
ARIMA(1,1,0)(0,1,1)[12]	AIC=263,000	Time=0,06 sec
ARIMA(2,1,1)(0,1,1)[12]	AIC=259,373	Time=0,24 sec
ARIMA(1,1,2)(0,1,1)[12]	AIC=259,458	Time=0,19 sec
ARIMA(0,1,0)(0,1,1)[12]	AIC=265,149	Time=0,07 sec
ARIMA(0,1,2)(0,1,1)[12]	AIC=257,508	Time=0,08 sec
ARIMA(0,1,2)(0,1,0)[12]	AIC=259,662	Time=0,05 sec
ARIMA(0,1,2)(1,1,1)[12]	AIC=inf	Time=0,33 sec
ARIMA(0,1,2)(0,1,2)[12]	AIC=inf	Time=0,50 sec
ARIMA(0,1,2)(1,1,0)[12]	AIC=259,332	Time=0,08 sec
ARIMA(0,1,2)(1,1,2)[12]	AIC=inf	Time=0,69 sec
ARIMA(0,1,3)(0,1,1)[12]	AIC=259,407	Time=0,13 sec
ARIMA(1,1,3)(0,1,1)[12]	AIC=261,406	Time=0,20 sec
ARIMA(0,1,2)(0,1,1)[12] intercept	AIC=inf	Time=0,42 sec

El modelo seleccionado fue el ARIMA(0,1,2)(0,1,1)[12] debido a que tiene el valor más bajo de AIC (257,508) entre todos los modelos probados en el análisis. Esto significa que este modelo tiene un buen equilibrio entre precisión, parsimonia y con menor varianza residual.

En la parte ARIMA(0,1,2), el componente autorregresivo en este caso es 0, lo que significa que no se están utilizando términos autoregresivos (AR) en el modelo. Los dos términos de media móvil (MA) del modelo capturan el comportamiento a corto plazo de los errores de la serie.

La parte final del modelo $ARIMA(0,1,1)[12]$ sugiere que se incluye un término de media móvil (MA) de orden 1 sin componente autorregresivo ($P = 0$), pero con una diferenciación estacional de orden 1 ($D = 1$) y un término de media móvil estacional (SMA) de orden 1 con un ciclo estacional de 12 períodos ($s = 12$).

Para el modelo seleccionado se encuentra un Mean Absolute Percentage Error (MAPE) de 2,35 % lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 2,35 % de los valores originales.

Materia particulada $PM_{2,5}$:

Tabla 5-2.: Salida de Materia particulada $PM_{2,5}$ ARIMA

Modelo	AIC	Tiempo
ARIMA(1,0,1)(0,1,1)[12] intercept	AIC=225,905	Time=0,55 sec
ARIMA(0,0,0)(0,1,0)[12] intercept	AIC=231,278	Time=0,06 sec
ARIMA(1,0,0)(1,1,0)[12] intercept	AIC=227,397	Time=0,25 sec
ARIMA(0,0,1)(0,1,1)[12] intercept	AIC=224,903	Time=0,25 sec
ARIMA(0,0,0)(0,1,0)[12]	AIC=230,132	Time=0,09 sec
ARIMA(0,0,1)(0,1,0)[12] intercept	AIC=228,342	Time=0,14 sec
ARIMA(0,0,1)(1,1,1)[12] intercept	AIC=inf	Time=0,65 sec
ARIMA(0,0,1)(0,1,2)[12] intercept	AIC=inf	Time=1,00 sec
ARIMA(0,0,1)(1,1,0)[12] intercept	AIC=228,321	Time=0,22 sec
ARIMA(0,0,1)(1,1,2)[12] intercept	AIC=inf	Time=2,04 sec
ARIMA(0,0,0)(0,1,1)[12] intercept	AIC=227,781	Time=0,24 sec
ARIMA(0,0,2)(0,1,1)[12] intercept	AIC=225,683	Time=0,37 sec
ARIMA(1,0,0)(0,1,1)[12] intercept	AIC=223,966	Time=0,28 sec
ARIMA(1,0,0)(0,1,0)[12] intercept	AIC=227,370	Time=0,08 sec
ARIMA(1,0,0)(1,1,1)[12] intercept	AIC=inf	Time=0,81 sec
ARIMA(1,0,0)(0,1,2)[12] intercept	AIC=inf	Time=1,00 sec
ARIMA(1,0,0)(1,1,2)[12] intercept	AIC=inf	Time=1,24 sec
ARIMA(2,0,0)(0,1,1)[12] intercept	AIC=225,888	Time=0,32 sec
ARIMA(2,0,1)(0,1,1)[12] intercept	AIC=227,869	Time=0,61 sec
ARIMA(1,0,0)(0,1,1)[12]	AIC=222,757	Time=0,22 sec
ARIMA(1,0,0)(0,1,0)[12]	AIC=225,754	Time=0,11 sec
ARIMA(1,0,0)(1,1,1)[12]	AIC=inf	Time=0,49 sec
ARIMA(1,0,0)(0,1,2)[12]	AIC=inf	Time=0,77 sec
ARIMA(1,0,0)(1,1,0)[12]	AIC=225,835	Time=0,19 sec
ARIMA(1,0,0)(1,1,2)[12]	AIC=inf	Time=1,12 sec
ARIMA(0,0,0)(0,1,1)[12]	AIC=227,147	Time=0,17 sec
ARIMA(2,0,0)(0,1,1)[12]	AIC=224,596	Time=0,19 sec
ARIMA(1,0,1)(0,1,1)[12]	AIC=224,620	Time=0,20 sec
ARIMA(0,0,1)(0,1,1)[12]	AIC=223,937	Time=0,18 sec
ARIMA(2,0,1)(0,1,1)[12]	AIC=226,623	Time=0,45 sec

El modelo seleccionado fue el ARIMA(1,0,0)(0,1,1)[12] debido a que tiene el valor más bajo de AIC (222,757) entre todos los modelos probados en el análisis.

La parte inicial del modelo $\text{ARIMA}(1,0,0)$ indica que se incluye un término autorregresivo (AR) de orden 1, lo que implica una relación lineal entre el valor actual y el valor en el período anterior, sin diferenciación en la serie temporal ($d = 0$).

La parte final del modelo $\text{ARIMA}(0,1,1)[12]$ incluye un término de media móvil estacional (MA) de orden 1 sin componente autorregresivo ($P = 0$), pero con una diferenciación estacional de orden 1 ($D = 1$) y un término de media móvil estacional (SMA) de orden 1 con un ciclo estacional de 12 períodos ($s = 12$). Esto implica una relación lineal entre el valor actual y los errores residuales de la predicción estacional inmediatamente anterior.

Para el modelo seleccionado se encuentra un Mean Absolute Percentage Error (MAPE) de 6,23% lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 6,23% de los valores originales.

5.1.3. Backtesting:

El metodo backtesting es una técnica utilizada en la modelización y predicción de series temporales, así como en otros campos como el trading financiero, para evaluar el rendimiento de un modelo o estrategia utilizando datos históricos. Este tiene como proposito principal evaluar cómo se comportaría un modelo si se hubiera aplicado en el pasado. Esto proporciona información sobre la capacidad predictiva del modelo y su habilidad para generalizar patrones históricos hacia el futuro.

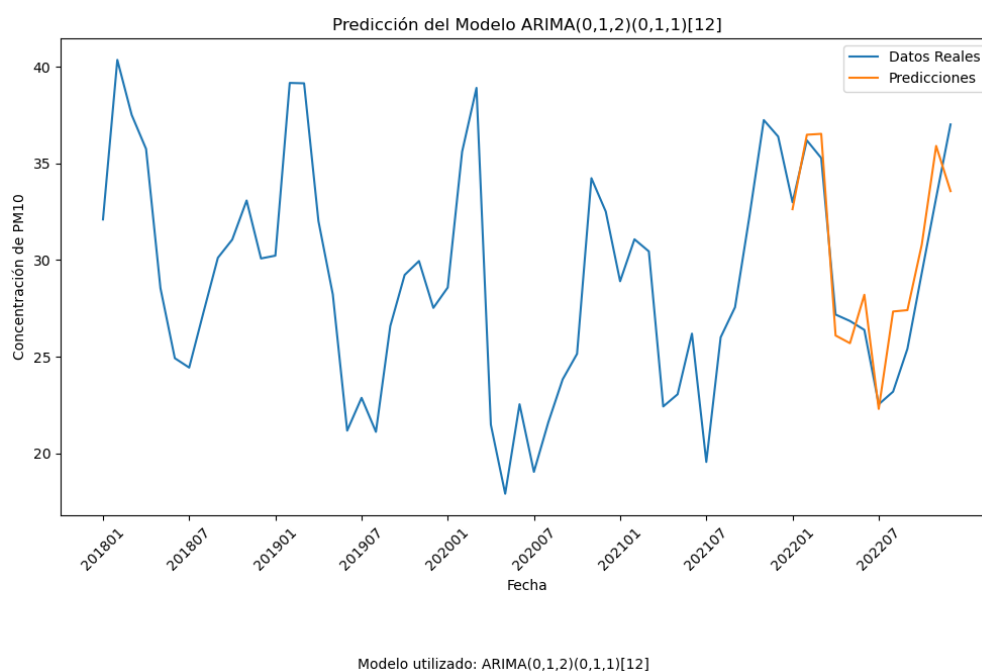


Figura 5-5.: Backtesting PM_{10}

En este caso, el procedimiento fue:

1. Se entrenó un modelo **ARIMA (0, 1, 2)(0, 1, 1, 12)** usando los datos históricos de concentración de **PM_{10}** .
2. **Horizonte de pronóstico:** 12 meses (último año de la serie).
3. **División Train/Test:**
 - a) **Train:** Desde el inicio de los datos hasta 12 meses antes del final.
 - b) **Test:** Los últimos 12 meses.
4. **Entrenamiento:** El modelo fue entrenado únicamente con el conjunto **train**. Los datos del conjunto **test** se utilizaron solo para la evaluación.

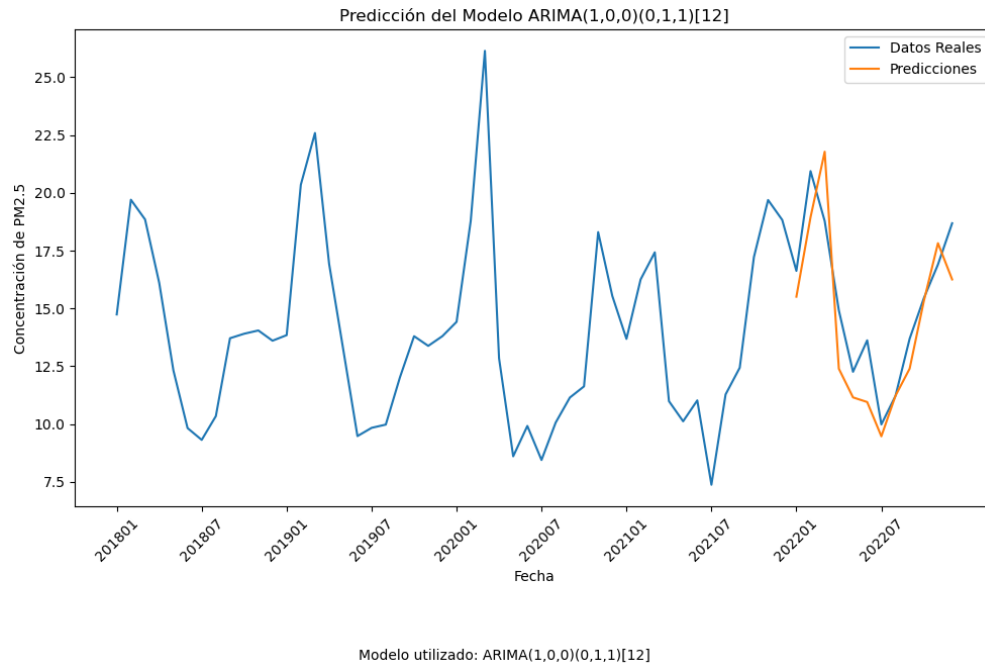


Figura 5-6.: Backtesting $PM_{2,5}$

En este caso, el procedimiento fue:

1. Se entrenó un modelo **ARIMA** $(1, 0, 0)(0, 1, 1, 12)$ usando los datos históricos de concentración de **PM_{2.5}**.
2. **Horizonte de pronóstico:** 12 meses (último año de la serie).
3. **División Train/Test:**
 - a) **Train:** Desde el inicio de los datos hasta 12 meses antes del final.
 - b) **Test:** Los últimos 12 meses.

Las Figuras 5-5 y 5-6 ilustran cómo los modelos ARIMA utilizados para las materias particuladas PM_{10} y $PM_{2,5}$ se ajustaron de manera óptima a los datos históricos. Este ajuste sugiere que las predicciones generadas por estos modelos para ambas materias particuladas serán probablemente precisas y efectivas.

5.1.4. Pronósticos y Raíz del Error Cuadrático Medio:

La raíz del error cuadrático medio es una medida de la precisión de un modelo predictivo. Se calcula como la raíz cuadrada del error cuadrático medio, que a su vez es la media de los cuadrados de las diferencias entre los valores reales y los valores predichos

por el modelo.

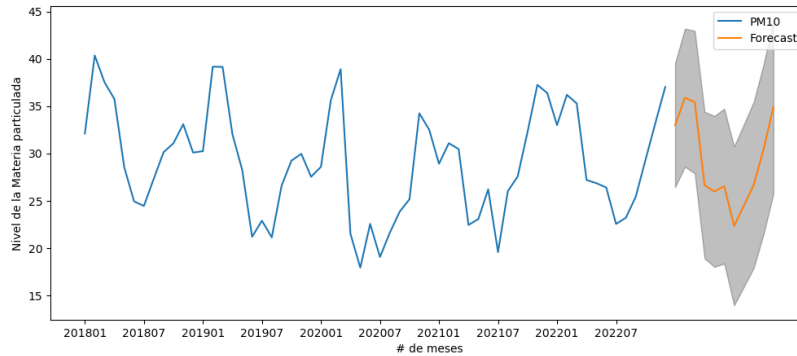


Figura 5-7.: Pronósticos de los modelos PM_{10}

Para la partícula PM_{10} el ECM es de aproximadamente 8,8698 unidades cuadradas, mientras que para $PM_{2,5}$ es de 5,3521. En este caso, los niveles de contaminación atmosférica están medidos en microgramos por metro cúbico ($\mu g/m^3$), por tanto el ECM se mide en $(\mu g/m^3)^2$.

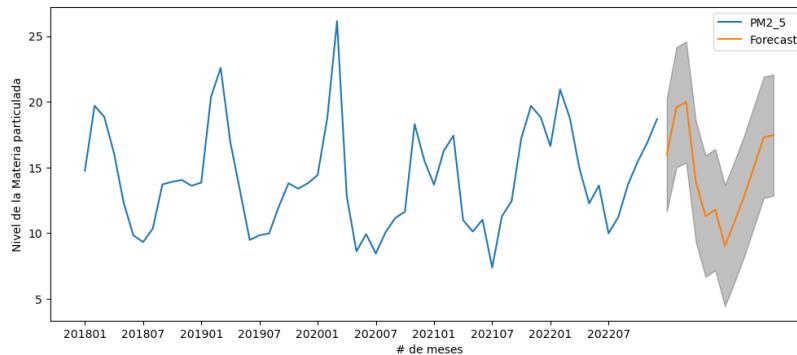


Figura 5-8.: Pronósticos de los modelos $PM_{2,5}$

6. Interpolación espacial de la concentración de material particulado (PM_{10} y $PM_{2,5}$) en las localidades de Bogotá D.C

Este capítulo se centra en determinar la concentración de materia particulada (PM_{10} y $PM_{2,5}$) en las localidades de Bogotá D.C para el periodo 2018 a 2022, a partir de los registros en las estaciones de la Red de Monitoreo de Calidad de Aire - RMCAB. Como se mencionó en capítulos anteriores se cuenta con los registros mensuales de 20 estaciones distribuidas en 12 de las 20 localidades del distrito, por tanto este capítulo busca por medio de métodos de interpolación espacial estimar la concentración del contaminante en diferentes ubicaciones, que permita entender su distribución en el espacio.

Para ello, se realiza un análisis descriptivo con la visualización geográfica para comprender la distribución y el comportamiento de las variables en el espacio, y la aplicación de métodos de interpolación espacial para predecir la concentración del contaminante en localidades sin mediciones.

6.1. Análisis descriptivo de los datos

Como se mencionó en el capítulo 1, se cuenta con una base de datos con las variables de: fecha, concentración de PM_{10} y $PM_{2,5}$, barrio o estación, latitud, longitud y localidad. Para el desarrollo de este capítulo se aplica la transformación de logaritmo a la concentración de las materias particuladas PM_{10} y $PM_{2,5}$ para su mejor procesamiento y análisis, y se agrega la variable de código de la localidad, con el fin de cruzar la base de datos con la capa geográfica de las localidades de Bogotá D.C.

Tabla 6-1.: Resumen concentración de PM_{10} y $PM_{2,5}$ en Bogotá D.C (2018-2022)

Index	Latitud	Longitud	PM_{10}	$PM_{2,5}$	$\log PM_{10}$	$\log PM_{2,5}$
Min.	4,532	-74,17	16,16	8,405	2,783	2,129
1st Qu.	4,602	-74,14	23,93	12,891	3,174	2,557
Median	4,645	-74,11	33,52	14,938	3,512	2,704
Mean	4,653	-74,10	32,23	15,609	3,428	2,711
3rd Qu.	4,701	-74,08	35,81	17,324	3,578	2,852
Max.	4,784	-74,03	57,21	29,011	4,047	3,368

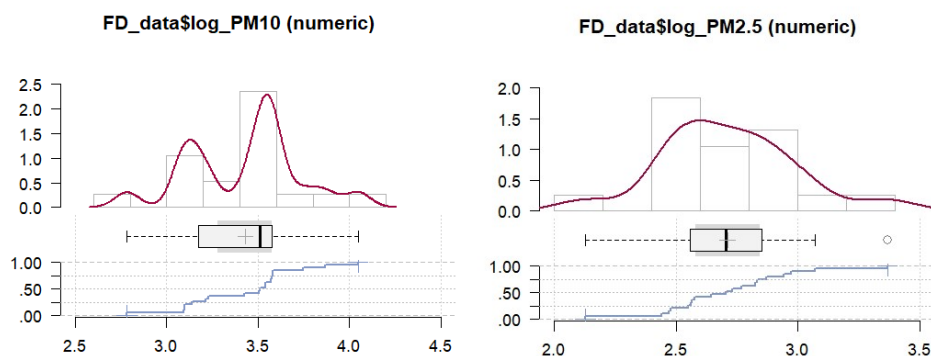


Figura 6-1.: Descripción del Logaritmo de la Materia Particulada PM_{10} y $PM_{2,5}$

La Figura 6-1 muestra dos conjuntos de datos relacionados con el logaritmo de la materia particulada (PM). En el lado izquierdo, se representa la partícula PM_{10} , y el lado derecho la partícula $PM_{2,5}$. Dichas gráficas presentan una combinación de un diagrama de caja y un gráfico de densidad. Para la partícula PM_{10} se evidencia una escala logarítmica de concentraciones entre 2,5 a 4,5 aproximadamente, y de 2,0 a 3,5 para la partícula $PM_{2,5}$. Por otro lado, el eje y ubicado en el lado izquierdo de cada gráfica muestra la densidad que alcanza su punto máximo alrededor del centro del rango de datos para ambas partículas.

Así mismo, los gráficos de densidad representados por las líneas rojas, muestran la distribución de los datos con picos que indican los valores más comunes, en donde la gráfica de partícula PM_{10} muestra picos más pronunciados que la partícula $PM_{2,5}$. Adicionalmente, en el gráfico se encuentra la función de distribución acumulativa (CDF) representada por una línea azul escalonada que aumenta de izquierda a derecha, mostrando el porcentaje acumulado de puntos de datos que caen por debajo de cierto valor. En consecuencia, se evidencia frente a la partícula PM_{10} una mayor concen-

tración en la escala logarítmica de 3,0 a 3,6 aproximadamente, lo que indica que en su mayoría, los registros de la concentración de este contaminante se encuentran dentro de los niveles máximos permisibles a nivel nacional (Tabla **2-2**). Así mismo, en este gráfico se observa la presencia de multimodalidad, muestra dos picos (modas), lo que indica que hay al menos dos grupos distintos de concentraciones, lo que puede sugerir la influencia de diferentes fuentes de emisión, como tráfico y actividades industriales.

Con respecto a la partícula $PM_{2,5}$ se evidencia una distribución más uniforme en comparación con PM_{10} . Así mismo se evidencia una mayor concentración en la escala logarítmica de 2,4 a 3,0 aproximadamente, lo que indica que hay algunos registros de la concentración de este contaminante que se encuentran dentro de los niveles máximos permisibles a nivel nacional (Tabla **2-2**), sin embargo, se encuentra una pequeña proporción que excede estos niveles.

Por otro lado, los diagramas de caja se encuentran en la mitad inferior de cada gráfica, evidenciando para la partícula PM_{10} una mediana (línea dentro de la caja) con tendencia hacia la derecha o valores más elevados. Con respecto a la partícula $PM_{2,5}$ se evidencia la presencia de un valor atípico, representado por un círculo pequeño más allá de los bigotes.

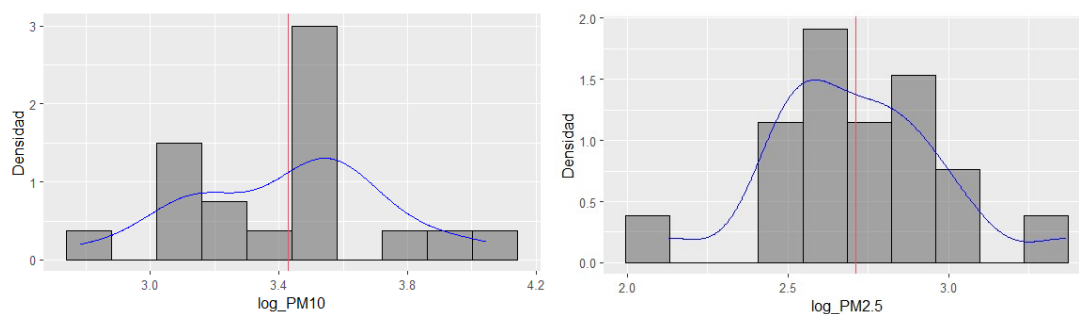


Figura 6-2.: Histograma del Logaritmo de la Materia Particulada PM_{10} y $PM_{2,5}$

De igual forma, la Figura **6-2** representa los dos conjuntos de datos relacionados con las materias particuladas PM_{10} y $PM_{2,5}$, en donde también se observan gráficos de densidad representados por las líneas azules, los cuales evidencian picos más pronunciados en la partícula PM_{10} . En términos generales, y como se evidenció en las figuras anteriores, se puede observar que los valores de $PM_{2,5}$ son relativamente más bajos en comparación a la partícula PM_{10} , sin embargo, la partícula $PM_{2,5}$ presenta mayores registros que exceden los niveles máximos permisibles; así mismo se observa la presencia de multimodalidad en la partícula PM_{10} , lo que podría indicar la agrupación de

diferentes concentraciones por la influencia de algunas fuentes de emisión o ubicaciones.

6.2. Análisis geoestadístico de los datos

Para el análisis descriptivo de la base de datos, se establecen las coordenadas para crear un objeto espacial. Como se mencionó anteriormente se cuenta con registros de la concentración de esta partícula en 12 localidades de la ciudad de Bogotá.

Con los valores promedios de la concentración de los contaminantes en las estaciones mencionadas durante los años 2018 a 2022, se construyen gráficos descriptivos, en ellos se puede evidenciar la existencia de una aleatoriedad espacial, ya que los puntos de todos los colores se encuentran dispersos en todo el mapa.

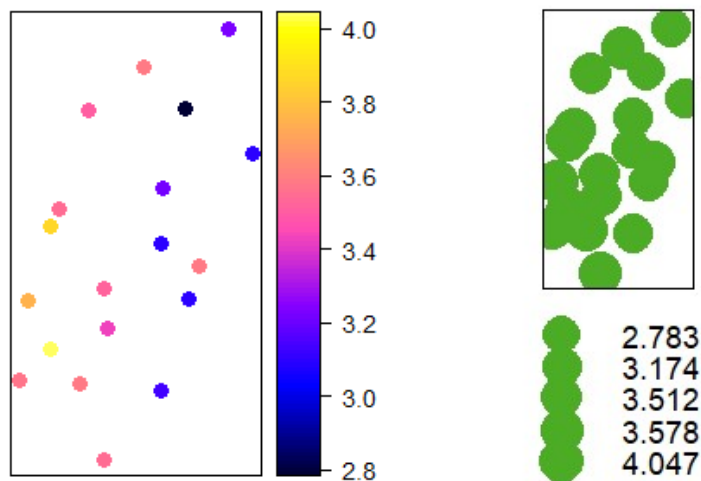


Figura 6-3.: Gráficos del Logaritmo de la Materia Particulada PM_{10}

Con respecto a la materia particulada PM_{10} se observan dos gráficos de dispersión, con una escala que va de 2.8 a 4.0 representada en una escala de colores que va desde puntos de color negro hasta puntos de color amarillo. Así mismo, se observa que la concentración del logaritmo de esta partícula presenta mayor cantidad de puntos de color rosa (entre 3.4 y 3.6 aproximadamente), los cuales se encuentran dispersos en todo el mapa.

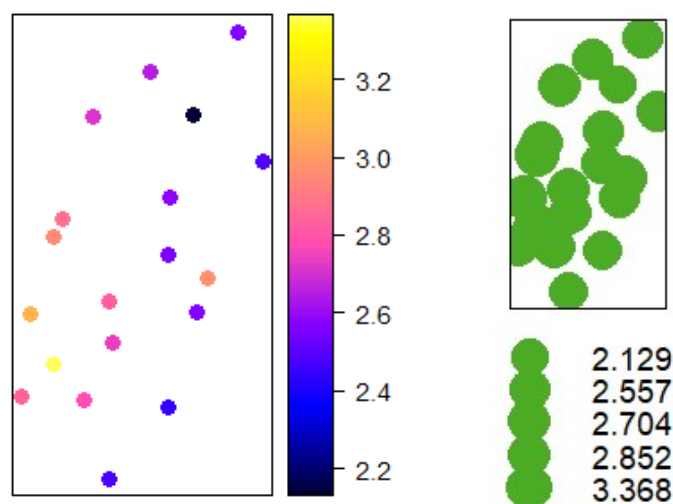


Figura 6-4.: Gráficos del Logaritmo de la Materia Particulada $PM_{2,5}$

De igual forma, para la materia particulada $PM_{2,5}$ se observan dos gráficos de dispersión, con una escala menor en comparación a la partícula PM_{10} , que va de 2.2 a 3.2 representada igualmente en una escala de colores que va desde puntos de color negro hasta puntos de color amarillo. Así mismo, se observa que la concentración del logaritmo de esta partícula presenta mayor cantidad de puntos de color rosa y violeta (entre 2.6 y 2.9 aproximadamente), los cuales se encuentran dispersos en todo el mapa. En general se observa un compartamiento similar para las partículas PM_{10} y $PM_{2,5}$, en donde no hay una agrupación clara y los valores se distribuyen aleatoriamente, lo que podría sugerir la no presencia de dependencia espacial.

6.2.1. Análisis de la tendencia

Por otro lado para confirmar si existe algún tipo de dependencia espacial, se construyen gráficos de dispersión contra las coordenadas (longitud y latitud), encontrando que en ambas partículas el gráfico frente a la longitud evidencia una tendencia decreciente que indica que aparentemente los datos no son independientes y que la correlación espacial está presente en ellos o la existencia de alguna tendencia que hace falta remover.

Frente a la latitud esta tendencia no es tan evidente, sin embargo, se procede a utilizar un modelo lineal de primer orden contra las direcciones con el fin de remover dicha tendencia.

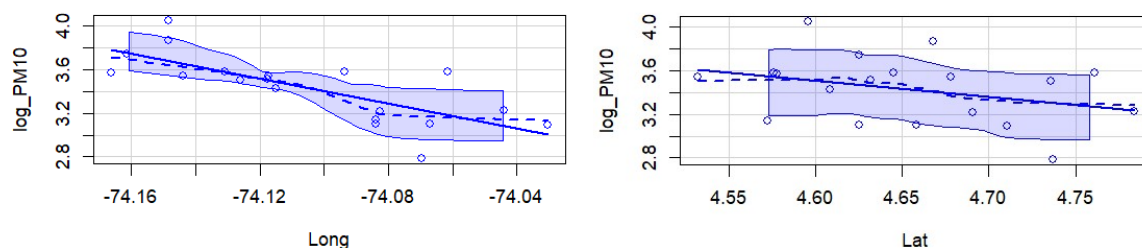


Figura 6-5.: Gráficos de dispersión contra la longitud y latitud del Logaritmo de PM_{10}

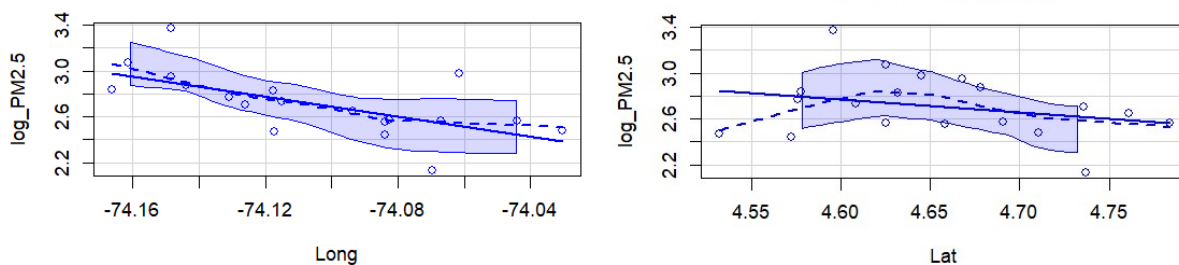


Figura 6-6.: Gráficos de dispersión contra la longitud y latitud del Logaritmo de $PM_{2,5}$

6.2.2. Ajuste de modelo para remover la tendencia

Teniendo en cuenta el análisis de la tendencia realizado, se encuentra necesario remover dicha tendencia antes de aplicar métodos de interpolación espacial, puesto que muchos de estos métodos, como el Kriging ordinario y el Inverso de la Distancia - IDW (Inverse Distance Weighting), asumen que los datos no tienen una estructura de tendencia global, sino que presentan una variabilidad local y aleatoria alrededor de un valor promedio, es decir variables estacionarias, con media y variabilidad que no cambian significativamente en el espacio; si llegara a existir una tendencia, esta violaría el supuesto de estacionariedad y podría sesgar las estimaciones de los valores interpolados.

Para remover dicha tendencia y teniendo en cuenta que en ambas partículas en sus gráficos de dispersión evidencian una tendencia decreciente frente a la longitud, se decide utilizar una de las técnicas más comunes, consistente en modelar la tendencia

utilizando regresiones que relacionan los valores de la variable de interés con las coordenadas espaciales.

Al ajustar un modelo de regresión lineal de primer orden contra las direcciones, se evidencia que la latitud no resulta significativa, por tanto se procede a estimar un modelo solo con el término de la longitud. Así mismo, con los residuales de este modelo se verifica que tengan media constante sobre la región de estudio, por medio de los gráficos de dispersión contra las coordenadas y con el ajuste de un modelo de primer orden para los residuales.

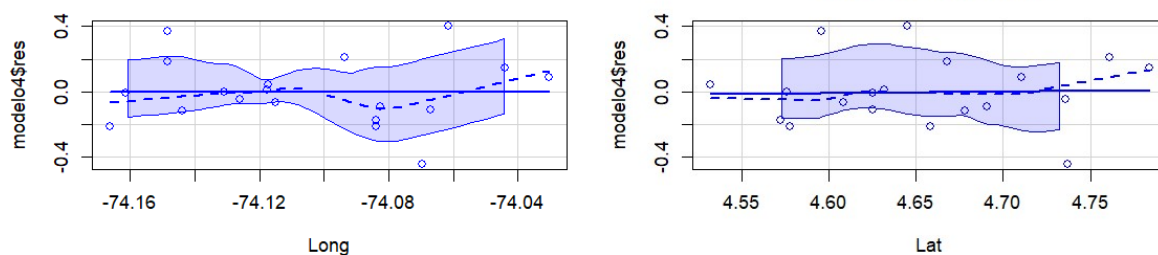


Figura 6-7.: Gráfico de dispersión contra la longitud y latitud de los residuales del modelo para el logaritmo de PM_{10}

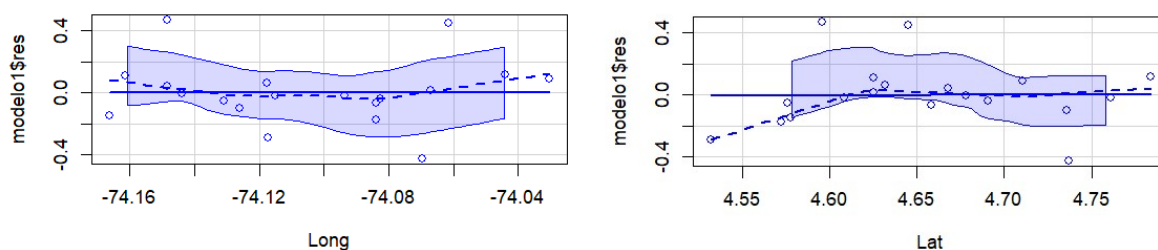


Figura 6-8.: Gráficos de dispersión contra la longitud y latitud de los residuales del modelo para el logaritmo de $PM_{2,5}$

Al observar el comportamiento de los gráficos de dispersión contra las coordenadas, se evidencia que ninguno reporta indicios de tendencia, así mismo los términos del modelo de primer orden para los residuales no resultan significativos, por tanto se concluye que dichos residuales tienen una media constante sobre la región de estudio, así como no existe evidencia estadísticamente significativa de dependencia espacial en los residuales.

Tabla 6-2.: Anova modelo para los residuales del modelo de primer orden con la longitud del logaritmo de PM_{10}

Dirección	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Long	1	0,00000	0,000000	0,0000	1,0000
Lat	1	0,00090	0,000901	0,019	0,8921

Tabla 6-3.: Anova modelo para los residuales del modelo de primer orden con la longitud del logaritmo de $PM_{2,5}$

Dirección	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Long	1	0,00000	0,000000	0,0000	1,0000
Lat	1	0,00024	0,000244	0,0048	0,9454

En consecuencia, y como se observa después de filtrar la tendencia mediante modelos de primer orden con la longitud, se evidencia que los residuales de estos modelos no presentan dependencia o correlación espacial, en este caso se puede inferir que la aparente dependencia espacial existente en los datos originales es en realidad resultado de la tendencia presente.

Así mismo, al estimar el semivariograma, el cual describe la relación entre la distancia espacial y la similitud de los valores, se evidencia que el modelo teórico que se ajusta es lineal, es decir un semivariograma constante, indicando que la semivarianza no cambia con la distancia entre las ubicaciones espaciales, lo que sugiere que no existe una correlación espacial entre los valores medidos, siendo estos aleatorios en el espacio, sin alguna estructura espacial discernible.

Teniendo en cuenta lo anterior, así como la presencia de poco registros dentro de cada localidad, se determina que para realizar la interpolación espacial de la concentración de estas partículas, no es necesario utilizar métodos espaciales sofisticados como el método de Kriging, para este caso se utilizarán los métodos de predicción espacial de inverso de la distancia y modelos lineales (para este método se requiere tener covariables que estén disponibles para cualquier punto dentro de la región de observación, por tanto, las coordenadas (longitud y latitud) se utilizarán para suplir esta necesidad).

6.3. Interpolación espacial

6.3.1. Construcción de la grilla a interpolar

Teniendo en cuenta las condiciones atípicas de la localidad de Sumapaz (considerada como la localidad 20 de carácter rural y situada en el extremo sur de Bogotá, geográficamente dividida por las cuencas del Río Blanco y del Río Sumapaz, destacada por su belleza natural y reconocida por su productividad agrícola y lechera), se considera pertinente suprimir dicha localidad y realizar la interpolación espacial con las 19 localidades restantes.

Para estimar el comportamiento de estos contaminantes sobre toda la región de estudio, se procede a generar una grilla con 100 puntos aleatorios distribuidos en cada polígono (localidades) para un total de 1.900 puntos a interpolar distribuidos en las 19 localidades, posteriormente se generan las estimaciones.



Figura 6-9.: Visualización grilla a interpolar localidades de Bogotá D.C

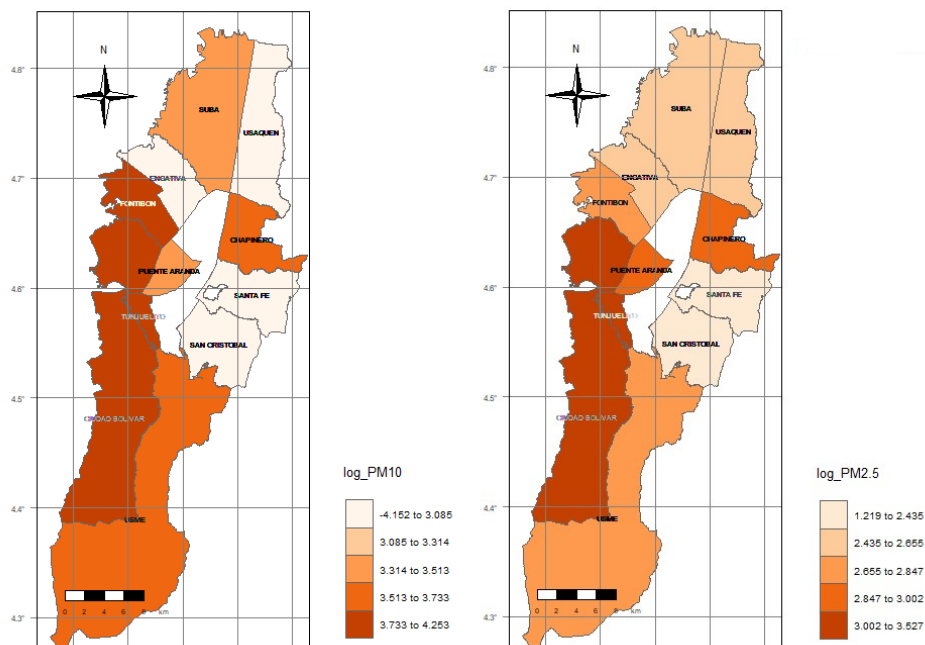


Figura 6-10.: Concentración del logaritmo de PM_{10} y $PM_{2,5}$ año 2022 en Bogotá D.C

Visualización de la capa geográfica con los datos originales

La Figura 6-10 muestra la capa geográfica del logaritmo de los registros (datos originales) correspondientes al año 2022 para los dos contaminantes en 12 localidades de Bogotá. Para la partícula PM_{10} se cuenta con cinco intervalos que van desde -4,152 a 4,253, en donde se ilustra a Ciudad Bolívar, Tunjuelito, Kennedy y Fontibón como las localidades con mayor concentración del contaminante; mientras que, se ubica en el menor intervalo las localidades de San Cristobal, Santa Fé, Usaquen y Engativá, con menores niveles de concentración del contaminante.

Un comportamiento similar se evidencia para la partícula $PM_{2,5}$; de igual forma se cuenta con cinco intervalos que van desde 1,219 a 3,527 (por debajo de la concentración de PM_{10}). Así como ocurre en la partícula anterior, se ilustra a Ciudad Bolívar, Tunjuelito y Kennedy como las localidades con mayor concentración del contaminante. De igual forma ocurre con las localidades de San Cristobal y Santa Fé ubicadas en el menor intervalo con los menores niveles de concentración del contaminante.

Para este proceso de interpolación, se tomó el último año correspondiente al 2022, en donde se realizaron cuatro cortes en los meses de enero, abril, julio y octubre, con el fin de interpolar la concentración de la partícula y observar cambios en estos cuatro

periodos del tiempo, para ello, a continuación se visualiza la capa geográfica con el logaritmo de los registros (datos originales) correspondientes al año 2022, en donde se realizaron estos cuatro cortes.

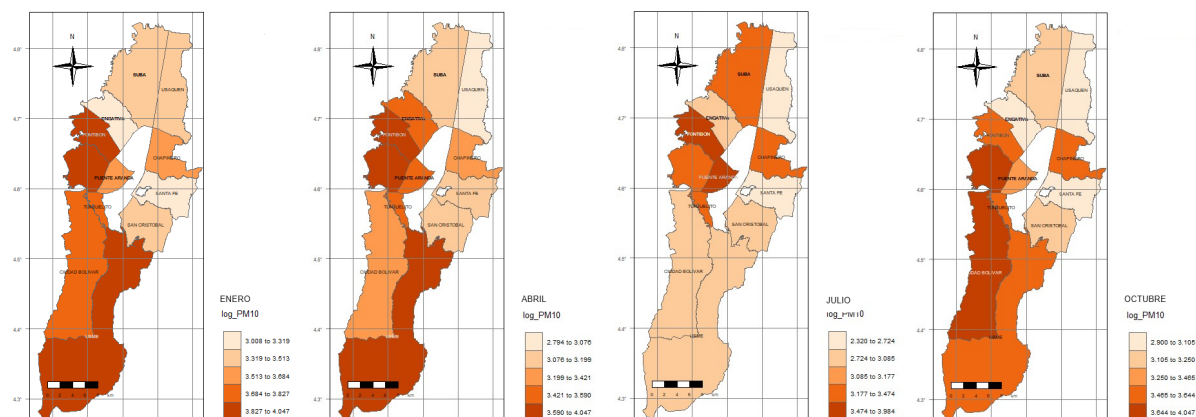


Figura 6-11.: Concentración del logaritmo de PM_{10} cortes año 2022 en Bogotá D.C

La Figura **6-11** ilustra la capa geográfica del logaritmo de la concentración de la partícula PM_{10} en cuatro cortes del año 2022 en 12 localidades de Bogotá. En los cuatro casos se cuenta con cinco intervalos, en los que se evidencia el mes de julio con los niveles más bajos, puesto que gran parte de las localidades en este mes se ubican en intervalos de 2,320 a 3,085. Así mismo, entre las localidades con mayor concentración del contaminante se mantienen Usme, Fontibón, Kennedy, Puente Aranda y Ciudad Bolívar. Contrario a ello, entre las localidades con menores niveles del contaminante se encuentran Engativá, Santa Fé y Usaquen.

La Figura **6-12** ilustra la capa geográfica del logaritmo de la concentración de la partícula $PM_{2,5}$ en cuatro cortes del año 2022 en 12 localidades de Bogotá. En los cuatro casos se cuenta con cinco intervalos, en los que se observa un comportamiento similar a la partícula anterior. Se evidencia el mes de julio con los niveles más bajos en intervalos de 1,630 a 2,913. Así mismo, entre las localidades con mayor concentración del contaminante se mantienen Kennedy, Fontibón, Tunjuelito, Ciudad Bolívar y Chapinero. Contrario a ello, entre las localidades con menores niveles del contaminante se encuentran Usaquen, Suba, Engativá y San Cristobal.

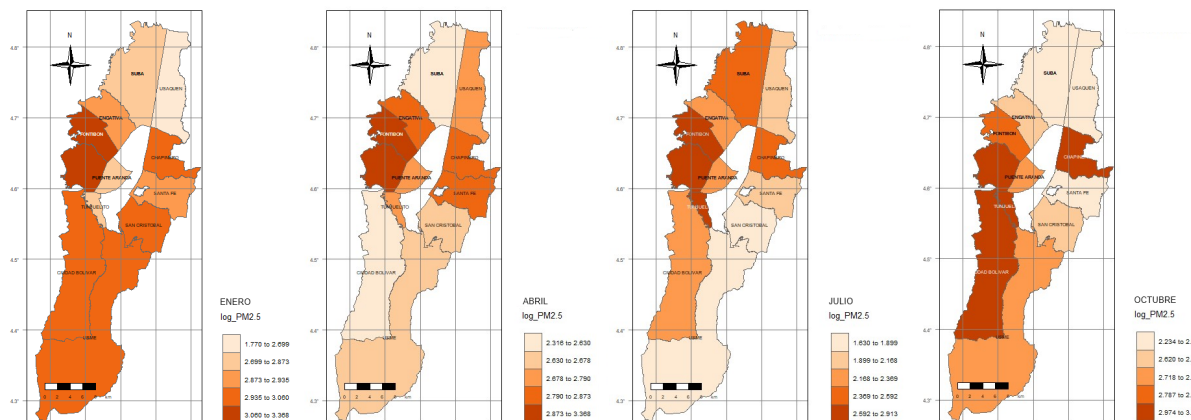


Figura 6-12.: Concentración del logaritmo de $PM_{2,5}$ cortes año 2022 en Bogotá D.C

6.3.2. Interpolación espacial de la Materia Particulada PM_{10} - Cortes 2022

Como se indicó anteriormente, para interpolar el comportamiento de la materia particulada PM_{10} sobre toda la región de estudio, se realiza el método inverso de la distancia, modelo lineal de primer orden y modelo lineal de segundo orden. Para este caso en particular, se tomó el último año correspondiente al 2022, en donde se realizaron cuatro cortes (Enero-Abril-Julio-Octubre), con el fin de interpolar la concentración de la partícula y observar cambios en estos cuatro periodos del tiempo durante el año observado.

Como resultado y para tener una idea de la capacidad de pronóstico de los métodos desarrollados, se realiza un proceso de validación cruzada, obteniendo los siguientes resultados:

Finalmente, por medio de la validación cruzada se logra seleccionar el mejor modelo según el periodo de tiempo de observación, obteniendo los siguientes resultados:

- Enero de 2022 para PM_{10} :** Como resultado del método seleccionado (Modelo lineal de segundo orden), se encuentra un Error Cuadrático Medio de 0,1516617 y un Mean Absolute Percentage Error (MAPE) de 3,392329 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 3,394 % de los valores originales.
- Abril de 2022 para PM_{10} :** Como resultado del método seleccionado (Modelo lineal de segundo orden), se encuentra un Error Cuadrático Medio de 0,2137303

Tabla 6-4.: Resultados validación cruzada métodos inverso de la distancia y modelos lineales de primer y segundo orden PM_{10}

Cortes	Método	ECM	MAPE
Enero	Inverso de la Distancia	0,2394177	6,769 %
	Modelo Lineal de 1er orden	0,1833533	4,612 %
	Modelo Lineal de 2do orden	0,1516617	3,392 %
Abril	Inverso de la Distancia	0,2360855	7,138 %
	Modelo Lineal de 1er orden	0,2327908	5,426 %
	Modelo Lineal de 2do orden	0,2137303	4,775 %
Julio	Inverso de la Distancia	0,06990515	2,279 %
	Modelo Lineal de 1er orden	0,2838712	7,554 %
	Modelo Lineal de 2do orden	0,277129	7,190 %
Octubre	Inverso de la Distancia	0,1555061	4,624 %
	Modelo Lineal de 1er orden	0,209458	5,074 %
	Modelo Lineal de 2do orden	0,1865089	4,046 %

y un Mean Absolute Percentage Error (MAPE) de 4,775064 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 4,77 % de los valores originales.

c) **Julio de 2022 para PM_{10} :** Como resultado del método seleccionado (Inverso de la distancia), se encuentra un Error Cuadrático Medio de 0,06990515 y un Mean Absolute Percentage Error (MAPE) de 2,27986 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 2,27 % de los valores originales.

d) **Octubre de 2022 para PM_{10} :** Como resultado del método seleccionado (Modelo lineal de segundo orden), se encuentra un Error Cuadrático Medio de 0,1865089 y un Mean Absolute Percentage Error (MAPE) de 4,046075 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 4,04 % de los valores originales.

El gráfico de los valores estimados o interpolados para los puntos de la grilla creada tiene el siguiente aspecto (**6-13**).

La Figura **6-13** ilustra gráficos de dispersión con la interpolación para el logaritmo de la concentración de la partícula PM_{10} en cuatro cortes del año 2022 para 1.900 puntos

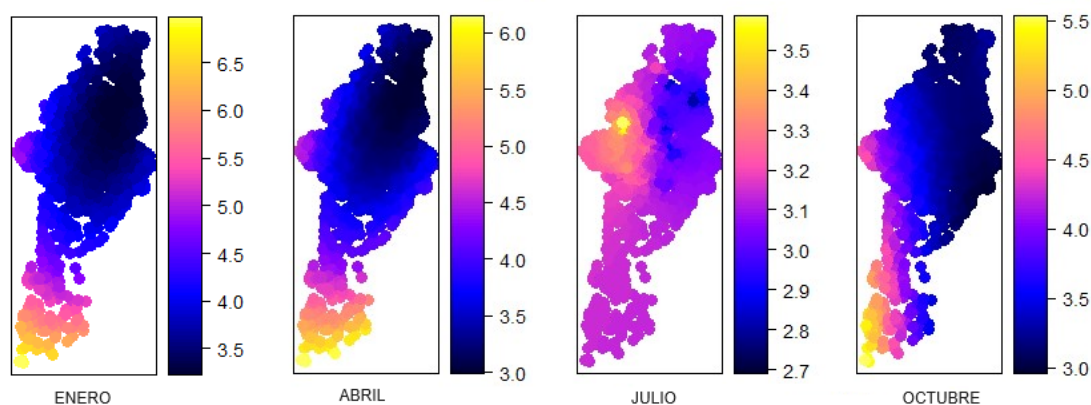


Figura 6-13.: Interpolación para el Logaritmo de la Materia Particulada PM_{10} cortes año 2022

aleatorios en 19 localidades de Bogotá. Dichos gráficos presentan una escala de colores que van desde puntos de color negro hasta puntos de color amarillo. De igual forma que en las figuras anteriores, se observa que el mes de julio presenta las concentraciones mas bajas con una escala entre 2,7 y 3,5, mientras que los meses de enero y abril presentan niveles más elevados llegando a concentraciones de hasta 6,5 y 6,0 respectivamente.

Para los meses de enero, abril y octubre se observa que la concentración de la partícula PM_{10} es menor para puntos en la parte nororiente o superior derecha y aumenta a medida que se desplaza a la parte sur de la ciudad. Por otro lado, para el mes de julio se observa mayor cantidad de puntos de color violeta y rosa (entre 3,0 y 3,4).

A continuación se visualiza la capa geográfica con los valores interpolados en los cuatro cortes del año 2022, por las 19 localidades de Bogotá D.C:

La Figura 6-14 ilustra la capa geográfica con la interpolación para el logaritmo de la concentración de la partícula PM_{10} en cuatro cortes del año 2022 en 19 localidades de Bogotá. En los cuatro cortes se cuenta con cinco intervalos, en donde se evidencia nuevamente el mes de julio con los niveles más bajos con intervalos entre 2,689 y 3,587. Así mismo, se observa que la concentración del contaminante es menor en localidades ubicadas en el nororiente de la ciudad y aumenta en localidades ubicadas en el sur de la ciudad. En particular, se observa que las localidades de Ciudad Bolívar, Usme, Bosa, Kennedy y Fontibón presentan mayor concentración del contaminante, mientras que las localidades de Teusaquillo, Barrios Unidos, Chapinero, Usaquen y Santa Fé presentan menores niveles del contaminante.

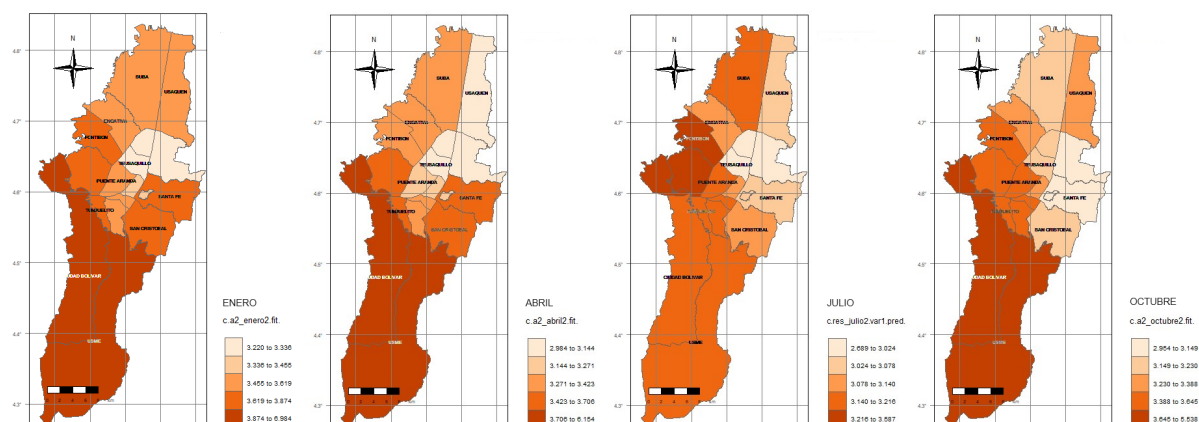


Figura 6-14.: Interpolación para el Logaritmo de la Materia Particulada PM_{10} cortes año 2022 por localidades

Así mismo, se presenta la Figura 6-15 que ilustra la interpolación de la concentración de PM_{10} en los cuatro cortes del año 2022 en las 19 localidades de Bogotá.

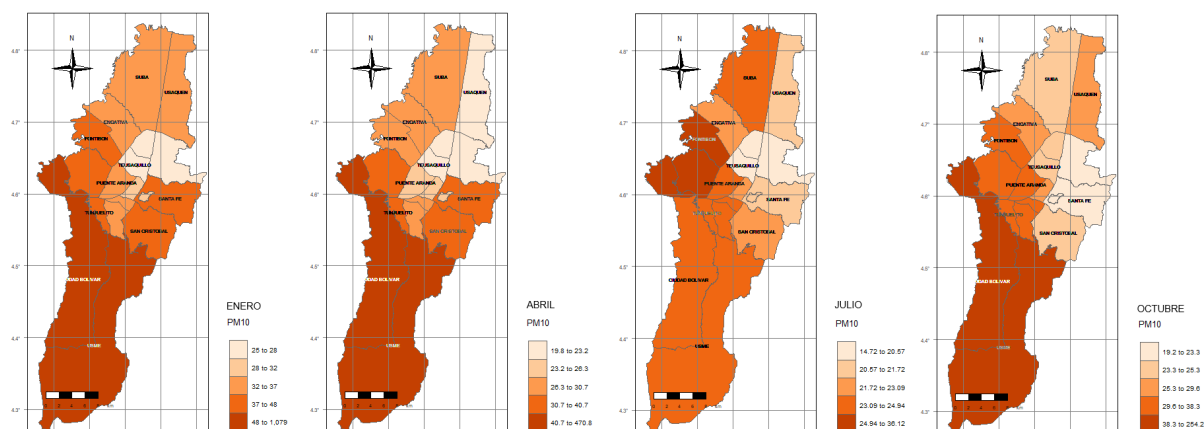


Figura 6-15.: Interpolación para la Materia Particulada PM_{10} cortes año 2022 por localidades

6.3.3. Interpolación espacial de la Materia Particulada $PM_{2,5}$ - Cortes 2022

Teniendo en cuenta lo indicado, para interpolar el comportamiento de la materia particulada $PM_{2,5}$ sobre toda la región de estudio, se realiza el método inverso de la

distancia, modelo lineal de primer orden y modelo lineal de segundo orden. De igual forma que con la partícula PM_{10} , se tomó el último año, correspondiente al 2022, en donde se realizaron cuatro cortes (Enero-Abril-Julio-October), con el fin de interpolar la concentración de la partícula y observar cambios en estos cuatro periodos del tiempo durante el año observado.

Como resultado y para tener una idea de la capacidad de pronóstico de los métodos desarrollados, se realiza un proceso de validación cruzada, obteniendo los siguientes resultados:

Tabla 6-5.: Resultados validación cruzada métodos inverso de la distancia y modelos lineales de primer y segundo orden $PM_{2,5}$

Cortes	Método	ECM	MAPE
Enero	Inverso de la Distancia	0,08451145	3,086 %
	Modelo Lineal de 1er orden	0,3063378	9,742 %
	Modelo Lineal de 2do orden	0,302486	9,336 %
Abril	Inverso de la Distancia	0,129015	4,703 %
	Modelo Lineal de 1er orden	0,2033405	5,376 %
	Modelo Lineal de 2do orden	0,1862926	4,174 %
Julio	Inverso de la Distancia	0,691423	31,401 %
	Modelo Lineal de 1er orden	0,2658385	9,302 %
	Modelo Lineal de 2do orden	0,2399491	8,646 %
October	Inverso de la Distancia	0,1162833	4,247 %
	Modelo Lineal de 1er orden	0,1838975	5,016 %
	Modelo Lineal de 2do orden	0,1732851	4,919 %

Finalmente, por medio de la validación cruzada se logra seleccionar el mejor modelo según el periodo de tiempo de observación, obteniendo los siguientes resultados:

- Enero de 2022 para $PM_{2,5}$:** Como resultado del método seleccionado (Inverso de la distancia), se encuentra un Error Cuadrático Medio de 0,08451145 y un Mean Absolute Percentage Error (MAPE) de 3,086205 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 3,086 % de los valores originales.
- Abril de 2022 para $PM_{2,5}$:** Como resultado del método seleccionado (Modelo lineal de segundo orden), se encuentra un Error Cuadrático Medio de 0,1862926

y un Mean Absolute Percentage Error (MAPE) de 4,174154 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 4,174 % de los valores originales.

- c) **Julio de 2022 para $PM_{2,5}$:** Como resultado del método seleccionado (Modelo lineal de segundo orden), se encuentra un Error Cuadrático Medio de 0,2399491 y un Mean Absolute Percentage Error (MAPE) de 8,646665 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 8,646 % de los valores originales.
- d) **Octubre de 2022 para $PM_{2,5}$:** Como resultado del método seleccionado (Inverso de la distancia), se encuentra un Error Cuadrático Medio de 0,1162833 y un Mean Absolute Percentage Error (MAPE) de 4,247618 lo que indica que es de esperar que las estimaciones dadas por este modelo se alejen un 4,247 % de los valores originales.

El gráfico de los valores estimados o interpolados para los puntos de la grilla creada tiene el siguiente aspecto:

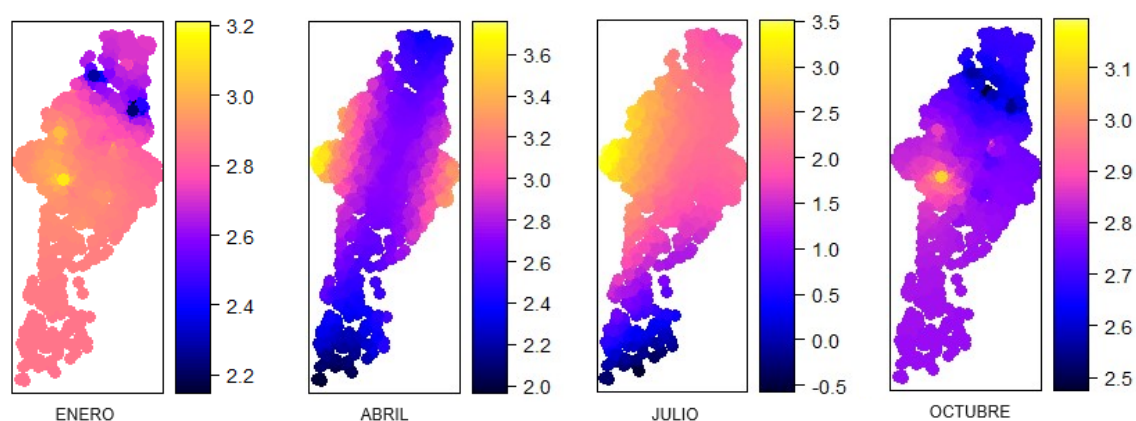


Figura 6-16.: Interpolación para el Logaritmo de la Materia Particulada $PM_{2,5}$ cortes año 2022

La Figura 6-16 ilustra gráficos de dispersión con la interpolación para el logaritmo de la concentración de la partícula $PM_{2,5}$ en cuatro cortes del año 2022 para 1.900 puntos aleatorios en 19 localidades de Bogotá. Dichos gráficos presentan una escala de colores que van desde puntos de color negro hasta puntos de color amarillo. Para esta

partícula, no es tan notorio algún mes o corte con menor nivel de concentración del contaminante, en comparación con la partícula PM_{10} , sin embargo, se puede observar que los meses de julio y enero presentan las concentraciones más bajas con una escala entre $[-0,5 - 3,5]$ y $[2,2 - 3,2]$ respectivamente, mientras que los meses de abril y octubre presentan niveles más elevados llegando a concentraciones de hasta 3,6 aproximadamente.

Para el mes de abril se observa que la concentración de la partícula $PM_{2,5}$ es menor para puntos en la parte sur y norte, y aumenta a medida que se desplaza a la parte oriente u occidente de la ciudad. Por otro lado, para los meses de enero y octubre se observa que la concentración es menor para puntos en la parte norte, y aumenta a medida que se desplaza a la parte sur de la ciudad. Finalmente, para el mes de julio se observa que la concentración es menor para puntos en la parte sur, y aumenta a medida que se desplaza a la parte noroccidental de la ciudad.

A continuación se visualiza la capa geográfica con los valores interpolados en los cuatro cortes del año 2022, en las 19 localidades de Bogotá D.C:

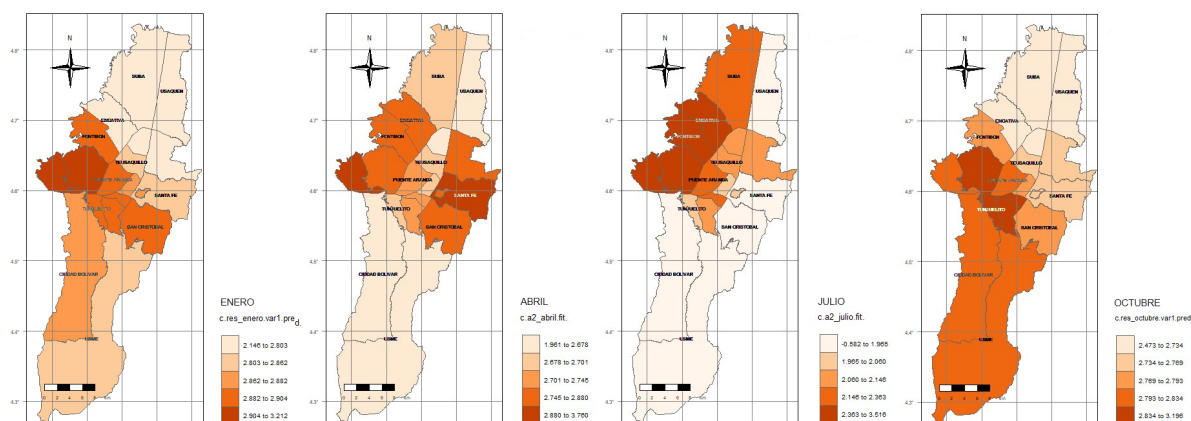


Figura 6-17.: Interpolación para el Logaritmo de la Materia Particulada $PM_{2,5}$ cortes año 2022 por localidades

La Figura 6-17 ilustra la capa geográfica con la interpolación para el logaritmo de la concentración de la partícula $PM_{2,5}$ en cuatro cortes del año 2022 en 19 localidades de Bogotá. En los cuatro cortes se cuenta con cinco intervalos, en donde se observa cierta dispersión de la concentración del contaminante en las localidades durante los cuatro cortes. En general, se observa que las localidades de Kennedy, Bosa, Santa Fé, Fontibón, Engativá, Tunjuelito, Rafael Uribe y Antonio Nariño presentan mayor concentración del contaminante, mientras que las localidades de Engativá, Suba, Usaquen,

Barrios Unidos, Chapinero, Ciudad Bolívar, Usme, San Cristobal y Santa Fé presentan menores niveles del contaminante.

Así mismo, se presenta la Figura 6-18 que ilustra la interpolación de la concentración de $PM_{2,5}$ en los cuatro cortes del año 2022 en las 19 localidades de Bogotá.

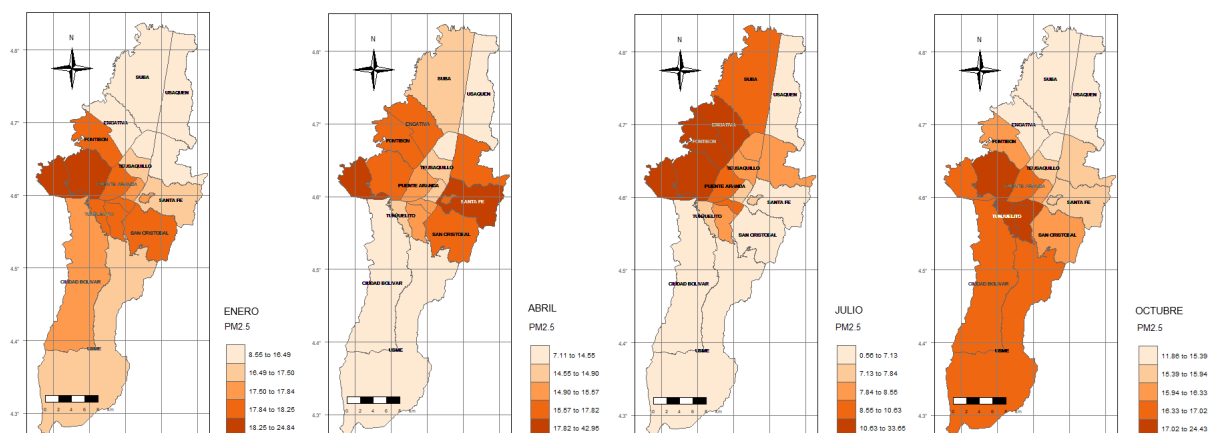


Figura 6-18.: Interpolación para la Materia Particulada $PM_{2,5}$ cortes año 2022 por localidades

7. Estimación de modelos espaciales y espacio temporales

Este capítulo tiene como objetivo estimar modelos estadísticos avanzados para comprender la relación entre la tasa de mortalidad y algunas variables demográficas, ambientales y sociales. Se emplean modelos espaciales como SAR (Spatial Autoregressive Model), SEM (Spatial Error Model) y modelos espacio-temporales (Spatial-Temporal Fixed Differences Model). La tasa de mortalidad se considera como la variable dependiente, mientras que las variables independientes incluyen el comportamiento de las partículas de materia PM_{10} y $PM_{2,5}$, el tipo de régimen y la localidad.

Para ello, se realiza visualización de la capa geográfica de la tasa de mortalidad en las localidades de Bogotá, para finalizar con el desarrollo de los modelos mencionados anteriormente y su análisis para comprender la relación entre la tasa de mortalidad, la concentración por material particulado y algunas variables demográficas y sociales.

7.1. Visualización de la Capa Geográfica de la Tasa de Mortalidad

La Figura 7-1 ilustra la capa geográfica con la tasa de mortalidad en las 20 localidades de Bogotá, durante el año 2022. Se cuenta con cinco intervalos entre 0,006 y 0,269, en donde se observa que la tasa de mortalidad es mayor en localidades ubicadas en el norte y sur de la ciudad, y disminuye en localidades ubicadas en el centro de la ciudad. En particular, se observa que las localidades de Sumapaz, Puente Aranda, La Candelaria y Santa Fé presentan mayores tasas de mortalidad, mientras que las localidades de Usme, Rafael Uribe y Teusaquillo presentan menores tasas.

Así mismo, y teniendo en cuenta los cuatro períodos estudiados (Enero, Abril, Julio, Octubre) en el capítulo anterior, se elaboraron visualizaciones geográficas de la tasa de mortalidad por enfermedades respiratorias agudas en las localidades de Bogotá, D.C.,

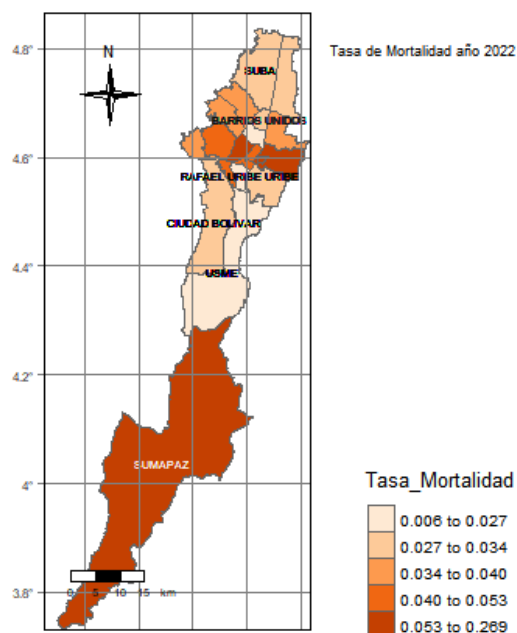


Figura 7-1.: Tasa de Mortalidad en las localidades de Bogotá D.C año 2022

durante estos períodos:

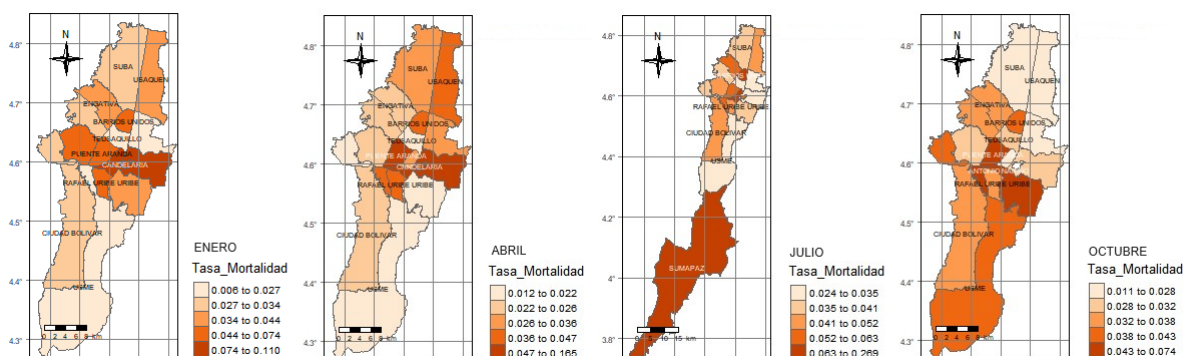


Figura 7-2.: Tasa de Mortalidades en las localidades de Bogotá D.C cortes año 2022

La Figura 7-2 ilustra la capa geográfica con la tasa de mortalidad en las 20 localidades de Bogotá en cuatro cortes del año 2022. En estos periodos se cuenta con cinco intervalos, en donde se evidencian los meses de enero y octubre con las tasas de mortalidad más bajas con intervalos entre 0,006 y 0,165. Así mismo, para los meses de enero, abril y julio se observa que la tasa mortalidad es menor en localidades ubicadas en el sur de la ciudad y aumenta en localidades ubicadas en el norte de la ciudad. Por el con-

trario, en el mes de octubre se observa que la tasa mortalidad es menor en localidades ubicadas en el norte del departamento y aumenta en localidades ubicadas en el sur de la ciudad. En general, se observa que las localidades de Santa Fé, La Candelaria, Los Martires, Antonio Nariño, Puente Aranda, Sumapaz, Barrios Unidos, San Cristobal y Tunjuelito presentan mayores tasas de mortalidad, mientras que las localidades de Usme, Rafael Uribe, Teusaquillo, Suba, Usaquen y Chapinero presentan menores tasas.

7.2. Modelos Espaciales SAR y SEM

En el marco de este estudio, se utilizó la Tabla 4-5 la cual contiene los registros de la contaminación de material particulado y la Tabla 4-6 con los registros de la tasa de mortalidad. El objetivo principal es desarrollar modelos espaciales y espacio-temporales que permitan comprender la relación entre la tasa de mortalidad y diversos factores como la localidad, el régimen y las materias particuladas, entre otros.

Para lograr este propósito, se llevó a cabo una exhaustiva preparación de los datos, que incluyó la elaboración de una matriz de ponderaciones espaciales. Esta matriz, también conocida como matriz de proximidad, juega un papel crucial en la modelación estadística al definir el peso o la influencia relativa de cada localidad respecto a las otras en función de su cercanía o conexión espacial. Por ejemplo, en un modelo de Autoregresión Espacial (SAR), la matriz de proximidad podría definir que localidades geográficamente cercanas tengan una mayor influencia en las predicciones de las localidades vecinas, reflejando la idea de que áreas próximas tienden a presentar características similares. Esta matriz es fundamental en la construcción de los modelos de regresión lineal, SAR (Autoregresión Espacial) y SEM (Modelos de Error Espacial), que proporcionan perspectivas significativas sobre la naturaleza de la relación entre los diferentes factores mencionados.

Resultados de los modelos SAR y SEM

Para comprender de manera integral la relación entre la tasa de mortalidad y los factores ambientales y sociales, se evaluó inicialmente la relación entre estas variables en diferentes puntos del tiempo, considerando la correlación espacial entre las localidades. Esta perspectiva implica examinar los datos de manera conjunta en dos dimensiones: el espacio y el tiempo. En el análisis espacial, se consideran las variaciones geográficas en la tasa de mortalidad y su relación con los factores ambientales y sociales en diferentes localidades. Por otro lado, el análisis temporal permite evaluar la evolución de la tasa

de mortalidad a lo largo del tiempo y su asociación con los cambios en los factores ambientales y sociales.

Para ello, y como se abordó en el capítulo anterior, se seleccionaron cuatro meses específicos (enero, abril, julio y octubre) como puntos de referencia a lo largo de todos los períodos de estudio. La elección de estos meses responden a la necesidad de considerar la variabilidad estacional en la tasa de mortalidad y en los factores ambientales, como la concentración de partículas PM_{10} y $PM_{2,5}$, que pueden presentar patrones estacionales distintivos.

Tabla 7-1.: Pruebas de significancia modelo SAR

Mes	Prueba	Coefficientes	Valor-p
Enero	Rho	-0,59563	
	AIC	-277,07	
	LR(Verosimilitud)	5,8701	0,015401
Abril	Rho	-0,10171	
	AIC	-352,94	
	LR(Verosimilitud)	6,192	0,012833
Julio	Rho	-0,34254	
	AIC	-306,62	
	LR(Verosimilitud)	5,675	0,017208
Octubre	Rho	-0,66991	
	AIC	-331,94	
	LR(Verosimilitud)	6,2204	0,012629

AIC (Criterio de Información Akaike): El modelo con el AIC más bajo es generalmente considerado como el mejor. En este caso, el AIC del modelo SAR en términos generales es ligeramente menor que el del modelo SEM. Si bien la diferencia es pequeña, indica una leve preferencia por el modelo SAR.

LR test value (Valor del test de Razón de Verosimilitud): Este valor se utiliza para evaluar la significancia de la autocorrelación espacial. Un valor p bajo indica que la autocorrelación espacial es significativa. En ambos modelos, el valor p del LR test es bajo, lo que indica que la autocorrelación espacial es importante en ambos casos. Sin embargo, un valor ligeramente más bajo en el modelo SAR sugiere que este modelo captura la autocorrelación espacial de manera un poco más efectiva.

Tabla 7-2.: Pruebas de significancia modelo SEM

Mes	Prueba	Coefficientes	Valor-p
Enero	Lamba	-0,90622	
	AIC	-276,89	
	LR(Verosimilitud)	5,6837	0,017123
Abril	Lamba	-0,028	
	AIC	-353,29	
	LR(Verosimilitud)	6,5347	0,013579
Julio	Lamba	-0,416	
	AIC	-306,46	
	LR(Verosimilitud)	5,5211	0,018789
Octubre	Lamba	-0,60071	
	AIC	-331,65	
	LR(Verosimilitud)	5,9333	0,014858

Rho para el modelo SAR: Este valor representa el coeficiente de autocorrelación espacial en el modelo SAR. Un valor de Rho cercano a -1 indica una fuerte autocorrelación espacial inversa o negativa, lo que significa que las observaciones cercanas tienen valores opuestos entre sí. Esto sugiere una estructura espacial en la que las diferencias en los valores de las observaciones tienden a disminuir a medida que las ubicaciones se vuelven más cercanas entre sí. Por tanto, los valores altos de la tasa de mortalidad en un área tienden a estar rodeados por otras áreas con tasas de mortalidad altas.

En conclusión, considerando los criterios de AIC, el valor del test LR y el coeficiente de autocorrelación espacial (Rho), el modelo SAR es la mejor opción para este análisis. Esto se debe a que el AIC es ligeramente menor, el valor p del test LR es comparable pero más bajo, y la presencia de un coeficiente Rho significativo sugiere que el modelo SAR captura de manera más efectiva la dependencia espacial entre las variables.

A continuación se presenta el modelo realizado y sus principales resultados:

$$\text{Tasa Mortalidad} = \beta_1 + \beta_2 \log(PM_{10}) + \beta_3 \log(PM_{2,5}) + \beta_4 \text{Localidad} \quad (1)$$

$$+ \beta_5 \text{Reg. Contributivo} + \beta_6 \text{Reg. Subsidiado} + \mu_i \quad (2)$$

Tabla 7-3.: Coeficientes Modelo Cortes (Enero-Abri-Julio-Octubre)

Variable	Coef.	Valor-p	Coef.	Valor-p	Coef.	Valor-p	Coef.	Valor-p
	Enero		Abril		Julio		Octubre	
Intercepto	0,146	0,028	0,045	0,12	0,23	0,00	-0,087	0,021
logPM10	0,024	0,079	0,031	0,00	0,015	0,042	0,041	0,00
logPM2.5	-0,008	0,546	-0,009	0,275	-0,012	0,087	0,005	0,667
Localidad Ciudad Bolívar	-0,025	0,046	-0,02	0,009	0,013	0,192	0,009	0,282
Localidad Engativá	-0,006	0,644	-0,005	0,437	0,004	0,674	0,020	0,018
Localidad Fontibón	-0,014	0,234	-0,008	0,172	-0,008	0,405	0,005	0,551
Localidad Kennedy	-0,013	0,278	-0,013	0,048	0,009	0,352	-0,007	0,409
Localidad Puente Aranda	0,019	0,103	0,021	0,001	0,037	0,00	0,027	0,001
Localidad San Cristobal	-0,004	0,763	0,001	0,913	0,008	0,411	0,042	0,00
Localidad Santa fe	0,031	0,021	0,019	0,01	0,036	0,00	0,037	0,00
Localidad Suba	-0,018	0,178	-0,01	0,133	-0,004	0,667	0,013	0,168
Localidad Tunjuelito	0,007	0,561	0,003	0,661	0,034	0,001	0,013	0,088
Localidad Usaquen	-0,005	0,739	-0,007	0,262	-0,005	0,569	0,023	0,038
Localidad Usme	-0,026	0,053	-0,023	0,005	0,003	0,775	0,008	0,384
Regimen Contributivo	-0,104	0,016	-0,037	0,03	-0,131	0,00	0,017	0,163
Regimen Subsidiado	-0,067	0,133	0,00	0,985	-0,149	0,00	0,013	0,308

Intercepto:

En ausencia de otros factores que influyan en la tasa de mortalidad, se estima que el valor promedio esperado es de 0,146, 0,23 y -0,087 unidades para los meses de enero, julio y octubre respectivamente.

 PM_{10} :

Se evidenció una asociación significativa entre la concentración de PM_{10} y la tasa de mortalidad en los meses de abril, julio y octubre, mientras que en enero la asociación fue débilmente significativa. Esto implica que, controlando el efecto espacial y otras variables independientes, un incremento de una unidad en la concentración de PM_{10} se asocia con un aumento en la tasa de mortalidad. El mayor efecto sobre la tasa de mortalidad se observó en octubre.

 $PM_{2,5}$:

No se encontró evidencia estadísticamente significativa de una asociación entre $PM_{2,5}$ y la tasa de mortalidad para ninguno de los meses estudiados.

Localidad:

- **Ciudad Bolívar y Usme:** En los meses de enero y abril, estas localidades presentan coeficientes negativos y estadísticamente significativos, lo que indica que la tasa de mortalidad en Ciudad Bolívar y Usme es menor que el promedio de la ciudad en comparación con otras localidades durante esos meses. Sin embargo, para los meses de julio y octubre, la tasa de mortalidad en estas dos localidades no presenta una diferencia significativa con respecto al promedio de la ciudad.
- **Engativá, San Cristobal y Usaquen:** Estas localidades presentan coeficiente positivo y significativo únicamente en el mes de octubre. Esto implica que, durante este mes, la tasa de mortalidad en estas localidades supera el promedio de la ciudad. En los meses restantes (enero, abril y julio), la tasa de mortalidad en estas tres localidades no presenta una diferencia estadísticamente significativa respecto al promedio.
- **Fontibón y Suba:** No se observa un patrón definido en los coeficientes de estas localidades y tampoco se detectan diferencias estadísticamente significativas.
- **Kennedy:** Esta localidad presenta un coeficiente negativo y significativo únicamente en el mes de abril. Esto sugiere que la tasa de mortalidad es menor que el promedio de la ciudad, en comparación con otras localidades, para el mes de abril. En los demás meses la tasa de mortalidad no presenta una diferencia significativa con respecto al promedio.
- **Tunjuelito:** Esta localidad presenta un coeficiente positivo y significativo únicamente en el mes de julio. Esto sugiere que la tasa de mortalidad es mayor que el promedio de la ciudad, en comparación con otras localidades. En los demás meses la tasa de mortalidad no presenta una diferencia significativa con respecto al promedio.
- **Puente Aranda y Santa Fé:** Estas localidades presentan coeficientes positivos y significativos en todos los meses (exceptuando el mes de enero en la localidad de Puente Aranda). Esto sugiere que la tasa de mortalidad en estas localidades es mayor que el promedio de la ciudad, en comparación con otras localidades, para todos los meses (exceptuando el mes de enero en la localidad de Puente Aranda).

Régimen Contributivo:

Se observó una asociación significativa entre el régimen contributivo y una menor tasa de mortalidad durante los meses de enero, abril y julio. Esto indica que, controlando el efecto espacial y las demás variables independientes estar vinculado al régimen contributivo se asocia con una caída en la tasa de mortalidad.

Régimen Subsidiado:

No es estadísticamente significativo en los meses de enero, abril y octubre, ya que su valor p está por encima de 0.05.

Hallazgos:

- El modelo SAR captura la autocorrelación espacial en la tasa de mortalidad, lo que significa que las tasas de mortalidad en áreas vecinas se influyen mutuamente.
- Los resultados sugieren que la contaminación del aire, particularmente PM_{10} , tiene un impacto significativo en la tasa de mortalidad, con un efecto más fuerte en los meses posteriores al invierno.
- Ciertas localidades presentan tasas de mortalidad consistentemente más altas o más bajas, lo que podría estar relacionado con factores socioeconómicos, ambientales o de acceso a la atención médica.
- El régimen contributivo se asocia con una menor tasa de mortalidad, lo que podría ser un indicador del acceso a una buena calidad de servicios de salud.

7.3. Modelo Espacio - Temporal STFDF

Para realizar el Modelo Espacio - Temporal STFDF se requiere presentar la información en tres diferentes tablas, así:

- a) **Periodo de tiempo:** Rango de tiempo en el cual se encuentra la investigación. Desde 2018-01 al 2022-12.

Start Date	End Date
2018-01-01 GMT	2022-12-01 GMT

- b) **Espacio (Coordenadas):** Columnas donde se reflejan diferentes ubicaciones o áreas de medición.

En este caso se tomaron los puntos espaciales (Coordenadas) de las localidades analizadas.

Localidad	Latitud	Longitud
Chapinero	4,6452239	-74,06152606
Ciudad Bolívar	4,5778351	-74,16624828
Engativá	4,695020267	-74,09741683
Fontibón	4,67315035	-74,14613022
Kennedy	4,61047115	-74,15488717
Puente Aranda	4,62016285	-74,11618439
San Cristobal	4,5725823	-74,08378439
Santa Fe	4,6255156	-74,06695106
Suba	4,760760033	-74,06899457
Tunjuelito	4,5762545	-74,13092606
Usaquen	4,7103795	-74,03038717
Usme	4,5320851	-74,11710939

c) **Datos de análisis:** Valores de la base a analizar (para este caso, se tomaron las variables de PM_{10} , $PM_{2,5}$, tasa de mortalidad, localidad y régimen de afiliación al sistema de salud).

Registros	Tasa Mortalidad	PM_{10}	$PM_{2,5}$	Localidad	Reg. Cont	Reg. Sub
1	0,0191	42,222	19,283	Chapinero	1	0
2	0,0391	35,580	16,986	Ciudad Bolivar	0,29	0,67
3	0,0593	30,953	14,334	Engativá	0,72	0,21
4	0,0355	41,089	18,331	Fontibón	0,38	0,38
5	0,0313	54,745	26,226	Kennedy	0,65	0,25
.....						
716	0,0929	25,606	18,766	Santa Fe	0,4	0,5
717	0,0368	30,902	14,914	Suba	0,70	0,19
718	0,0440	47,191	24,747	Tunjuelito	0,75	0,125
719	0,0465	27,753	14,445	Usaquén	0,77	0,18
720	0,0274	37,098	14,191	Usme	0,45	0,27

7.3.1. Resultados Modelo Espacio - Temporal STFDF

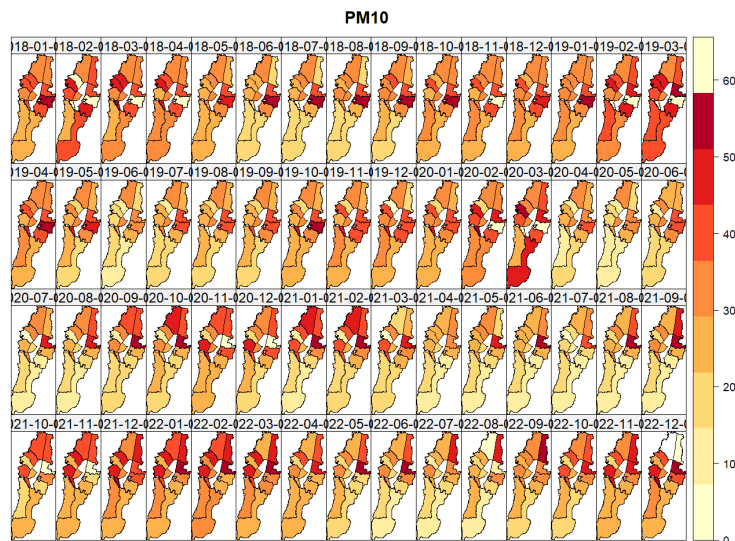


Figura 7-3.: Modelo Espacio - Temporal Material Particulado PM_{10}

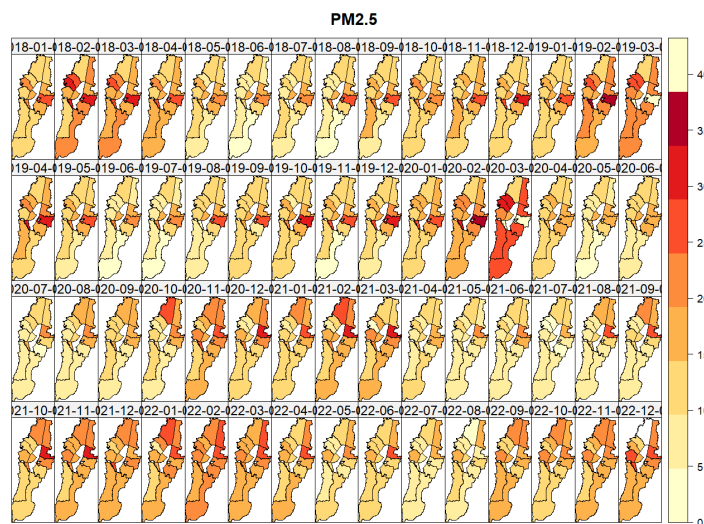


Figura 7-4.: Modelo Espacio - Temporal Material Particulado $PM_{2.5}$

En las Figuras 7-3 y 7-4 se observa un patrón consistente en los primeros tres meses de cada año, donde se registra un pico de contaminación por partículas PM_{10} y $PM_{2,5}$. De igual manera, se nota una tendencia a la disminución de la contaminación de estas partículas desde abril hasta septiembre. Por otro lado, en los últimos tres meses del año se aprecia un aumento en comparación con los meses intermedios.

Es relevante destacar que en marzo del año 2020 se identifica el pico más alto de todos los períodos analizados. Sin embargo, posteriormente a este evento, se observa una disminución en los niveles de partículas de PM_{10} y $PM_{2,5}$ hasta el mes de marzo del año 2021. Este descenso coincide con la llegada de la pandemia por COVID-19, lo que podría explicar la reducción en la contaminación general durante esos períodos en la ciudad de Bogotá.

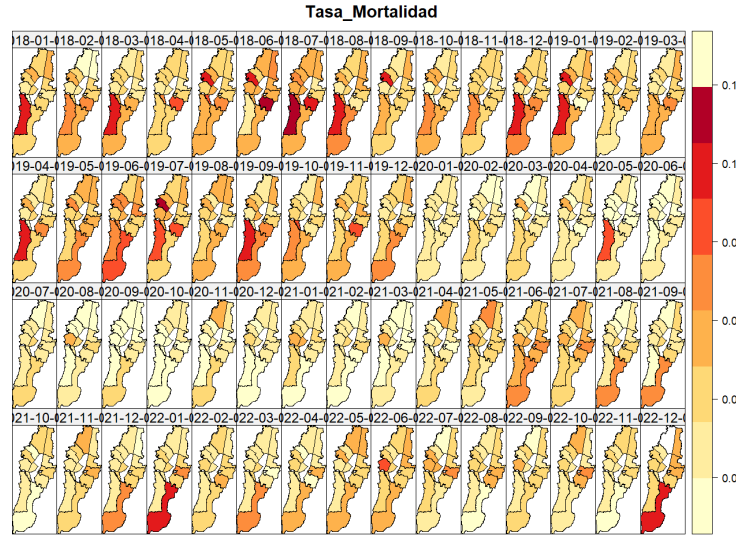


Figura 7-5.: Modelo Espacio Temporal Tasa de Mortalidad por IRA

Finalmente, en la Figura 7-5 se aprecia que hasta el mes de diciembre del año 2019, la tasa de mortalidad por enfermedades respiratorias era significativamente alta. Sin embargo, durante la pandemia por COVID-19 parece ser que estas tasas disminuyeron considerablemente. Este fenómeno podría atribuirse a la implementación de diversas restricciones, cuidados como el uso de tapabocas y otras medidas preventivas que tuvieron un impacto positivo en la reducción de fallecimientos relacionados con dichas enfermedades.

A continuación se presenta el modelo realizado y sus principales resultados:

$$\text{Tasa de Mortalidad} = \beta_1 + \beta_2 \text{Log}(PM_{10}) + \beta_3 \text{Log}(PM_{2,5}) + \beta_4 \text{Localidad} \quad (3)$$

$$+ \beta_5 \text{Reg}_\text{Contributivo} + \beta_6 \text{Reg}_\text{Subsidiado} + \mu_i \quad (4)$$

Tabla 7-4.: Coeficientes Modelo STFDF

Variable	Coef.	Valor-p
logPM10	0,001	0,189
logPM2.5	-0,005	0,002
Localidad Ciudad Bolívar	-0,004	0,010
Localidad Engativá	0,0007	0,77
Localidad Fontibón	-0,005	0,052
Localidad Kennedy	-0,002	0,27
Localidad Puente Aranda	0,01	6.136e-05
Localidad San Cristóbal	0,001	0,69
Localidad Santa Fe	0,013	1.194e-05
Localidad Suba	-0,010	8.280e-05
Localidad Tunjuelito	0,007	0,009
Localidad Usaquén	-0,055	0,03
Localidad Usme	-0,012	6.123e-05
Régimen Contributivo	0,002	0,67
Régimen Subsidiado	0,003	0,55

PM_{10} :

En el modelo STFDF, el coeficiente de $\log PM_{10}$ es positivo (0,001) pero no significativo (p-valor 0,189). Esto sugiere que, aunque hay una relación positiva entre las concentraciones de PM_{10} y la tasa de mortalidad, esta no es estadísticamente significativa en el período y localidades analizadas.

$PM_{2,5}$:

El coeficiente de $\text{Log}(PM_{2,5})$ es negativo (-0,005) y significativo (p-valor 0,002), lo que indica una relación inversa entre las concentraciones de $PM_{2,5}$ y la tasa de mortalidad. Es posible que en este contexto la concentración de $PM_{2,5}$ esté correlacionada con otros factores (como la reducción de movilidad durante la pandemia por COVID-19) que disminuyen la mortalidad por enfermedades respiratorias.

Variables espaciales:

Localidades como Santa Fé (coeficiente positivo y significativo) y Usaquén (coeficiente negativo y significativo) muestran claras diferencias en su impacto en la tasa de mortalidad por enfermedades respiratorias. Localidades con coeficientes positivos (Santa Fé, Tunjuelito) indican un mayor riesgo en comparación con localidades como Usaquén y Suba, que tienen coeficientes negativos.

Régimen de afiliación:

Ni el régimen contributivo ni el subsidiado muestran una relación estadísticamente significativa con la tasa de mortalidad, con p-valores muy por encima de los umbrales convencionales de significancia.

8. Conclusiones y recomendaciones

- El análisis exploratorio permitió visualizar e identificar patrones y tendencias en los registros de contaminación por material particulado PM_{10} y $PM_{2,5}$ a nivel distrital, destacando una amplia variabilidad en los niveles de concentración de estos contaminantes a lo largo del tiempo y en diferentes localidades de Bogotá.
- Para ambas partículas se destaca la presencia de valores atípicos, lo que indica la dispersión de concentración dentro de las localidades, evidenciándose registros que se encuentran significativamente alejados de la mayoría de las observaciones.
- Así mismo, se observa que las localidades de Kennedy, Chapinero, Tunjuelito y Fontibón presentan mayores niveles de concentración de la partícula PM_{10} especialmente en el primer trimestre de los años 2019 a 2021. De igual forma, la partícula $PM_{2,5}$ presenta mayor nivel de concentración en el primer trimestre de los años 2019 y 2022.
- El análisis exploratorio reveló la presencia de grupos etarios de mayor riesgo, como la población de 70 años en adelante, la cual mostró una mayor concentración de defunciones por enfermedades respiratorias en comparación con los demás grupos etarios, lo que sugiere la necesidad de implementar medidas específicas para proteger a estos grupos vulnerables.
- Los modelos ARIMA seleccionados para cada materia particulada tienen en cuenta tanto componentes autorregresivos como de medias móviles, así como componentes estacionales. Estos modelos proporcionan un buen ajuste a los datos históricos y muestran una capacidad predictiva prometedora.
- El Backtesting realizado confirma que los modelos ARIMA se ajustan adecuadamente a los datos históricos, lo que sugiere que las predicciones generadas por estos modelos son probablemente precisas y efectivas.
- Los pronósticos realizados muestran resultados favorables, con un error cuadrático medio (ECM) aceptable tanto para PM_{10} como para $PM_{2,5}$. Esto indica que los modelos son capaces de predecir con precisión los niveles de contaminación atmosférica en Bogotá D.C.
- La utilización de métodos de interpolación espacial permitió estimar la concentración de materia particulada PM_{10} y $PM_{2,5}$ en diferentes ubicaciones de Bogotá,

lo que facilita la comprensión de la distribución espacial de estos contaminantes en la ciudad. Se identificaron variaciones significativas en los niveles de concentración de materia particulada en las localidades de Bogotá, destacando áreas con menores niveles de contaminación como San Cristóbal y Santa Fé; y áreas con mayor concentración del contaminante como las localidades de Ciudad Bolívar, Tunjuelito, Fontibón y Kennedy lo que sugiere la existencia de disparidades en la calidad del aire a nivel local.

- Por medio del proceso de interpolación se observan menores niveles de concentración de la partícula $PM_{2,5}$, frente al PM_{10} , destacándose el mes de julio como uno de los meses con menores niveles de contaminación. Así mismo, se observa que para la partícula PM_{10} su concentración es menor en localidades ubicadas en el nororiente de la ciudad y aumenta en zonas ubicadas en el sur de la ciudad; mientras que para la partícula $PM_{2,5}$ los meses de julio y enero presentan las concentraciones más bajas, frente a los meses de abril y octubre con niveles más elevados.
- Con respecto a la tasa de mortalidad, se observa que su concentración es mayor en localidades ubicadas en el norte y sur de la ciudad, y disminuye en localidades ubicadas en el centro de la ciudad. Así mismo, se observa que las localidades de Santa Fé, La Candelaria, Los Mártires, Antonio Nariño, Puente Aranda, Sumapaz, Barrios Unidos, San Cristóbal y Tunjuelito presentan mayores tasas de mortalidad, mientras que las localidades de Usme, Rafael Uribe, Teusaquillo, Suba, Usaquén y Chapinero presentan menores tasas.
- Los resultados de los modelos SAR y SEM muestran que ambos capturan la autocorrelación espacial en la tasa de mortalidad, sin embargo, el modelo SAR parece ser ligeramente preferido según criterios como el AIC y el valor p del LR test. Se encontró una asociación significativa entre la concentración de PM_{10} y la tasa de mortalidad en varios meses, mientras que, no se encontró evidencia significativa para $PM_{2,5}$. Además, ciertas localidades mostraron tasas de mortalidad más altas o bajas, y el régimen contributivo se asoció con una menor tasa de mortalidad.
- El análisis espacio-temporal revela picos de contaminación en los primeros meses del año, especialmente entre enero y marzo, coincidiendo con las épocas más secas y menos lluviosas en Bogotá. Este aumento de partículas en el aire está asociado con una mayor mortalidad por IRA durante estos meses. En contraste, se observa una reducción en los niveles de contaminación y mortalidad en los meses de abril a septiembre, probablemente debido a una mejora en la calidad del aire relacionada con mayor humedad y lluvias. En marzo de 2020 se observó el pico más alto de contaminación coincidiendo con el inicio de la pandemia de COVID-19, lo que resalta cómo los cambios en los patrones de movilidad y las políticas restrictivas pudieron influir en la calidad del aire y posteriormente en la tasa de mortalidad.

- Los hallazgos del modelo espacio-temporal destacan la fuerte relación entre los niveles de material particulado, especialmente $PM_{2,5}$, y las tasas de mortalidad por enfermedades respiratorias en las localidades más contaminadas de Bogotá. Las variaciones temporales y espaciales de la contaminación influyen en los desenlaces de salud, mientras que el régimen de afiliación parece tener un impacto limitado en la mitigación de estos efectos.
- Se considera relevante implementar medidas efectivas para controlar las emisiones de PM_{10} y $PM_{2,5}$, especialmente en las localidades con mayor concentración de estos contaminantes. Es pertinente la creación de políticas ambientales que regulen las fuentes fijas y móviles de emisión, tales como la industria y el transporte, promoviendo el uso de tecnologías más limpias.
- Es necesario fortalecer el monitoreo de la calidad del aire, particularmente durante los meses donde se presentan picos de contaminación, para tomar medidas preventivas que reduzcan la exposición de la población a niveles peligrosos de PM .
- Es importante aumentar la concientización entre los ciudadanos sobre el impacto de la contaminación en la salud, promoviendo comportamientos que reduzcan su contribución a la contaminación, como el uso de transporte público y la reducción del uso de vehículos privados, además de promover el uso de mascarillas en periodos de alta contaminación para proteger a las poblaciones más vulnerables.

A. Anexo A: Promedio mensual de material particulado PM_{10}

Fecha	Chapinero	Ciudad Bolivar	Engativá	Fontibón	Kennedy	Puente aranda	San Cristóbal	Santa Fe	Suba	Tunjuelito	Usaquén	Usme
201801	42,22	0,00	19,87	0,00	54,75	22,79	24,08	0,00	25,79	37,16	38,33	0,00
201802	60,13	0,00	30,01	0,00	61,97	30,16	38,55	0,00	33,37	49,95	59,91	0,00
201803	52,24	0,00	25,23	0,00	60,49	26,07	35,72	26,21	28,73	49,50	47,87	0,00
201804	48,48	0,00	21,47	0,00	61,40	24,65	34,36	32,51	22,08	46,10	39,91	0,00
201805	29,90	0,00	14,28	0,00	47,44	19,52	26,47	28,31	19,85	27,50	31,27	0,00
201806	20,92	0,00	8,09	0,00	52,15	14,34	17,03	20,59	18,78	23,35	25,83	0,00
201807	4,11	0,00	9,58	0,00	51,00	13,60	17,40	20,87	18,66	25,46	34,56	0,00
201808	0,00	0,00	8,82	0,00	55,08	11,97	16,81	15,81	24,00	25,10	36,19	0,00
201809	0,00	0,00	16,85	0,00	53,26	17,19	20,13	25,54	24,70	33,30	36,55	0,00
201810	0,00	0,00	21,42	0,00	52,81	17,48	25,05	33,24	18,62	42,97	27,21	0,00
201811	0,00	0,00	23,22	0,00	60,57	25,10	26,46	32,76	19,99	42,68	32,29	0,00
201812	35,28	0,00	20,79	16,26	49,99	23,15	21,96	31,27	25,53	39,24	27,44	0,00
201901	34,53	0,00	20,19	18,57	53,79	20,02	25,68	31,85	24,11	32,45	31,46	0,00
201902	59,93	0,00	30,64	26,24	62,61	29,07	36,07	38,86	30,38	44,42	41,63	0,00
201903	60,13	0,00	30,88	27,14	58,97	28,42	37,76	36,91	31,13	44,56	43,71	0,00
201904	45,47	0,00	22,21	19,46	52,35	21,81	26,90	33,48	21,75	39,30	31,37	0,00
201905	46,14	0,00	17,72	16,38	50,17	19,30	20,62	26,87	19,45	31,67	20,28	0,00
201906	0,00	0,00	9,77	12,38	36,07	10,99	12,18	17,91	15,57	21,02	12,30	0,00
201907	0,00	0,00	11,02	14,16	43,26	10,51	16,34	18,55	18,59	22,59	13,51	0,00
201908	4,05	0,00	12,52	15,31	38,46	11,91	16,51	22,21	22,09	23,86	16,37	0,00
201909	42,34	0,00	16,73	17,52	40,39	17,69	20,01	27,80	21,53	26,12	18,83	0,00
201910	48,27	0,00	17,52	18,88	51,69	20,87	24,71	26,20	23,03	32,47	17,05	0,00
201911	49,59	0,00	17,31	17,39	42,76	21,98	27,97	32,35	23,27	41,37	15,31	0,00
201912	47,40	0,00	9,96	18,31	41,54	22,23	26,01	31,77	24,44	38,57	0,01	0,00
202001	47,92	0,00	13,68	18,17	37,95	20,35	27,20	25,02	25,23	42,05	15,32	0,00
202002	57,82	0,00	22,48	23,37	61,02	23,57	35,73	29,47	31,94	50,48	21,27	0,00
202003	62,48	0,00	27,19	25,60	65,52	23,87	44,45	27,63	31,63	56,49	31,91	0,00
202004	27,79	0,00	11,58	12,22	37,23	12,54	17,95	13,93	16,74	25,08	12,66	0,00
202005	22,95	0,00	7,29	9,40	36,30	9,71	11,65	10,48	9,39	20,13	7,59	0,00
202006	31,20	0,00	11,73	14,29	42,51	14,38	19,02	15,91	13,15	28,32	9,93	0,00
202007	25,60	0,00	8,23	5,00	44,01	12,42	12,65	9,76	11,90	21,89	7,01	0,00
202008	29,18	0,00	10,06	12,79	48,71	8,33	15,19	12,79	14,74	28,35	9,03	0,00
202009	33,59	0,00	11,26	15,15	51,63	13,29	15,44	14,95	16,08	33,20	11,29	0,00
202010	38,17	18,87	22,21	16,36	55,53	25,92	16,59	14,64	18,01	33,75	11,57	30,22
202011	49,35	46,73	32,70	17,37	58,79	37,07	27,78	22,06	21,15	47,33	20,40	30,12
202012	42,77	42,75	32,03	19,71	64,09	38,29	23,00	19,63	18,65	45,21	14,60	29,41
202101	30,34	40,71	27,94	16,59	55,16	32,16	19,72	14,32	21,58	38,09	12,80	37,46
202102	43,94	39,93	27,58	16,86	52,58	35,34	25,41	19,20	20,74	38,60	13,68	39,07
202103	45,61	40,57	25,27	22,02	50,53	35,38	26,75	21,03	20,80	41,58	0,01	35,88
202104	21,11	26,17	20,54	31,12	31,14	25,21	13,37	13,58	18,26	23,26	0,00	23,33
202105	28,83	23,49	15,78	21,77	44,17	22,05	15,26	10,87	18,65	24,35	0,00	29,45
202106	32,59	30,34	19,79	34,13	51,80	27,91	12,72	11,35	16,79	27,43	0,00	27,47
202107	26,64	21,86	13,84	33,44	42,35	19,58	10,48	8,72	17,25	19,08	0,67	20,84
202108	23,10	29,44	22,12	40,18	53,57	23,75	13,99	12,81	21,00	27,87	15,08	29,18
202109	25,73	31,41	23,49	43,86	53,60	28,31	13,15	14,52	20,29	29,53	15,45	31,40
202110	26,81	40,67	24,52	38,13	58,72	33,76	21,74	22,02	21,07	39,58	19,97	40,12
202111	37,81	48,79	32,17	41,92	62,48	21,79	27,64	27,62	25,62	48,55	27,33	45,20
202112	36,65	52,79	33,84	46,83	38,02	19,76	28,75	26,60	33,02	48,74	26,75	44,92
202201	33,72	40,71	30,86	44,24	26,82	16,96	27,95	25,05	29,94	45,88	27,73	46,00
202202	42,90	39,93	31,63	48,83	25,20	28,40	28,72	30,37	32,41	50,35	30,60	44,98
202203	36,96	40,57	31,38	48,62	25,75	32,53	30,25	27,91	30,61	48,63	23,01	47,13
202204	29,03	26,17	23,85	38,46	19,66	25,29	24,11	22,80	23,79	35,83	20,37	36,84
202205	28,53	23,49	18,38	39,77	52,50	14,54	19,14	19,08	22,56	35,23	15,41	33,58
202206	26,34	30,34	17,12	38,75	53,79	13,75	19,21	13,59	19,85	35,21	17,28	31,47
202207	23,98	21,86	16,80	45,63	34,93	11,04	17,29	13,05	20,81	29,40	13,86	21,86
202208	24,46	29,44	20,60	49,22	13,65	12,65	20,37	14,49	22,13	28,70	17,62	25,02
202209	4,81	31,41	25,16	52,91	6,46	16,44	24,13	21,25	28,76	35,88	21,98	35,81
202210	0,00	40,67	23,88	40,73	19,55	30,83	23,38	18,30	22,68	37,71	21,76	36,67
202211	0,00	48,79	31,03	48,47	21,81	32,95	27,28	24,51	25,09	44,15	16,56	42,14
202212	0,00	52,79	34,79	58,79	25,97	38,38	29,06	25,61	30,90	47,19	27,75	37,10

B. Anexo B: Promedio mensual de material particulado $PM_{2,5}$

Fecha	Chapinero	Ciudad Bolívar	Engativá	Fontibón	Kennedy	Puente aranda	San Cristóbal	Santa Fe	Suba	Tunjuelito	Usaquén	Usme
201801	19,28	0,00	9,36	0,00	26,23	8,14	14,17	0,00	11,30	21,28	12,22	0,00
201802	29,48	0,00	14,78	0,00	30,75	11,95	23,30	0,00	15,91	31,15	24,21	0,00
201803	27,07	0,00	13,60	0,00	32,31	11,39	19,58	16,89	15,48	26,92	21,04	0,00
201804	24,52	0,00	10,68	0,00	27,37	9,46	14,46	17,74	10,51	22,99	13,57	0,00
201805	16,28	0,00	7,07	0,00	22,09	7,87	8,19	14,04	7,12	14,59	8,64	0,00
201806	13,66	0,00	3,10	0,00	22,59	4,76	4,10	7,42	6,40	7,44	6,44	0,00
201807	2,08	0,00	4,84	0,00	21,94	4,84	5,20	10,57	4,66	10,87	4,80	0,00
201808	0,00	0,00	4,79	0,00	24,92	4,12	4,72	7,37	4,74	6,66	5,26	0,00
201809	0,00	0,00	8,96	0,00	26,53	5,84	9,07	14,41	7,96	18,90	11,36	0,00
201810	0,00	0,00	9,30	0,00	21,78	7,10	10,24	11,95	8,40	22,25	14,35	0,00
201811	0,00	0,00	9,69	0,00	24,79	9,42	10,95	15,76	7,81	13,96	14,68	0,00
201812	25,34	0,00	7,51	5,47	29,18	8,91	9,63	11,71	9,44	16,17	11,05	0,00
201901	20,59	0,00	9,24	6,74	26,50	8,09	12,81	13,60	9,64	15,91	14,16	0,00
201902	37,28	0,00	15,11	10,60	36,74	12,61	17,99	20,82	16,15	24,75	23,35	0,00
201903	39,47	0,00	18,63	11,90	43,16	13,25	20,44	23,05	17,81	27,76	26,79	0,00
201904	29,41	0,00	11,99	10,73	32,24	9,70	13,89	18,10	10,62	20,10	17,51	0,00
201905	22,91	0,00	7,82	8,68	28,42	7,91	8,51	12,49	6,80	14,76	11,40	0,00
201906	0,00	0,00	3,38	6,34	22,01	5,14	3,74	6,08	4,70	6,72	7,19	0,00
201907	0,00	0,00	3,46	6,79	22,59	4,94	4,36	6,24	5,92	7,21	8,08	0,00
201908	2,05	0,00	5,74	8,21	26,58	4,75	5,38	9,54	8,58	10,28	9,74	0,00
201909	18,74	0,00	6,59	9,20	27,55	0,45	0,00	10,02	8,29	11,33	11,73	0,00
201910	22,12	0,00	7,34	9,62	30,69	0,00	0,00	11,00	8,26	14,73	9,51	0,00
201911	26,47	0,00	8,79	9,19	27,12	3,00	4,34	13,80	8,69	16,14	14,17	0,00
201912	25,34	0,00	8,84	9,91	28,88	7,41	7,66	12,52	9,60	15,17	11,43	0,00
202001	26,08	0,00	9,99	9,61	25,98	7,35	11,04	13,17	10,18	15,32	15,48	0,00
202002	32,30	0,00	12,03	13,38	33,83	13,41	17,37	16,59	14,29	21,27	22,19	0,00
202003	42,71	0,00	21,56	18,55	40,80	18,48	28,20	28,51	20,93	31,91	33,27	0,00
202004	18,32	0,00	7,84	9,03	18,70	9,37	11,24	12,53	9,86	12,66	15,71	0,00
202005	12,63	0,00	3,67	5,92	15,55	6,08	4,48	6,68	5,01	7,59	6,77	0,00
202006	14,58	0,00	4,93	6,95	15,97	7,60	6,07	8,60	5,75	9,93	9,78	0,00
202007	11,72	0,00	3,96	2,65	19,39	6,35	3,96	6,20	4,51	7,01	6,72	0,00
202008	13,94	0,00	5,58	6,88	20,86	7,37	5,44	7,64	5,83	9,03	9,32	0,00
202009	17,57	0,00	6,76	8,27	18,81	9,68	6,27	8,92	7,25	11,29	10,07	0,00
202010	17,44	10,12	9,22	8,51	23,13	14,90	8,70	9,39	6,96	11,57	7,72	11,84
202011	25,63	24,93	16,10	9,97	27,32	21,46	16,49	16,78	10,22	20,40	12,24	18,05
202012	23,24	20,91	14,29	10,64	30,84	21,44	12,79	13,58	8,78	14,69	7,43	7,84
202101	16,14	18,97	14,24	8,90	24,04	18,47	10,53	11,79	9,67	14,19	6,94	10,29
202102	21,98	19,50	15,39	9,81	29,18	18,99	14,54	15,48	12,41	17,25	7,59	12,92
202103	25,50	20,52	16,24	12,20	33,07	22,20	15,83	15,38	13,14	20,43	0,01	14,60
202104	12,12	12,20	10,87	12,08	17,62	14,85	8,54	9,69	7,81	8,33	0,00	5,79
202105	11,15	12,23	8,04	8,51	19,44	12,42	6,72	8,65	6,63	7,63	0,00	8,04
202106	12,11	13,84	10,01	11,99	21,06	14,69	6,50	8,69	7,65	7,05	0,00	6,74
202107	9,70	9,04	6,68	10,75	16,78	9,48	4,21	6,44	5,00	6,01	0,25	4,08
202108	14,02	12,73	11,43	16,77	24,05	12,53	8,67	7,88	7,44	8,76	5,06	5,97
202109	15,41	13,96	11,67	18,38	26,63	14,16	7,11	8,12	9,07	10,01	7,06	7,62
202110	20,78	19,93	14,66	19,26	31,73	19,62	12,68	13,24	12,01	17,15	12,14	13,41
202111	22,27	23,23	18,22	21,57	32,53	18,72	18,03	15,64	13,53	20,07	16,38	16,09
202112	21,54	25,18	18,48	23,14	17,48	19,89	19,01	17,45	16,06	17,29	14,02	16,39
202201	20,37	18,97	16,49	22,89	11,72	17,45	20,05	18,22	11,82	16,14	5,87	19,47
202202	26,85	19,50	18,76	27,77	13,26	23,42	21,65	23,38	18,34	20,93	19,39	18,05
202203	21,52	20,52	16,90	25,77	13,04	20,08	19,09	18,70	16,09	19,91	16,52	17,23
202204	16,31	12,20	13,63	20,85	10,41	16,35	14,23	16,65	13,31	15,66	15,13	14,23
202205	13,39	12,23	10,20	15,16	22,57	12,18	9,55	10,95	9,37	16,68	6,65	8,14
202206	10,51	13,84	11,27	17,30	25,61	13,58	10,05	11,68	10,20	19,00	9,59	10,79
202207	12,29	9,04	8,54	15,81	14,53	9,54	6,54	8,10	8,40	15,07	6,77	5,10
202208	12,48	12,73	11,79	18,25	6,07	11,92	7,74	9,33	10,75	16,54	8,58	8,43
202209	3,38	13,96	14,59	21,79	3,86	16,39	12,54	14,98	15,49	20,85	12,83	13,62
202210	0,00	19,93	14,56	18,02	8,67	16,18	14,18	13,13	12,06	22,42	11,30	15,23
202211	0,00	23,23	16,74	19,39	8,98	17,57	15,96	16,09	13,22	23,35	13,25	15,45
202212	0,00	25,18	17,89	21,63	13,81	20,60	18,46	18,77	14,91	24,75	14,45	14,19

C. Anexo C: Promedio de medición de material particulado PM_{10} -Interpolada

Fecha	Chapinero	Ciudad Bolívar	Engativá	Fontibón	Kennedy	Puente aranda	San Cristóbal	Santa Fe	Suba	Tunjuelito	Usaquén	Usme
201801	42,22	35,58	19,87	27,93	54,75	22,79	24,08	22,19	25,79	37,16	38,33	34,54
201802	60,13	35,58	30,01	27,93	61,97	30,16	38,55	22,19	33,37	49,95	59,91	34,54
201803	52,24	35,58	25,23	27,93	60,49	26,07	35,72	26,21	28,73	49,50	47,87	34,54
201804	48,48	35,58	21,47	27,93	61,40	24,65	34,36	32,51	22,08	46,10	39,91	34,54
201805	29,90	35,58	14,28	27,93	47,44	19,52	26,47	28,31	19,85	27,50	31,27	34,54
201806	20,92	35,58	8,09	27,93	52,15	14,34	17,03	20,59	18,78	23,35	25,83	34,54
201807	4,11	35,58	9,58	27,93	51,00	13,60	17,40	20,87	18,66	25,46	34,56	34,54
201808	35,88	35,58	8,82	27,93	55,08	11,97	16,81	15,81	24,00	25,10	36,19	34,54
201809	35,88	35,58	16,85	27,93	53,26	17,19	20,13	25,54	24,70	33,30	36,55	34,54
201810	35,88	35,58	21,42	27,93	52,81	17,48	25,05	33,24	18,62	42,97	27,21	34,54
201811	35,88	35,58	23,22	27,93	60,57	25,10	26,46	32,76	19,99	42,68	32,29	34,54
201812	35,28	35,58	20,79	16,26	49,99	23,15	21,96	31,27	25,53	39,24	27,44	34,54
201901	34,53	35,58	20,19	18,57	53,79	20,02	25,68	31,85	24,11	32,45	31,46	34,54
201902	59,93	35,58	30,64	26,24	62,61	29,07	36,07	38,86	30,38	44,42	41,63	34,54
201903	60,13	35,58	30,88	27,14	58,97	28,42	37,76	36,91	31,13	44,56	43,71	34,54
201904	45,47	35,58	22,21	19,46	52,35	21,81	26,90	33,48	21,75	39,30	31,37	34,54
201905	46,14	35,58	17,72	16,38	50,17	19,30	20,62	26,87	19,45	31,67	20,28	34,54
201906	35,88	35,58	9,77	12,38	36,07	10,99	12,18	17,91	15,57	21,02	12,30	34,54
201907	35,88	35,58	11,02	14,16	43,26	10,51	16,34	18,55	18,59	22,59	13,51	34,54
201908	4,05	35,58	12,52	15,31	38,46	11,91	16,51	22,21	22,09	23,86	16,37	34,54
201909	42,34	35,58	16,73	17,52	40,39	17,69	20,01	27,80	21,53	26,12	18,83	34,54
201910	48,27	35,58	17,52	18,88	51,69	20,87	24,71	26,20	23,03	32,47	17,05	34,54
201911	49,59	35,58	17,31	17,39	42,76	21,98	27,97	32,35	23,27	41,37	15,31	34,54
201912	47,40	35,58	9,96	18,31	41,54	22,23	26,01	31,77	24,44	38,57	0,01	34,54
202001	47,92	35,58	13,68	18,17	37,95	20,35	27,20	25,02	25,23	42,05	15,32	34,54
202002	57,82	35,58	22,48	23,37	61,02	23,57	35,73	29,47	31,94	50,48	21,27	34,54
202003	62,48	35,58	27,19	25,60	65,52	23,87	44,45	27,63	31,63	56,49	31,91	34,54
202004	27,79	35,58	11,58	12,22	37,23	12,54	17,95	13,93	16,74	25,08	12,66	34,54
202005	22,95	35,58	7,29	9,40	36,30	9,71	11,65	10,48	9,39	20,13	7,59	34,54
202006	31,20	35,58	11,73	14,29	42,51	14,38	19,02	15,91	13,15	28,32	9,93	34,54
202007	25,60	35,58	8,23	5,00	44,01	12,42	12,65	9,76	11,90	21,89	7,01	34,54
202008	29,18	35,58	10,06	12,79	48,71	8,33	15,19	12,79	14,74	28,35	9,03	34,54
202009	33,59	35,58	11,26	15,15	51,63	13,29	15,44	14,95	16,08	33,20	11,29	34,54
202010	38,17	18,87	22,21	16,36	55,53	25,92	16,59	14,64	18,01	33,75	11,57	30,22
202011	49,35	46,73	32,70	17,37	58,79	37,07	27,78	22,06	21,15	47,33	20,40	30,12
202012	42,77	42,75	32,03	19,71	64,09	38,29	23,00	19,63	18,65	45,21	14,60	29,41
202101	30,34	40,71	27,94	16,59	55,16	32,16	19,72	14,32	21,58	38,09	12,80	37,46
202102	43,94	39,93	27,58	16,86	52,58	35,34	25,41	19,20	20,74	38,60	13,68	39,07
202103	45,61	40,57	25,27	22,02	50,53	35,38	26,75	21,03	20,80	41,58	0,01	35,88
202104	21,11	26,17	20,54	31,12	31,14	25,21	13,37	13,58	18,26	23,26	22,06	23,33
202105	28,83	23,49	15,78	21,77	44,17	22,05	15,26	10,87	18,65	24,35	22,06	29,45
202106	32,59	30,34	19,79	34,13	51,80	27,91	12,72	11,35	16,79	27,43	22,06	27,47
202107	26,64	21,86	13,84	33,44	42,35	19,58	10,48	8,72	17,25	19,08	0,67	20,84
202108	23,10	29,44	22,12	40,18	53,57	23,75	13,99	12,81	21,00	27,87	15,08	29,18
202109	25,73	31,41	23,49	43,86	53,60	28,31	13,15	14,52	20,29	29,53	15,45	31,40
202110	26,81	40,67	24,52	38,13	58,72	33,76	21,74	22,02	21,07	39,58	19,97	40,12
202111	37,81	48,79	32,17	41,92	62,48	21,79	27,64	27,62	25,62	48,55	27,33	45,20
202112	36,65	52,79	33,84	46,83	38,02	19,76	28,75	26,60	33,02	48,74	26,75	44,92
202201	33,72	40,71	30,86	44,24	26,82	16,96	27,95	25,05	29,94	45,88	27,73	46,00
202202	42,90	39,93	31,63	48,83	25,20	28,40	28,72	30,37	32,41	50,35	30,60	44,98
202203	36,96	40,57	31,38	48,62	25,75	32,53	30,25	27,91	30,61	48,63	23,01	47,13
202204	29,03	26,17	23,85	38,46	19,66	25,29	24,11	22,80	23,79	35,83	20,37	36,84
202205	28,53	23,49	18,38	39,77	52,50	14,54	19,14	19,08	22,56	35,23	15,41	33,58
202206	26,34	30,34	17,12	38,75	53,79	13,75	19,21	13,59	19,85	35,21	17,28	31,47
202207	23,98	21,86	16,80	45,63	34,93	11,04	17,29	13,05	20,81	29,40	13,86	21,86
202208	24,46	29,44	20,60	49,22	13,65	12,65	20,37	14,49	22,13	28,70	17,62	25,02
202209	4,81	31,41	25,16	52,91	6,46	16,44	24,13	21,25	28,76	35,88	21,98	35,81
202210	35,88	40,67	23,88	40,73	19,55	30,83	23,38	18,30	22,68	37,71	21,76	36,67
202211	35,88	48,79	31,03	48,47	21,81	32,95	27,28	24,51	25,09	44,15	16,56	42,14
202212	35,88	52,79	34,79	58,79	25,97	38,38	29,06	25,61	30,90	47,19	27,75	37,10

D. Anexo D: Promedio de medición de material Particulado $PM_{2,5}$ -Interpolada

Fecha	Chapinero	Ciudad Bolivar	Engativá	Fontibón	Kennedy	Puente aranda	San Cristóbal	Santa Fe	Suba	Tunjuelito	Usaquén	Usme
201801	19,28	16,99	9,36	13,10	26,23	8,14	14,17	12,97	11,30	21,28	12,22	11,84
201802	29,48	16,99	14,78	13,10	30,75	11,95	23,30	12,97	15,91	31,15	24,21	11,84
201803	27,07	16,99	13,60	13,10	32,31	11,39	19,58	16,89	15,48	26,92	21,04	11,84
201804	24,52	16,99	10,68	13,10	27,37	9,46	14,46	17,74	10,51	22,99	13,57	11,84
201805	16,28	16,99	7,07	13,10	22,09	7,87	8,19	14,04	7,12	14,59	8,64	11,84
201806	13,66	16,99	3,10	13,10	22,59	4,76	4,10	7,42	6,40	7,44	6,44	11,84
201807	2,08	16,99	4,84	13,10	21,94	4,84	5,20	10,57	4,66	10,87	4,80	11,84
201808	19,57	16,99	4,79	13,10	24,92	4,12	4,72	7,37	4,74	6,66	5,26	11,84
201809	19,57	16,99	8,96	13,10	26,53	5,84	9,07	14,41	7,96	18,90	11,36	11,84
201810	19,57	16,99	9,30	13,10	21,78	7,10	10,24	11,95	8,40	22,25	14,35	11,84
201811	19,57	16,99	9,69	13,10	24,79	9,42	10,95	15,76	7,81	13,96	14,68	11,84
201812	25,34	16,99	7,51	5,47	29,18	8,91	9,63	11,71	9,44	16,17	11,05	11,84
201901	20,59	16,99	9,24	6,74	26,50	8,09	12,81	13,60	9,64	15,91	14,16	11,84
201902	37,28	16,99	15,11	10,60	36,74	12,61	17,99	20,82	16,15	24,75	23,35	11,84
201903	39,47	16,99	18,63	11,90	43,16	13,25	20,44	23,05	17,81	27,76	26,79	11,84
201904	29,41	16,99	11,99	10,73	32,24	9,70	13,89	18,10	10,62	20,10	17,51	11,84
201905	22,91	16,99	7,82	8,68	28,42	7,91	8,51	12,49	6,80	14,76	11,40	11,84
201906	19,57	16,99	3,38	6,34	22,01	5,14	3,74	6,08	4,70	6,72	7,19	11,84
201907	19,57	16,99	3,46	6,79	22,59	4,94	4,36	6,24	5,92	7,21	8,08	11,84
201908	2,05	16,99	5,74	8,21	26,58	4,75	5,38	9,54	8,58	10,28	9,74	11,84
201909	18,74	16,99	6,59	9,20	27,55	0,45	11,50	10,02	8,29	11,33	11,73	11,84
201910	22,12	16,99	7,34	9,62	30,69	11,96	11,50	11,00	8,26	14,73	9,51	11,84
201911	26,47	16,99	8,79	9,19	27,12	3,00	4,34	13,80	8,69	16,14	14,17	11,84
201912	25,34	16,99	8,84	9,91	28,88	7,41	7,66	12,52	9,60	15,17	11,43	11,84
202001	26,08	16,99	9,99	9,61	25,98	7,35	11,04	13,17	10,18	15,32	15,48	11,84
202002	32,30	16,99	12,03	13,38	33,83	13,41	17,37	16,59	14,29	21,27	22,19	11,84
202003	42,71	16,99	21,56	18,55	40,80	18,48	28,20	28,51	20,93	31,91	33,27	11,84
202004	18,32	16,99	7,84	9,03	18,70	9,37	11,24	12,53	9,86	12,66	15,71	11,84
202005	12,63	16,99	3,67	5,92	15,55	6,08	4,48	6,68	5,01	7,59	6,77	11,84
202006	14,58	16,99	4,93	6,95	15,97	7,60	6,07	8,60	5,75	9,93	9,78	11,84
202007	11,72	16,99	3,96	2,65	19,39	6,35	3,96	6,20	4,51	7,01	6,72	11,84
202008	13,94	16,99	5,58	6,88	20,86	7,37	5,44	7,64	5,83	9,03	9,32	11,84
202009	17,57	16,99	6,76	8,27	18,81	9,68	6,27	8,92	7,25	11,29	10,07	11,84
202010	17,44	10,12	9,22	8,51	23,13	14,90	8,70	9,39	6,96	11,57	7,72	11,84
202011	25,63	24,93	16,10	9,97	27,32	21,46	16,49	16,78	10,22	20,40	12,24	18,05
202012	23,24	20,91	14,29	10,64	30,84	21,44	12,79	13,58	8,78	14,69	7,43	7,84
202101	16,14	18,97	14,24	8,90	24,04	18,47	10,53	11,79	9,67	14,19	6,94	10,29
202102	21,98	19,50	15,39	9,81	29,18	18,99	14,54	15,48	12,41	17,25	7,59	12,92
202103	25,50	20,52	16,24	12,20	33,07	22,20	15,83	15,38	13,14	20,43	0,01	14,60
202104	12,12	12,20	10,87	12,08	17,62	14,85	8,54	9,69	7,81	8,33	11,91	5,79
202105	11,15	12,23	8,04	8,51	19,44	12,42	6,72	8,65	6,63	7,63	11,91	8,04
202106	12,11	13,84	10,01	11,99	21,06	14,69	6,50	8,69	7,65	7,05	11,91	6,74
202107	9,70	9,04	6,68	10,75	16,78	9,48	4,21	6,44	5,00	6,01	0,25	4,08
202108	14,02	12,73	11,43	16,77	24,05	12,53	8,67	7,88	7,44	8,76	5,06	5,97
202109	15,41	13,96	11,67	18,38	26,63	14,16	7,11	8,12	9,07	10,01	7,06	7,62
202110	20,78	19,93	14,66	19,26	31,73	19,62	12,68	13,24	12,01	17,15	12,14	13,41
202111	22,27	23,23	18,22	21,57	32,53	18,72	18,03	15,64	13,53	20,07	16,38	16,09
202112	21,54	25,18	18,48	23,14	17,48	19,89	19,01	17,45	16,06	17,29	14,02	16,39
202201	20,37	18,97	16,49	22,89	11,72	17,45	20,05	18,22	11,82	16,14	5,87	19,47
202202	26,85	19,50	18,76	27,77	13,26	23,42	21,65	23,38	18,34	20,93	19,39	18,05
202203	21,52	20,52	16,90	25,77	13,04	20,08	19,09	18,70	16,09	19,91	16,52	17,23
202204	16,31	12,20	13,63	20,85	10,41	16,35	14,23	16,65	13,31	15,66	15,13	14,23
202205	13,39	12,23	10,20	15,16	22,57	12,18	9,55	10,95	9,37	16,68	6,65	8,14
202206	10,51	13,84	11,27	17,30	25,61	13,58	10,05	11,68	10,20	19,00	9,59	10,79
202207	12,29	9,04	8,54	15,81	14,53	9,54	6,54	8,10	8,40	15,07	6,77	5,10
202208	12,48	12,73	11,79	18,25	6,07	11,92	7,74	9,33	10,75	16,54	8,58	8,43
202209	3,38	13,96	14,59	21,79	3,86	16,39	12,54	14,98	15,49	20,85	12,83	13,62
202210	19,57	19,93	14,56	18,02	8,67	16,18	14,18	13,13	12,06	22,42	11,30	15,23
202211	19,57	23,23	16,74	19,39	8,98	17,57	15,96	16,09	13,22	23,35	13,25	15,45
202212	19,57	25,18	17,89	21,63	13,81	20,60	18,46	18,77	14,91	24,75	14,45	14,19

E. Anexo E: Clasificación Códigos CIE - Relacionados con IRA

Codigo CIE	Descripción
JJ181	NEUMONIA LOBAR, NO ESPECIFICADA
J189	NEUMONIA, NO ESPECIFICADA
J449	ENFERMEDAD PULMONAR OBSTRUCTIVA CRONICA, NO ESPECIFICADA
J841	OTRAS ENFERMEDADES PULMONARES INTERSTICIALES CON FIBROSIS
J984	OTROS TRASTORNOS DEL PULMON
J440	ENFERMEDAD PULMONAR OBSTRUCTIVA CRONICA CON INFECCION AGUDA DE LAS VIAS RESPIRATORIAS INFERIORES
J110	INFLUENZA CON NEUMONIA, VIRUS NO IDENTIFICADO
J849	ENFERMEDAD PULMONAR INTERSTICIAL, NO ESPECIFICADA
J81X	EDEMA PULMONAR
J689	AFECCION RESPIRATORIA NO ESPECIFICADA, DEBIDA A INHALACION DE GASES, HUMOS, VAPORES Y SUSTANCIAS QUIMICAS
J439	ENFISEMA, NO ESPECIFICADO
J46X	ESTADO ASMATICO
J159	NEUMONIA BACTERIANA, NO ESPECIFICADA
J180	BRONCONEUMONIA, NO ESPECIFICADA
J441	ENFERMEDAD PULMONAR OBSTRUCTIVA CRONICA CON EXACERBACION AGUDA, NO ESPECIFICADA
J985	ENFERMEDADES DEL MEDIASTINO, NO CLASIFICADAS EN OTRA PARTE
J850	GANGRENA Y NECROSIS DEL PULMON
J969	INSUFICIENCIA RESPIRATORIA, NO ESPECIFICADA
J690	NEUMONITIS DEBIDA A ASPIRACION DE ALIMENTO O VOMITO
J219	BRONQUIOLITIS AGUDA, NO ESPECIFICADA
J869	PIOTORAX SIN FISTULA
J60X	NEUMOCONIOSIS DE LOS MINEROS DEL CARBON
J677	NEUMONITIS DE LA VENTILACION DEBIDA AL ACONDICIONADOR Y HUMIDIFICADOR DEL AIRE
J980	ENFERMEDADES DE LA TRAQUEA Y DE LOS BRONQUIOS, NO CLASIFICADAS EN OTRA PARTE
J628	NEUMOCONIOSIS DEBIDA A OTROS POLVOS QUE CONTIENEN SILICE
J988	OTROS TRASTORNOS RESPIRATORIOS ESPECIFICADOS
J40X	BRONQUITIS, NO ESPECIFICADA COMO AGUDA O CRONICA
J448	OTRAS ENFERMEDADES PULMONARES OBSTRUCTIVAS CRONICAS ESPECIFICADAS
J459	ASMA, NO ESPECIFICADO
J852	ABSCESO DEL PULMON SIN NEUMONIA
J42X	BRONQUITIS CRONICA NO ESPECIFICADA
J986	TRASTORNOS DEL DIAFRAGMA
J960	INSUFICIENCIA RESPIRATORIA AGUDA
J939	NEUMOTORAX, NO ESPECIFICADO
J64X	NEUMOCONIOSIS, NO ESPECIFICADA
J90X	DERRAME PLEURAL NO CLASIFICADO EN OTRA PARTE
J680	BRONQUITIS Y NEUMONITIS DEBIDAS A INHALACION DE GASES, HUMOS, VAPORES Y SUSTANCIAS QUIMICAS
J158	OTRAS NEUMONIAS BACTERIANAS
J981	COLAPSO PULMONAR
J129	NEUMONIA VIRAL, NO ESPECIFICADA
J182	NEUMONIA HIPOSTATICA, NO ESPECIFICADA
J22X	INFECCION AGUDA NO ESPECIFICADA DE LAS VIAS RESPIRATORIAS INFERIORES
J209	BRONQUITIS AGUDA, NO ESPECIFICADA
J041	TRAQUEITIS AGUDA
J387	OTRAS ENFERMEDADES DE LA LARINGE
J848	OTRAS ENFERMEDADES PULMONARES INTERSTICIALES ESPECIFICADAS
J941	FIBROTORAX
J398	OTRAS ENFERMEDADES ESPECIFICADAS DE LAS VIAS RESPIRATORIAS SUPERIORES
J13X	NEUMONIA DEBIDA A STREPTOCOCCUS PNEUMONIAE
J111	INFLUENZA CON OTRAS MANIFESTACIONES RESPIRATORIAS, VIRUS NO IDENTIFICADO
J688	OTRAS AFECCIONES RESPIRATORIAS DEBIDAS A INHALACION DE GASES, HUMOS, VAPORES Y SUSTANCIAS QUIMICAS

Código CIE	Descripción
J121	NEUMONIA DEBIDA A VIRUS SINCITIAL RESPIRATORIO
J942	HEMOTORAX
J120	NEUMONIA DEBIDA A ADENOVIRUS
J679	NEUMONITIS DEBIDA A HIPERSENSIBILIDAD A POLVO ORGANICO NO ESPECIFICADO
J860	PIOTORAX CON FISTULA
J384	EDEMA DE LARINGE
J80X	SINDROME DE DIFICULTAD RESPIRATORIA DEL ADULTO
J840	AFECCIONES ALVEOLARES Y ALVEOLOPARIETALES
J930	NEUMOTORAX ESPONTANEO A PRESION
J47X	BRONQUIECTASIA
J36X	ABSCESO PERIAMIGDALINO
J931	OTROS TIPOS DE NEUMOTORAX ESPONTANEO
J431	ENFISEMA PANLOBULAR
J329	SINUSITIS CRONICA, NO ESPECIFICADA
J152	NEUMONIA DEBIDA A ESTAFILOCOCCOS
J210	BRONQUIOLITIS AGUDA DEBIDA A VIRUS SINCITIAL RESPIRATORIO
J069	INFECCION AGUDA DE LAS VIAS RESPIRATORIAS SUPERIORES, NO ESPECIFICADA
J61X	NEUMOCONIOSIS DEBIDA AL ASBESTO Y A OTRAS FIBRAS MINERALES
J704	TRASTORNOS PULMONARES INTERSTICIALES NO ESPECIFICADOS INDUCIDOS POR DROGAS
J698	NEUMONITIS DEBIDA A ASPIRACION DE OTROS SOLIDOS Y LIQUIDOS
J430	SINDROME DE MACLEOD
J961	INSUFICIENCIA RESPIRATORIA CRONICA
J150	NEUMONIA DEBIDA A KLEBSIELLA PNEUMONIAE
J65X	NEUMOCONIOSIS ASOCIADA CON TUBERCULOSIS
J188	OTRAS NEUMONIAS, DE MICROORGANISMO NO ESPECIFICADO
J068	OTRAS INFECCIONES AGUDAS DE SITIOS MULTIPLES DE LAS VIAS RESPIRATORIAS SUPERIORES
J940	QUILOTORAX
J938	OTROS NEUMOTORAX
J156	NEUMONIA DEBIDA A OTRAS BACTERIAS AEROBICAS GRAMNEGATIVAS
J410	BRONQUITIS CRONICA SIMPLE
J208	BRONQUITIS AGUDA DEBIDA A OTROS MICROORGANISMOS ESPECIFICADOS
J154	NEUMONIA DEBIDA A OTROS ESTREPTOCOCCOS
J386	ESTENOSIS LARINGEA
J051	EPIGLOTITIS AGUDA
J983	ENFISEMA COMPENSATORIO
J948	OTRAS AFECCIONES ESPECIFICADAS DE LA PLEURA
J392	OTRAS ENFERMEDADES DE LA FARINGE
J678	NEUMONITIS DEBIDAS A HIPERSENSIBILIDAD A OTROS POLVOS ORGANICOS
J039	AMIGDALITIS AGUDA, NO ESPECIFICADA
J122	NEUMONIA DEBIDA A VIRUS PARAINFLUENZA
J040	LARINGITIS AGUDA
J348	OTROS TRASTORNOS ESPECIFICADOS DE LA NARIZ Y DE LOS SENOS PARANASALES
J218	BRONQUIOLITIS AGUDA DEBIDA A OTROS MICROORGANISMOS ESPECIFICADOS
J390	ABSCESO RETROFARINGEO Y PARAFARINGEO
J050	LARINGITIS OBSTRUCTIVA, AGUDA [CRUP]
J989	TRASTORNO RESPIRATORIO, NO ESPECIFICADO
J920	PAQUIPLEURITIS CON ASBESTOSIS
J684	AFECCIONES RESPIRATORIAS CRONICAS DEBIDAS A INHALACION DE GASES, HUMOS, VAPORES Y SUSTANCIAS QUIMICAS
J418	BRONQUITIS CRONICA MIXTA SIMPLE Y MUCOPURULENTA
J205	BRONQUITIS AGUDA DEBIDA A VIRUS SINCITIAL RESPIRATORIO
J90	DERRAME PLEURAL NO CLASIFICADO EN OTRA PARTE

Código CIE	Descripción
J09X	INFLUENZA DEBIDA A VIRUS DE LA INFLUENZA ZONOTICO O PANDEMICO IDENTIFICADOS
J851	ABSCESO DEL PULMON CON NEUMONIA
J151	NEUMONIA DEBIDA A PSEUDOMONAS
J634	SIDEROSIS
J042	LARINGOTRAQUEITIS AGUDA
J324	PANSINUSITIS CRONICA
J100	INFLUENZA CON NEUMONIA, DEBIDA A OTRO VIRUS DE LA INFLUENZA ESTACIONAL IDENTIFICADO
J450	ASMA PREDOMINANTEMENTE ALERGICA
J391	OTROS ABSCESOS DE LA FARINGE
J206	BRONQUITIS AGUDA DEBIDA A RINOVIRUS
J668	ENFERMEDAD DE LAS VIAS AEREAS DEBIDA A OTROS POLVOS ORGANICOS ESPECIFICOS
J019	SINUSITIS AGUDA, NO ESPECIFICADA
J708	AFECCIONES RESPIRATORIAS DEBIDAS A OTROS AGENTES EXTERNOS ESPECIFICADOS
J168	NEUMONIA DEBIDA A OTROS MICROORGANISMOS INFECCIOSOS ESPECIFICADOS
J399	ENFERMEDAD DE LAS VIAS RESPIRATORIAS SUPERIORES, NO ESPECIFICADA
J411	BRONQUITIS CRONICA MUCOPURULENTA
J393	REACCION DE HIPERSENSIBILIDAD DE LAS VIAS RESPIRATORIAS SUPERIORES, SITIO NO ESPECIFICADO
J385	ESPASMO LARINGEO
J660	BISINOSIS
J030	AMIGDALITIS ESTREPTOCOCICA
J128	NEUMONIA DEBIDA A OTROS VIRUS
J638	NEUMOCONIOSIS DEBIDA A OTROS POLVOS INORGANICOS ESPECIFICADOS
J853	ABSCESO DEL MEDIASTINO
J157	NEUMONIA DEBIDA A MYCOPLASMA PNEUMONIAE
J014	PANSINUSITIS AGUDA
J670	PULMON DEL GRANJERO
J709	AFECCIONES RESPIRATORIAS DEBIDAS A AGENTES EXTERNOS NO ESPECIFICADOS
J46	ESTADO ASMATICO
J42	BRONQUITIS CRONICA NO ESPECIFICADA
J47	BRONQUIECTASIA
J64	NEUMOCONIOSIS, NO ESPECIFICADA
J65	NEUMOCONIOSIS ASOCIADA CON TUBERCULOSIS
J22	INFECCION AGUDA NO ESPECIFICADA DE LAS VIAS RESPIRATORIAS INFERIORES
J81	EDEMA PULMONAR
J80	SINDROME DE DIFICULTAD RESPIRATORIA DEL ADULTO
J60	NEUMOCONIOSIS DE LOS MINEROS DEL CARBON
J202	BRONQUITIS AGUDA DEBIDA A ESTREPTOCOCOS
J13	NEUMONIA DEBIDA A STREPTOCOCCUS PNEUMONIAE
J40	BRONQUITIS, NO ESPECIFICADA COMO AGUDA O CRONICA
J14	NEUMONIA DEBIDA A HAEMOPHILUS INFLUENZAE
J929	PAQUIPLEURITIS SIN ASBESTOSIS
J320	SINUSITIS MAXILAR CRONICA
J61	NEUMOCONIOSIS DEBIDA AL ASBESTO Y A OTRAS FIBRAS MINERALES
J029	FARINGITIS AGUDA, NO ESPECIFICADA
J010	SINUSITIS MAXILAR AGUDA
J123	NEUMONIA DEBIDA A METANEUMOVIRUS HUMANO
J383	OTRAS ENFERMEDADES DE LAS CUERDAS VOCALES
J101	INFLUENZA CON OTRAS MANIFESTACIONES RESPIRATORIAS, DEBIDA A VIRUS DE LA INFLUENZA ESTACIONAL IDENTIFICADO
J682	INFLAMACION RESPIRATORIA SUPERIOR DEBIDA A INHALACION DE GASES, HUMOS, VAPORES Y SUSTANCIAS QUIMICAS, NO CLASIFICADAS EN OTRA PARTE
J987	Infecciones respiratorias, clasificadas en otra parte
J432	ENFISEMA CENTROLOBULAR
J108	INFLUENZA, CON OTRAS MANIFESTACIONES, DEBIDA A VIRUS DE LA INFLUENZA ESTACIONAL IDENTIFICADO
J00X	RINOFARINGITIS AGUDA [RESFRIADO COMUN]
J328	OTRAS SINUSITIS CRONICAS
J458	ASMA MIXTA
J323	SINUSITIS ESFENOIDAL CRONICA
J982	ENFISEMA INTERSTICIAL
J155	NEUMONIA DEBIDA A ESCHERICHIA COLI
J631	FIBROSIS (DEL PULMON) DEBIDA A BAUXITA

Bibliografía

- [Alcaldía de Bogotá, 2008] Alcaldía de Bogotá, . (2008). Boletín informativo localidad fontibón no. 34. bogotá d.c. Technical report, Secretaría de Salud de Bogotá y Hospital de Fontibón.
- [Alcaldía de Bogotá, 2019] Alcaldía de Bogotá, . (2019). Informe anual de calidad del aire de bogotá. Technical report, Secretaría Distrital de Ambiente de Bogotá.
- [Alcaldía de Bogotá, 2021] Alcaldía de Bogotá, . (2021). Plan estratégico para la gestión integral de la calidad del aire de bogotá 2030. bogotá d.c. Technical report, Secretaría Distrital de Ambiente de Bogotá.
- [Alcaldía de Bogotá, sf] Alcaldía de Bogotá, . (s.f.). Localidad de sumapaz. <https://bogota.gov.co/mi-ciudad/localidades/sumapaz>.
- [Anselin, 1998] Anselin, L. (1998). *Spatial Econometrics: Methods and Models*. Springer Science Business Media.
- [Barrios Casas et al., 2004] Barrios Casas, S., Peña Cortés, F., and Osses Bustingorry, S. (2004). Efectos de la contaminación atmosférica por material particulado en las enfermedades respiratorias agudas en menores de 5 años. *Revista Ciencia y Enfermería*, 10(2):21–29.
- [Becerra Moreno et al., 2021] Becerra Moreno, D., Ramirez Rios, L. F., Plaza Galvez, L. F., Oviedo Sanabria, C. H., and Niño Ovalles, M. V. (2021). Relación entre la calidad del aire y la incidencia de enfermedades respiratorias en las comunas 4, 6, 7 y 8 del municipio de cúcuta, norte de santander. *Revista Científica y Tecnológica: Ingeniería y Competitividad*, 23(2).
- [Blangiardo and Cameletti, 2015] Blangiardo, M. and Cameletti, M. (2015). Spatio-temporal models. In *Spatial and Spatio-temporal Bayesian Models with R-INLA*, pages 235–258. John Wiley Sons, Ltd, first edition.
- [Centro Iberoamericano de Desarrollo Estratégico Urbano CIDEU, sf] Centro Iberoamericano de Desarrollo Estratégico Urbano CIDEU, . (s.f.). Bogotá. <https://www.cideu.org/miembro/bogota/>.
- [Cressie, 1993] Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley Sons, Inc., United States of America, revised edition.

- [Cárdaba Arranz, 2012] Córdaba Arranz, M. (2012). Carga de enfermedad por contaminación atmosférica debida a materia particulada (pm₁₀ y pm_{2,5}) y ozono (o₃) en la ciudad de valladolid durante el periodo 1999-2008.
- [Datos Macro, sf] Datos Macro, . (s.f.). Tasa de mortalidad - mortalidad. <https://datosmacro.expansion.com/diccionario/tasa-de-mortalidad>.
- [Díaz, 2007] Díaz, M. A. B. (2007). Diseño de un modelo espacio-temporal basado en razonamiento espacial. *Instituto Politécnico Nacional*.
- [Echeverri Londoño and Maya Vasco, 2008] Echeverri Londoño, C. A. and Maya Vasco, G. J. (2008). Relación entre las partículas finas (pm_{2.5}) y respirables (pm₁₀) en la ciudad de medellín. *Revista Ingenierías Universidad De Medellín*, 7(12):23–42.
- [Gaviria G et al., 2011] Gaviria G, C. F., Benavides C, P. C., and Tangarife, C. A. (2011). Contaminación por material particulado (pm_{2,5} y pm₁₀) y consultas por enfermedades respiratorias en medellín (2008-2009). *Revista Facultad Nacional de Salud Pública*, 29(3):241–250.
- [Giraldo Henao, 2002] Giraldo Henao, R. (2002). *Introducción a la Geoestadística: Teoría y Aplicación*. Universidad Nacional de Colombia, Bogotá.
- [Gujarati and Porter, 2010] Gujarati, D. N. and Porter, D. C. (2010). *Econometría*. McGraw Hill, México, 5ta edition.
- [Gómez Rivadeneira, 2015] Gómez Rivadeneira, A. (2015). Clasificación internacional de enfermedades (cie): Descifrando la cie-10 y esperando la cie-11. *Revista Monitor Estratégico: Superintendencia Nacional de Salud*, (7):66–73.
- [Gómez Rubio, 2022] Gómez Rubio, V. (2022). Una introducción a la estadística espacial. *BEIO: Boletín de Estadística e Investigación Operativa*, 38(1).
- [Hernández Florez et al., 2013] Hernández Florez, L. J., Aristizábal Duque, G., Quiroz, L., Medina, K., Rodríguez Moreno, N., Sarmiento, R., and Osorio García, S. D. (2013). Contaminación del aire y enfermedad respiratoria en menores de cinco años de bogotá, 2007. *Revista de Salud Pública*, 15(4):503–516.
- [LeSage and Pace, 2009] LeSage, J. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Springer Science y Business Media, 1st edition.
- [Martinez Beneito and Botella Rocamora, 2019] Martinez Beneito, M. A. and Botella Rocamora, P. (2019). *Disease Mapping: From Foundations to Multidimensional Modeling*. CRC Press - Taylor y Francis Group, 1st edition.
- [Martínez L et al., 2011] Martínez L, E., Quiroz, C. M., and Rúa, J. A. (2011). Morbilidad respiratoria asociada con la exposición a material particulado en el ambiente. *Revista Facultad Nacional de Salud Pública*, 29(4):454–460.

- [Melo Martínez, 2012] Melo Martínez, C. E. (2012). *Análisis Geoestadístico Espacio Tiempo Basado en Distancias y Splines con Aplicaciones*. Tesis doctoral, Universitat de Barcelona, Barcelona.
- [Ministerio de Ambiente y Desarrollo Sostenible, 2017] 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.
- [Ministerio de Salud y Protección Social de Colombia, sf] Ministerio de Salud y Protección Social de Colombia, . (s.f.). Infecciones respiratorias agudas (ira). [https://www.minsalud.gov.co/salud/Paginas/Infecciones-Respiratorias-Agudas-\(IRA\).aspx](https://www.minsalud.gov.co/salud/Paginas/Infecciones-Respiratorias-Agudas-(IRA).aspx).
- [Moreno Serrano and Vayá Valcarce, 2000] Moreno Serrano, R. and Vayá Valcarce, E. (2000). *Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial*. Universidad de Barcelona.
- [Organización Mundial de la Salud, 2022] Organización Mundial de la Salud, O. (2022). Contaminación del aire ambiente (exterior). [https://www.who.int/es/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health).
- [Pebesma, 2012] Pebesma, E. (2012). Spacetime: Spatio-temporal data in r. *Journal of Statistical Software*, 51(7).
- [Pinzón Hassan and Tique Ortiz, 2019] Pinzón Hassan, A. D. and Tique Ortiz, V. (2019). Análisis de series de tiempo de contaminantes atmosféricos en la ciudad de bogotá a partir del desarrollo de modelos estadísticos arima.
- [QGIS Documentation, sf] QGIS Documentation, . (s.f.). Análisis espacial (interpolación). https://docs.qgis.org/3.34/es/docs/gentle_gis_introduction/spatial_analysis_interpolation.html.
- [Quintana Romero and Mendoza, 2016] Quintana Romero, L. and Mendoza, M. (2016). *Econometría aplicada utilizando R*. Universidad Nacional Autónoma de México, 1ra edition.
- [Quintero Forero, 2020] Quintero Forero, J. M. (2020). ¿cuál es la relación existente entre las concentraciones de material particulado y las admisiones a centros médicos por enfermedades respiratorias agudas en bogotá?
- [Rojas and Galvis, 2005] Rojas, N. and Galvis, B. (2005). Relación entre pm2.5 y pm10 en la ciudad de bogotá. *Revista De Ingeniería*, 1(22):54–60.
- [Sandoval Meneses, 2021] Sandoval Meneses, J. X. (2021). Evaluación de la concentración de material particulado (pm2.5) y su relación con enfermedades respiratorias en el municipio de sogamoso, boyacá-colombia.

- [Sanhueza H et al., 2006] Sanhueza H, P., Vargas R, C., and Mellado G, P. (2006). Impacto de la contaminación del aire por pm10 sobre la mortalidad diaria en temuco, chile. *Revista Médica de Chile*, 134(6):754–761.
- [Secretaría Distrital de Ambiente de Bogotá, sf] Secretaría Distrital de Ambiente de Bogotá, . (s.f.). Red de monitoreo de calidad del aire de bogotá - rmcab. <https://rmcab.ambientebogota.gov.co/>.
- [Secretaría Distrital de Salud de Bogotá, sfa] Secretaría Distrital de Salud de Bogotá, . (s.f.a). Observatorio de salud de bogotá - saludata. <https://saludata.saludcapital.gov.co/osb/>.
- [Secretaría Distrital de Salud de Bogotá, sfb] Secretaría Distrital de Salud de Bogotá, O. d. S. (s.f.b). Enfermedades respiratorias en niños menores de 5 años en bogotá d.c. <https://saludata.saludcapital.gov.co/osb/indicadores/enfermedades-respiratorias/>.
- [Verdoy, 2005] Verdoy, P. J. (2005). *Técnicas Geoestadísticas Espectrales. Análisis de la Estacionariedad e Independencia*. Tesis doctoral, Universitat Jaume I de Castellón, España.
- [Warrick et al., 1986] Warrick, A., Myers, D., and Nielsen, D. (1986). Geostatistical methods applied to soil science. w: Methods of soil analysis, part 1. physical and mineralogical methods. *Agronomy Monograph*, (9):53–81.
- [Xin, 2010] Xin, W. H. (2010). Survey on three-dimensional spatial data models in gis. pages 492–495.