



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

ANÁLISIS DESCRIPTIVO PARA DATOS FUNCIONALES DE LA CONTAMINACIÓN DEL AIRE EN EL CENTRO HISTÓRICO DE QUITO. PERÍODO 2019 – 2020.

FABIÁN ROBERTO ALLAUCA PANCHO

**Trabajo de Titulación modalidad: Proyecto de Investigación y Desarrollo,
presentado ante el Instituto de Posgrado y Educación Continua de la ESPOCH,
como requisito parcial para la obtención del grado de:**

**MAGÍSTER EN MATEMÁTICA MENCIÓN MODELACIÓN Y
DOCENCIA**

Riobamba-Ecuador

Diciembre – 2021

© 2021, Fabián Roberto Allauca Pancho

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
CERTIFICACIÓN

EL TRIBUNAL DEL TRABAJO DE TITULACION CERTIFICA QUE:

El trabajo de titulación modalidad Proyectos de Investigación y Desarrollo, denominado: Análisis descriptivo para datos funcionales de la contaminación del aire en el centro histórico de Quito. Período 2019 – 2020, de responsabilidad del señor Fabián Roberto Allauca Pancho, ha sido prolijamente revisado y se autoriza su presentación.

Tribunal:

Ing. Ángel Geovanny Guamán Lozano; Mag.

PRESIDENTE



Firmado electrónicamente por:
**MANUEL ANTONIO
MENESES FREIRE**

Dr. Manuel Antonio Meneses Freire; Ph. D.

DIRECTOR



Firmado electrónicamente por:
**KLEBER AUGUSTO
JARAMILLO
GALARZA**

Ing. Kleber Augusto Jaramillo Galarza; Mag.

MIEMBRO



Firmado electrónicamente por:
**MARCO MARCEL
PAREDES HERRERA**

Ing. Marco Marcel Paredes Herrera; Mag.

MIEMBRO

Riobamba, diciembre 2021

DERECHOS INTELECTUALES

Yo, Fabián Roberto Allauca Pancho soy responsable de las ideas, doctrinas y resultados expuestos en este **Trabajo de Titulación modalidad Proyectos de Investigación y Desarrollo** y que el patrimonio intelectual generado por la misma pertenece exclusivamente a la Escuela Superior Politécnica de Chimborazo.

FABIÁN ROBERTO ALLAUCA PANCHO

N° Cédula: 060344847-3

DECLARACIÓN DE AUTENTICIDAD

Yo, Fabián Roberto Allauca Pancho, declaro que el presente **Trabajo de Titulación modalidad Proyectos de Investigación y Desarrollo**, es de mi autoría y que los resultados del mismo son auténticos y originales. Los textos constantes en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor, asumo la responsabilidad legal y académica de los contenidos de este proyecto de investigación de Maestría.

FABIÁN ROBERTO ALLAUCA PANCHO

N° Cédula: 060344847-3

DEDICATORIA

A mi familia, en especial a mis padres por haber creído en mí y apoyarme en cada reto que la vida me ha presentado, siendo el pilar y motivación para seguir adelante y lograr cada sueño, en especial aquellos que parecen imposibles.

A mis amigos y a cada persona que colaboro de una u otra manera para que este trabajo tenga un contenido técnico, ético y lógico. Haciendo del mismo una base para futuras investigaciones y proyectos.

AGRADECIMIENTO

Quiero expresar mi más sincero agradecimiento a la Escuela Superior Politécnica de Chimborazo y al personal académico por su buena predisposición en enseñar, por tenernos paciencia y ayudarnos a crecer profesionalmente en el campo de la matemática.

A mi director y miembros, quienes me guiaron de la mejor manera en la realización del presente trabajo de investigación, así también quiero expresar mi agradecimiento a las coordinadoras del presente programa de maestría, quienes me apoyaron desde el primer día de clases, hasta la culminación del mismo.

A mi familia, en especial a mis padres que estuvieron a mi lado durante todo el proceso de este programa, apoyándome y animándome para que termine con éxito esta maestría; a mis amigos que me apoyaron con sus conocimientos y me ayudaron a expandir los míos.

TABLA DE CONTENIDOS

RESUMEN	xvii
ABSTRACT	xviii

CAPÍTULO I

1.	INTRODUCCIÓN.....	1
1.1	Planteamiento del problema	1
1.1.1	Situación problemática	1
1.1.2	Formulación del problema.....	2
1.1.3	Preguntas directrices o específicas	2
1.2	Justificación de la investigación	2
1.3	Objetivos	3
1.3.1	Objetivo general	3
1.3.2	Objetivos específicos.....	3

CAPÍTULO II

2.	MARCO TEÓRICO.....	4
2.1	Antecedentes investigativos	4
2.2	Fundamentación epistemológica	7
2.3	Bases teóricas	7
2.3.1	Contaminación del aire	7
2.3.2	Clasificación de la contaminación del aire.....	8
2.3.3	Contaminantes del aire registrados por la Red Metropolitana de Monitoreo Atmosférico de Quito.....	8

2.3.3.1	<i>Monóxido de carbono</i>	9
2.3.3.2	<i>Dióxido de azufre</i>	9
2.3.3.3	<i>Dióxido de nitrógeno</i>	10
2.3.3.4	<i>Ozono</i>	11
2.3.3.5	<i>Material particulado</i>	11
2.3.4	Instrumentos de medida	12
2.3.5	Ubicación de las estaciones de la Red Metropolitana de Monitoreo Atmosférico de Quito	13
2.3.6	Teoría del análisis de datos funcionales	14
2.3.6.1	<i>Espacios de Hilbert</i>	15
2.3.6.2	<i>Configuración del sistema de bases para la construcción de funciones</i>	17
2.3.6.3	<i>Suavizamiento mediante un operador diferencial lineal</i>	18
2.3.7	Herramientas del software libre R para análisis estadístico de datos	18
2.3.7.1	<i>Librerías para Análisis de Datos Funcionales (ADF)</i>	19
2.3.7.2	<i>Funciones usadas en el tratamiento de datos</i>	19
2.4	Hipótesis	19
2.5	Identificación de variables	20
2.5.1	Variable	20

CAPÍTULO III

3.	METODOLOGÍA DE INVESTIGACIÓN	21
3.1	Enfoque de investigación	21
3.2	Diseño de investigación	21
3.3	Tipo de investigación	22
3.4	Modalidad de investigación	22
3.5	Población y muestra	22
3.6	Operacionalización de las variables	24

3.7	Matriz de consistencia	25
3.8	Recolección de información	26
3.9	Procesamiento de la información	26

CAPÍTULO IV

4.	RESULTADOS Y DISCUSIÓN	27
4.1	Análisis general	27
4.1.1	Monóxido de carbono (CO)	27
4.1.1.1	<i>Detección de valores atípicos funcionales</i>	28
4.1.1.2	<i>Curvas medias mensuales</i>	29
4.1.1.3	<i>Curva media funcional con entorno de confianza bootstrap al 95%</i>	30
4.1.2	Dióxido de azufre (SO₂)	31
4.1.2.1	<i>Detección de valores atípicos funcionales</i>	32
4.1.2.2	<i>Curvas medias mensuales</i>	33
4.1.2.3	<i>Curva media funcional con entorno de confianza bootstrap al 95%</i>	35
4.1.3	Dióxido de nitrógeno (NO₂)	35
4.1.3.1	<i>Detección de valores atípicos funcionales</i>	36
4.1.3.2	<i>Curvas medias mensuales</i>	37
4.1.3.3	<i>Curva media funcional con entorno de confianza bootstrap al 95%</i>	39
4.1.4	Ozono (O₃)	39
4.1.4.1	<i>Detección de valores atípicos funcionales</i>	40
4.1.4.2	<i>Curvas medias mensuales</i>	41
4.1.4.3	<i>Curva media funcional con entorno de confianza bootstrap al 95%</i>	43
4.1.5	Material particulado fino (PM_{2.5})	43
4.1.5.1	<i>Detección de valores atípicos funcionales</i>	44
4.1.5.2	<i>Curvas medias mensuales</i>	45
4.1.5.3	<i>Curva media funcional con entorno de confianza bootstrap al 95%</i>	47

4.2	Análisis por períodos	47
4.2.1	Monóxido de carbono (CO)	48
4.2.1.1	<i>Períodos escolares y vacaciones escolares</i>	48
4.2.1.2	<i>Períodos pandemia y no pandemia</i>	49
4.2.1.3	<i>Períodos lunes-viernes y sábados-domingos</i>	50
4.2.2	Dióxido de azufre (SO₂)	50
4.2.1.1	<i>Períodos escolares y vacaciones escolares</i>	50
4.2.1.2	<i>Períodos pandemia y no pandemia</i>	51
4.2.1.3	<i>Períodos lunes-viernes y sábados-domingos</i>	52
4.2.3	Dióxido de nitrógeno (NO₂)	53
4.2.1.1	<i>Períodos escolares y vacaciones escolares</i>	53
4.2.1.2	<i>Períodos pandemia y no pandemia</i>	54
4.2.1.3	<i>Períodos lunes-viernes y sábados-domingos</i>	55
4.2.4	Ozono (O₃)	56
4.2.1.1	<i>Períodos escolares y vacaciones escolares</i>	56
4.2.1.2	<i>Períodos pandemia y no pandemia</i>	56
4.2.1.3	<i>Períodos lunes-viernes y sábados-domingos</i>	57
4.2.5	Material particulado fino (PM_{2.5})	58
4.2.1.1	<i>Períodos escolares y vacaciones escolares</i>	58
4.2.1.2	<i>Períodos pandemia y no pandemia</i>	59
4.2.1.3	<i>Períodos lunes-viernes y sábados-domingos</i>	59

CAPÍTULO V

5.	PROPUESTA	61
5.1	Ajuste de modelos lineales funcionales con respuesta escalar de la variable: media diaria del material particulado fino (variable escalar “y”), en función de las	

	variables funcionales: monóxido de carbono, dióxido de azufre y dióxido de nitrógeno.....	61
5.1.1	Modelo lineal de la variable escalar “y” en función de la variable funcional “monóxido de carbono”	61
5.1.2	Modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de azufre”	63
5.1.3	Modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de nitrógeno”	64
5.1.4	Modelo lineal múltiple de la variable escalar “y” en función de las variables funcionales: monóxido de carbono, dióxido de azufre y dióxido de nitrógeno	65
5.2	<i>Ajuste del modelo lineal funcional con respuesta escalar de la variable: media diaria del ozono (variable escalar “y”), en función de la variable funcional: dióxido de nitrógeno.</i>	66
	CONCLUSIONES.....	67
	RECOMENDACIONES.....	69
	BIBLIOGRAFÍA	

ÍNDICE DE TABLAS

Tabla 1-2:	Detalle de los analizadores de gases y partículas de la REMMAQ	13
Tabla 2-2:	Ubicación de las estaciones de la REMMAQ	14
Tabla 3-2:	Formulación de hipótesis en estudios cuantitativos con diferentes alcances	20
Tabla 4-3:	Operacionalización de variables	24
Tabla 5-3:	Matriz de consistencia	25
Tabla 1-5:	Resultados del modelo lineal de la variable escalar “y” en función de la variable funcional “monóxido de carbono”	62
Tabla 2-5:	Resultados del modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de azufre”	63
Tabla 3-5:	Resultados del modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de nitrógeno”	64
Tabla 4-5:	Resultados del modelo lineal múltiple de la variable escalar “y” en función de las variables funcionales: monóxido de carbono, dióxido de azufre y dióxido de nitrógeno	65
Tabla 5-5:	Resultados del modelo lineal de la variable escalar “media diaria del ozono” en función de la variable funcional “dióxido de nitrógeno”	66

ÍNDICE DE FIGURAS

Figura 1-4:	Curvas diarias del monóxido de carbono para los años 2019 y 2020.....	27
Figura 2-4:	Funciones atípicas del monóxido de carbono para los años 2019 y 2020	28
Figura 3-4:	Gráfica exploratoria del monóxido de carbono sin curvas atípicas	29
Figura 4-4:	Curvas medias mensuales del monóxido de carbono para el año 2019.....	29
Figura 5-4:	Curvas medias mensuales del monóxido de carbono para el año 2020.....	30
Figura 6-4:	Estimador de curva media con entorno de confianza bootstrap al 95% del monóxido de carbono para los años 2019 y 2020.	31
Figura 7-4:	Curvas diarias del dióxido de azufre para los años 2019 y 2020	31
Figura 8-4:	Funciones atípicas del dióxido de azufre para los años 2019 y 2020.....	32
Figura 9-4:	Gráfica exploratoria del dióxido de azufre sin curvas atípicas.....	33
Figura 10-4:	Curvas medias mensuales del dióxido de azufre para el año 2019	34
Figura 11-4:	Curvas medias mensuales del dióxido de azufre para el año 2020	34
Figura 12-4:	Estimador de curva media con entorno de confianza bootstrap al 95% del dióxido de azufre para los años 2019 y 2020.	35
Figura 13-4:	Curvas diarias del dióxido de nitrógeno para los años 2019 y 2020	36
Figura 14-4:	Funciones atípicas del dióxido de nitrógeno para los años 2019 y 2020.....	36
Figura 15-4:	Gráfica exploratoria del dióxido de nitrógeno sin curvas atípicas	37
Figura 16-4:	Curvas medias mensuales del dióxido de nitrógeno para el año 2019	38
Figura 17-4:	Curvas medias mensuales del dióxido de nitrógeno para el año 2020	38
Figura 18-4:	Estimador de curva media con entorno de confianza bootstrap al 95% del dióxido de nitrógeno para los años 2019 y 2020.....	39
Figura 19-4:	Curvas diarias del ozono para los años 2019 y 2020	40
Figura 20-4:	Funciones atípicas del ozono para los años 2019 y 2020.....	40
Figura 21-4:	Gráfica exploratoria del ozono sin curvas atípicas	41
Figura 22-4:	Curvas medias mensuales del ozono para el año 2019	42
Figura 23-4:	Curvas medias mensuales del ozono para el año 2020	42

Figura 24-4:	Estimador de curva media con entorno de confianza bootstrap al 95% del ozono para los años 2019 y 2020.	43
Figura 25-4:	Curvas diarias del material particulado fino para los años 2019 y 2020	44
Figura 26-4:	Funciones atípicas del material particulado fino para los años 2019 y 2020.....	44
Figura 27-4:	Gráfica exploratoria del material particulado fino sin curvas atípicas	45
Figura 28-4:	Curvas medias mensuales del material particulado fino para el año 2019	46
Figura 29-4:	Curvas medias mensuales del material particulado fino para el año 2020	46
Figura 30-4:	Estimador de curva media con entorno de confianza bootstrap al 95% del material particulado fino para los años 2019 y 2020.....	47
Figura 31-4:	Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el monóxido de carbono	48
Figura 32-4:	Comparación de medias funcionales entre los periodos pandemia y no pandemia para el monóxido de carbono	49
Figura 33-4:	Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el monóxido de carbono	50
Figura 34-4:	Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el dióxido de azufre	51
Figura 35-4:	Comparación de medias funcionales entre los periodos pandemia y no pandemia para el dióxido de azufre.....	52
Figura 36-4:	Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el dióxido de azufre	53
Figura 37-4:	Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el dióxido de nitrógeno	53
Figura 38-4:	Comparación de medias funcionales entre los periodos pandemia y no pandemia para el dióxido de nitrógeno	54
Figura 39-4:	Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el dióxido de nitrógeno.....	55
Figura 40-4:	Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el ozono	56
Figura 41-4:	Comparación de medias funcionales entre los periodos pandemia y no pandemia para el ozono	57

Figura 42-4:	Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el ozono	57
Figura 43-4:	Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el material particulado fino	58
Figura 44-4:	Comparación de medias funcionales entre los periodos pandemia y no pandemia para el material particulado fino	59
Figura 45-4:	Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el material particulado fino.....	60
Figura 1-5:	Curvas diarias del monóxido de carbono para los años 2019 y 2020 (izquierda) y su representación en bases de Fourier (derecha)	62
Figura 2-5:	Curvas diarias del dióxido de azufre para los años 2019 y 2020 (izquierda) y su representación en bases de Fourier (derecha)	63
Figura 3-5:	Curvas diarias del dióxido de nitrógeno para los años 2019 y 2020 (izquierda) su representación en bases de Fourier (derecha)	64

ABSTRACT

This research study aims to descriptively analyze the air pollution functional data in the historic center of Quito, for the years 2019 – 2020. For which a statistical software R, with the package "fda.usc" was used. The pollutants analyzed were: carbon monoxide, sulfur dioxide, nitrogen dioxide, ozone, and fine particulate matter. The analysis started by downloading the pollutant database from the Secretaría de Ambiente de Quito. Then a selection, ordering, and filling of missing data was carried out, by linear interpolation. The reading of the database was carried out with the function read.table and the transformation to functional data by the fdata function. In the general analysis, exploratory graphs were made, obtaining atypical curves, monthly mean curves with bands of confidence. In the analysis by periods, the mean curves for the class school periods were compared with school holidays, the pandemic period and no pandemic, and periods from Monday to Friday, and Saturday to Sunday. To finally present linear models with a scalar response. The results obtained from the study indicated that the behavior of the pollutants is irregular throughout the day. From the analysis by periods, it was possible to conclude that the functional mean is higher in the school classes period, not pandemic period, and Monday to Friday. Linear models with a scalar response of the variable: the daily mean of fine particulate matter, resulted in the variability of 31.61%, 35.08% and for carbon monoxide, sulfur dioxide, and nitrogen dioxide a 57.45% respectively. While the percentage of variability of the daily mean for the fine particulate matter explained by the variables of the three pollutants showed a percentage of 60.83%.

Keywords:

<DESCRIPTIVE ANALYSIS>, <FUNCTIONAL DATA>, <AIR POLLUTION>, <STATISTICAL SOFTWARE R>, <FDA.USC PACKAGE>.

CAPÍTULO I

1. INTRODUCCIÓN

1.1 Planteamiento del problema

1.1.1 Situación problemática

La contaminación del aire es una mezcla de partículas sólidas y gases en el aire que respiramos. Las principales causas se relacionan con la quema de combustibles fósiles como el carbón, el petróleo y el gas cuyo origen principal se encuentra en el sector industrial, la extracción de pozos petrolíferos y el transporte por carretera. (Fundación Aquae, 2021)

Para monitorear los contaminantes presentes en el aire el Municipio del Distrito Metropolitano de Quito crea la Red Metropolitana de Monitoreo Atmosférico (REMMAQ) con el préstamo 822/OC-EC del Banco Interamericano de Desarrollo otorgado al Gobierno de la República del Ecuador el 13 de octubre de 1994. (Municipio del Distrito Metropolitano de Quito, 2021)

La puesta en marcha definitiva de la Red estuvo a cargo de una unidad especial creada por la Empresa de Desarrollo del Centro Histórico (ECH) el 5 de noviembre del 2002. Desde el 1 de enero del 2004 hasta 31 de octubre del 2010, la REMMAQ formó parte de la Corporación para el Mejoramiento del Aire en Quito (CORPAIRE). Actualmente la REMMAQ formar parte de la Secretaría de Ambiente del Municipio de Quito desde el 01 de noviembre del 2010. (Municipio del Distrito Metropolitano de Quito, 2021)

La finalidad de la REMMAQ es producir datos confiables sobre la concentración de contaminantes atmosféricos como es el caso del monóxido de carbono (CO), dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂), Ozono (O₃) y material particulado fino o de diámetro menor a 2,5 micrómetros (PM_{2,5}); en el territorio del Distrito Metropolitano de Quito, datos disponibles al público en general desde su página web www.quitoambiente.gob.ec apartado Red de Monitoreo Atmosférico, descarga datos históricos.

El problema es que la REMMAQ no analiza los datos obtenidos por las distintas estaciones de monitoreo, por tal motivo existe un desconocimiento del comportamiento de los contaminantes presentes en el aire en el centro histórico de Quito en las distintas horas del día, esto se puede deber principalmente a la falta de interés y de presupuesto por parte del Municipio del Distrito

Metropolitano de Quito, como también a la carencia de estudios por parte de la comunidad científica, así también como al desconocimiento de herramientas para análisis funcional de datos.

El desconocimiento del comportamiento de los contaminantes presentes en el aire trae consigo efectos perjudiciales a la salud de la población, principalmente en enfermedades pulmonares, puesto que al no saber en qué horas se presentan las máximas concentraciones de contaminantes, la población circula libremente a cualquier hora del día.

1.1.2 Formulación del problema

¿Cuál es el análisis descriptivo para datos funcionales de la contaminación del aire en el centro histórico de Quito, periodo 2019 - 2020?

1.1.3 Preguntas directrices o específicas

¿Cómo se examinan los datos de la contaminación del aire para su tratamiento en el software estadístico R?

¿Cómo se describe el comportamiento de los datos funcionales de la contaminación del aire en el software estadístico R?

¿Cuál es el modelo matemático para datos funcionales de la contaminación del aire registrados en el centro histórico de Quito en los años 2019 y 2020?

1.2 Justificación de la investigación

La presente investigación tendrá **interés** en analizar descriptivamente los datos funcionales de la contaminación del aire en el centro histórico de Quito para los años 2019 y 2020; como también proponer un modelo matemático que represente a estos datos.

La **importancia** radica en determinar los datos atípicos, la media, la desviación, así también como la banda de confianza por medio del método de análisis de datos funcionales (ADF), con la finalidad de establecer un modelo matemático que represente a los datos funcionales de la contaminación del aire en el centro histórico de Quito para los años antes mencionados.

Existe la **factibilidad** para realizar la presente investigación porque se dispone de los conocimientos suficientes por parte del investigador, la materia de Análisis Funcional dictada en la presente maestría abala tales conocimientos; facilidad para acceder a la información que respaldará esta investigación, la misma que fue tomada de la página web de la Secretaría de

Ambiente de Quito, www.quitoambiente.gob.ec apartado Red de Monitoreo Atmosférico, descarga datos históricos; bibliografía especializada; recursos tecnológicos (software estadístico R), económicos y el tiempo previsto para culminar con el trabajo de posgrado.

El trabajo de investigación tiene **utilidad práctica** porque se pone en ejecución los conceptos de análisis de datos funcionales, dictados en el módulo de Análisis Funcional del presente programa de maestría, dando solución al problema de desconocimiento del comportamiento de los contaminantes del aire que monitorea la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ).

Los **beneficiarios** de la investigación son las personas que viven y visitan el centro histórico de Quito, así también el Municipio del Distrito Metropolitano; puesto que conocerán el comportamiento de los datos de la contaminación del aire y de esta manera tomar las acciones pertinentes.

1.3 Objetivos

1.3.1 Objetivo general

Analizar descriptivamente los datos funcionales de la contaminación del aire en el centro histórico de Quito. Periodo 2019 – 2020.

1.3.2 Objetivos específicos

- Examinar los datos de la contaminación del aire para su tratamiento en el software estadístico R.
- Describir el comportamiento de los datos funcionales de la contaminación del aire en el software estadístico R.
- Plantear un modelo matemático para datos funcionales de la contaminación del aire registrados en el centro histórico de Quito en los años 2019 y 2020.

CAPÍTULO II

2. MARCO TEÓRICO

2.1 Antecedentes investigativos

La tesis de grado presentada como requisito para la obtención del título de Magíster en Matemáticas Aplicadas titulada “Análisis de datos meteorológicos del Valle de los Chillos usando Datos Funcionales” de la Universidad San Francisco de Quito, muestra el uso del análisis de datos funcionales (ADF) para modelar el comportamiento de las variables atmosféricas en este caso la temperatura registrada por la estación meteorológica de la hacienda El Prado ubicada en el valle de los Chillos; para este análisis el autor utiliza software estadístico de libre distribución R, con énfasis en el uso de las librerías `fda` y `fda.usc`; dentro de las recomendaciones el autor menciona que sería conveniente desarrollar un curso de Análisis de Datos Funcionales en el país, ya que no existen docentes suficientes que conozcan el tema. (Guerrón Varela, 2015)

La tesis antes mencionada aplica el Análisis de Datos Funcionales (ADF o FDA), que es una parte de la Estadística relativamente nueva, desarrollada por los profesores Jim O. Ramsey y Bernard W. Silverman en el año de 1997, por ser una disciplina nueva el autor recomienda que se dicten cursos en el país puesto que su aplicación es amplia como en medicina (cardiología), finanzas, economía, meteorología entre otras. (Guerrón Varela, 2015)

Dentro de la provincia de Chimborazo se encontró la tesis de grado presentada para obtener el título de magíster en Matemática Básica, titulada “Análisis exploratorio funcional de los datos de radiación solar 2011 como soporte para la simulación de su comportamiento” de la Escuela Superior Politécnica de Chimborazo, en donde se describe el comportamiento de la radiación solar global (RSG) registrada en la Estación Meteorológica ESPOCH, durante el 2011, fenómeno tratable con el Método Análisis de Datos Funcionales (ADF), la autora selecciona este análisis por su confiabilidad en resultados, aplicando software libre R con su librería `fda.usc`, las conclusiones a las que llega es a la detección de funciones atípicas producidas los días 14 de febrero, 26 de febrero, 8 de abril y 20 de octubre de 2011; la media funcional muestra un comportamiento acampanado para la estación de verano (junio-septiembre) y semiacampanado para el invierno (octubre-mayo), los meses de marzo, abril y mayo muestran comportamiento igual a los de verano, la autora concluye a que esto se debe por que el país no posee estaciones claramente definidas. (Escudero Villa, 2014)

La autora en esta tesis recalca la importancia del Análisis de Datos Funcionales (ADF) para manejo de grandes cantidades de datos, encontrando aplicaciones en el área de la economía (cotizaciones de bolsas), ingeniería en producción (demanda de energía eléctrica), medio ambiente (vigilancia atmosférica, fluvial o meteorológica), entre otras; como se puede observar cuando se tiene volúmenes altos de datos, lo más recomendable es tratarlos con estadística funcional L^2 , en donde se toma los datos discretizados y se transforma en funciones. (Escudero Villa, 2014)

Revisando la literatura internacional, encontramos un estudio realizado en España por la Universidad de Cádiz, departamento de Estadística e Investigación Operativa, titulado “Estudio y aplicación de técnicas de Análisis de Datos Funcionales de geoposicionamiento”, donde el objetivo general es aplicar métodos del Análisis de Datos Funcionales al estudio de problemas de geoposicionamiento; el estudio tiene un enfoque de carácter metodológico. El estudio concluye que dada la naturaleza de los datos procedentes de observaciones GPS, el marco de análisis que proporciona el FDA es muy útil para el tratamiento de problemas del ámbito de la Geodesia. Cabe recalcar que el estudio es una adaptación del artículo publicado en el año 2018 en la revista “Mathematical Geosciences” con el título “Analysis of a GPS network based on Functional Data Analysis”. Que es uno de los primeros trabajos publicados donde se tratan datos de carácter geodésico usando métodos FDA. (Pérez Plaza, 2020)

Podemos observar que el Análisis de Datos Funcionales (ADF o FDA) puede ser aplicable a variables de geoposicionamiento como muestra el estudio antes mencionado, como así también el estudio “Analysis of a GPS network based on Functional Data Analysis”, mostrando resultados satisfactorios puesto que pudieron analizar problemas del ámbito de la Geodesia, trayendo consigo una aplicación más del ADF.

La tesis “Análisis de Datos Funcionales aplicado a datos de temperatura en España” realizado por la Universidad de Valladolid, Facultad de Ciencias, estudia una serie de procedimientos relacionados con el Análisis de Datos Funcionales, que van desde la transformación de conjuntos de valores discretos en observaciones funcionales a la exploración de sus características. Los datos estudiados son valores diarios de temperatura medidos por distintas estaciones meteorológicas en España en 2013, con el objetivo de extraer información sobre las características de los datos funcionales obtenidos. (Picón Llamas, 2019)

Para este estudio el autor menciona que el Análisis de Datos Funcional son técnicas estadísticas específicamente diseñadas para tratar conjuntos de datos creados mediante la observación de funciones o curvas. Estas técnicas nacen como una alternativa más eficiente al tratamiento

estadístico multivariante de dichas funciones a partir de simples valores puntuales que toman estas funciones en momentos determinados. De esta forma, al cambiar del enfoque multivariante al funcional, se pueden extraer y estudiar sus características mejor. (Picón Llamas, 2019)

La tesis de master titulada “Métodos Avanzados de Análisis de Datos Funcionales” realizada en la Universidad de Cádiz, Facultad de Ciencias, para la obtención del Máster en Matemática; muestra una revisión de las principales técnicas relacionadas con el análisis y tratamiento estadístico de datos funcionales. El estudio comienza con el estado actual de la investigación en el campo de la estadística funcional, retos, desafíos y nuevos enfoques. Dentro de los principales métodos avanzados, se hace hincapié en las herramientas de reducción de la dimensión, especialmente en el análisis de componentes principales funcionales (ACPF) y en la reconstrucción de curvas mediante modelos lineales. A continuación, se desarrolla distintas técnicas de detección de outliers funcionales, para seguidamente estudiar y profundizar en varios métodos de clasificación de curvas o "análisis clúster funcional", en los que tiene gran protagonismo el algoritmo de K-medias. Todas y cada una de los conceptos vienen acompañados de ejemplos sobre conjuntos de datos reales, con el fin de evidenciar sus puntos fuertes y limitaciones. Para ello se emplea distintos paquetes del software R cuyo código puede consultarse en su apéndice. (Pérez Montilla, 2018)

No solo la aplicación del Análisis de Datos Funcionales es temática de estudio dentro de la comunidad científica, sino también la metodología de este análisis, como lo podemos comprobar con el estudio antes citado. El estudio abarca distintas técnicas de detección de atípicos, así también como el de clasificación de curvas, lo novedoso y útil es que todos los conceptos vienen acompañados de sus ejemplos empleando el software R cuya codificación se la puede encontrar en el apéndice de este estudio. La información hallada en este estudio será de gran importancia para el presente trabajo.

De la bibliografía consultada podemos concluir que el Análisis de Datos Funcionales (ADF o FDA) es ampliamente utilizado para describir comportamientos de distintas variables, donde se tenga como característica principal el volumen elevado de datos, las aplicaciones pueden ser tan variadas como estrellas en el universo, yendo desde la medicina hasta la economía, pasando por la meteorología, producción industrial entre tantas muchas más. En la contaminación del aire se tiene gran cantidad de datos, dependiendo de las estaciones de monitoreo, de los contaminantes que registren y del intervalo de tiempo de cada registro, se puede tener millones de datos a ser analizados. El Análisis de Datos Funcionales responde claramente a este inconveniente como se pudo observar en esta investigación bibliográfica.

2.2 Fundamentación epistemológica

El investigador para realizar el trabajo de grado se acoge a la **investigación científica**, puesto que es un proceso en el que el investigador se plantea algunas preguntas y obtiene conocimiento de la realidad, para el presente estudio se pretende conocer el comportamiento de los contaminantes del aire que monitorea la Red Metropolitana de Monitoreo Atmosférico de Quito en el centro histórico en las distintas horas del día para los años 2019 y 2020, es decir obtener conocimiento acerca de esta realidad. (Arias Odón, 2012)

El paradigma de investigación escogido para el presente estudio será el **paradigma positivista** y la metodología cuantitativa, puesto que sus características se acoplan de mejor manera al desarrollo del presente trabajo de investigación; siendo estas características las siguientes; la realidad es lo observable, lo medible; separación entre el sujeto y el objeto; prevalece el método hipotético-deductivo; y finalmente, el investigador busca prescindir de sus valores para que no interfiera en la investigación. (Ricoy Lorenzo, 2006)

2.3 Bases teóricas

2.3.1 Contaminación del aire

La contaminación del aire es actualmente uno de los problemas ambientales más severos a nivel mundial. Está presente en todas las sociedades, independientemente del nivel de desarrollo socioeconómico, y constituye un fenómeno que tiene particular incidencia sobre la salud del hombre. (Romero Placeres, Diego Olite, & Álvarez Toste, 2006)

La contaminación puede definirse como cualquier modificación indeseable del ambiente, causada por la introducción a este de agentes físicos, químicos o biológicos (contaminantes) en cantidades superiores a las naturales, que resulta nociva para la salud humana, daña los recursos naturales o altera el equilibrio ecológico. (Yassi, Kjellstrom, de Kok, & Guidotti, 2002)

La contaminación atmosférica o contaminación del aire es, por consiguiente, una de las formas principales en que puede ser degradado o afectado parte del ambiente. Yassi A., la describe como “la emisión al aire de sustancias peligrosas a una tasa que excede la capacidad de los procesos naturales de la atmósfera para transformarlos, precipitarlos y depositarlos o diluirlos por medio del viento y el movimiento del aire”. (Yassi, Kjellstrom, de Kok, & Guidotti, 2002)

2.3.2 Clasificación de la contaminación del aire

Existen muchas clasificaciones, entre las principales podemos mencionar;

- Según su origen: puede ser clasificada por causas naturales o antropogénicas. Las naturales siempre han existido, mientras que las antropogénicas, como su nombre lo indica, son causadas por las actividades humanas. (Gutiérrez, Romieu, Corey, & Fortoul, 1997)
- Según su fuente: pueden clasificarse en; fuentes naturales (polvo que contiene materias biológicas, esporas, polen y bacterias); fuentes agrícolas (insecticidas y herbicidas empleados en la agricultura); fuentes tecnológicas (procesos industriales de todo tipo, consumo industrial y doméstico de combustibles fósiles, vehículos de motor). (Romero Placeres, Diego Olite, & Álvarez Toste, 2006)
- Según su forma física: gases y aerosoles (líquidos y sólidos). (Romero Placeres, Diego Olite, & Álvarez Toste, 2006)
- Según su origen: primarios (aquellos emitidos a la atmósfera como resultado de un proceso natural o antropogénico, como por ejemplo: dióxido de azufre, monóxido de carbono) y secundarios (se forman en la atmósfera producto de alguna reacción; por ejemplo: sulfatos y ozono). (Meneses, Turtós, & Molina, 2004)

2.3.3 Contaminantes del aire registrados por la Red Metropolitana de Monitoreo Atmosférico de Quito

La Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ) tiene como finalidad producir datos confiables sobre la concentración de contaminantes atmosféricos en el territorio del Distrito Metropolitano de Quito que sirvan como insumo para la planificación, formulación, ejecución y evaluación de políticas y acciones orientadas al mejoramiento de la calidad del aire y difundir esta información en condiciones comprensibles para el público en general. (Municipio del Distrito Metropolitano de Quito, 2021)

La REMMAQ incluye nueve estaciones remotas de monitoreo con capacidad para analizar continua y automáticamente los siguientes contaminantes comunes del aire:

- Monóxido de carbono (CO);
- Dióxido de azufre (SO₂);
- Óxidos de nitrógeno (NO, NO₂ y NO_x);
- Ozono (O₃); y,

- Material particulado fino o de diámetro menor a 2.5 micrómetros (PM2.5) y menor a 10 micrómetros (PM10)

2.3.3.1 Monóxido de carbono

El monóxido de carbono es un gas incoloro e inodoro que se forma por la combustión incompleta de material orgánico, en presencia deficitaria de oxígeno. Es considerado uno de los mayores contaminantes de la atmósfera terrestre, y uno de los mayores problemas ambientales de América Latina (Organización Panamericana de la Salud, 2005). Las principales fuentes productoras de este contaminante son los vehículos automotores que utilizan como combustible gasolina o diésel; los procesos industriales; los incendios forestales y urbanos y la incineración de materia orgánica. Los vehículos automotores y los procesos industriales son responsables de aproximadamente 80 % de las emisiones de monóxido de carbono a la atmósfera (Córdoba & Ramos, 2001).

De acuerdo con la Environmental Protection Agency, en los Estados Unidos las emisiones de monóxido de carbono al ambiente son producidas en un 95 % por los vehículos automotores (U.S Environmental Protection Agency, 2001). La corporación para el mejoramiento del aire de Quito (CORPAIRE) informó que el monóxido de carbono es el contaminante más abundante en peso, con 94 593 toneladas. En cuanto a su origen, el 97% del CO es aportado por las fuentes móviles, principalmente particulares livianos (43%), camionetas (26%) y taxis a gasolina (12%). (Corporación para el Mejoramiento del Aire de Quito, 2006)

2.3.3.2 Dióxido de azufre

El dióxido de azufre es el principal causante de la lluvia ácida ya que en la atmósfera es transformado en ácido sulfúrico. Es liberado en muchos procesos de combustión ya que los combustibles como el carbón, el petróleo, el diésel o el gas natural contienen ciertas cantidades de compuestos azufrados. Por estas razones se intenta eliminar estos compuestos antes de su combustión por ejemplo mediante la hidrodesulfuración en los derivados del petróleo o con lavados del gas natural haciéndolo más dulce (Ministerio de Ambiente de Colombia , 2021).

También los procesos metalúrgicos liberan ciertas cantidades de este gas debido a que se emplean frecuentemente los metales en forma de sulfuros. En la naturaleza el dióxido de azufre se encuentra sobre todo en las proximidades de los volcanes y las erupciones pueden liberar cantidades importantes. El dióxido de azufre se utiliza para fines muy diversos, por ejemplo, como agente reductor en metalurgia, como frigorigeno en la industria del frío, como desinfectante y blanqueador, para la conservación de sustancias alimenticias, como decolorante y fumigante. El

dióxido de azufre es uno de los compuestos más importantes de la industria química. 98% del SO₂ técnico se utiliza para la producción de trióxido de azufre como precursor del ácido sulfúrico (LUMITOS AG, 2021).

En el inventario de emisiones del Distrito Metropolitano de Quito 2003 indica que se generaron 6 009 toneladas anuales de dióxido de azufre, el 56% de las cuales tuvieron su origen en las fuentes móviles, principalmente a diésel (40%); las fuentes fijas contribuyeron con el 44%, la mayor parte de lo cual (25,7%) corresponde a las centrales de generación termoeléctrica (Corporación para el Mejoramiento del Aire de Quito, 2006).

2.3.3.3 Dióxido de nitrógeno

El dióxido de nitrógeno NO₂ es un compuesto químico gaseoso de color marrón amarillento formado por la combinación de un átomo de nitrógeno y dos de oxígeno. Es un gas tóxico e irritante. El NO₂ junto al NO-óxido nitroso- son conocidos como NO_x y son algunos de los principales contaminantes en las ciudades (Instituto para la Salud Geoambiental, 2021).

En la naturaleza se produce por los incendios forestales o las erupciones volcánicas. También se produce de forma natural por la descomposición de nitratos orgánicos. El volumen total que se produce de forma natural es infinitamente menor que el que se produce por efecto del hombre (Gobierno de España, 2021).

La mayor parte tiene su origen en la oxidación del NO que se produce en la combustión de los motores de los vehículos, fundamentalmente los diésel. El NO emitido por los motores, una vez en la atmosfera, se oxida y se convierte en NO₂. Es también un potenciador del material particulado, sobre todo de partículas finas PM 2,5 que son las más perjudiciales. En su reacción con la luz UV del sol es un precursor de O₃ ozono troposférico (Gobierno de España, 2021).

La exposición continuada NO₂ se relaciona con diversas enfermedades de las vía respiratorias como disminución de la capacidad pulmonar, bronquitis agudas, asma y se considera el culpable de los procesos alérgicos, sobre todo en niños. Se ha relacionado las exposiciones crónicas a bajo nivel con el enfisema pulmonar. Otros efectos menores son la irritación ocular y de las mucosas. Existen algunos estudios que apuntan a un incremento en la mortalidad aunque aún no está suficientemente bien establecido que la causa sea solamente la exposición a NO₂ (Instituto para la Salud Geoambiental, 2021).

El Municipio del Distrito Metropolitano de Quito por medio de la Corporación para el Mejoramiento del Aire de Quito (CORPAIRE), menciona que se generaron 21 699 toneladas de NO_x , 80% de las cuales por fuentes móviles, principalmente buses y pesados (42%) y particulares livianos, taxis y camionetas (35%). Otra fuente significativa son las centrales de generación termoeléctrica (17%). (Corporación para el Mejoramiento del Aire de Quito, 2006)

2.3.3.4 Ozono

El ozono (O_3) es una sustancia cuya molécula está compuesta por tres átomos de oxígeno, formada al disociarse los dos átomos que componen el gas oxígeno. Cada átomo de oxígeno liberado se une a otra molécula de oxígeno gaseoso (O_2), formando moléculas de ozono (O_3). (Torres, 2002)

A temperatura y presión ambientales, el ozono es un gas que desprende olores fuertes (similar al de los mariscos en estado de descomposición avanzado) y generalmente sin coloración, pero en grandes concentraciones puede volverse ligeramente azulado. Si se respira en grandes cantidades puede provocar una irritación en los ojos o la garganta, la cual suele pasar después de respirar aire fresco y rico en oxígeno durante algunos minutos. (Sanchiz, 2014)

El límite recomendado de exposición de ozono es de 0,1 partes por millón (ppm), o sea 0,2 miligramos por metro cúbico, calculado como una concentración ponderada de ocho horas, y a corto plazo de 0,3 ppm (0,6 mg. por metro cúbico), como una concentración ponderada de quince minutos. En función de la concentración puede producir diversos efectos en el organismo. Incluso en concentraciones muy bajas, el ozono puede ser nocivo para el tracto respiratorio superior y los pulmones, aun tratándose de una exposición de corta duración. A concentraciones extremadamente bajas, el ozono puede producir hipersensibilidad bronquial y respuesta inflamatoria en el tejido respiratorio. Además, la exposición intermitente puede causar una inflamación en bronquios y pulmones. (Espín Mayorga & Veloz Cuichán, 2013)

2.3.3.5 Material particulado

El material particulado (PM) por sus siglas en inglés es un contaminante atmosférico de tipo partículas suspendidas, estas partículas suspendidas en las atmósferas se denomina “respirables” a las de un diámetro menor o igual a $10\ \mu\text{m}$ (PM10) por su capacidad de introducirse en las vías respiratorias (Salini Calderón, 2014).

Los principales componentes del material particulado son hollín, plomo, sulfato e hidrocarburos. Además, la contaminación aérea por material particulado afecta negativamente la calidad del aire

y la salud de las personas, animales y bienes materiales. La cantidad que se transporta por material particulado puede representar niveles biológicamente significativos y su presencia en el material particulado fino, si es inhalado, puede depositarse profundamente en los pulmones en los seres vivos. (Foster & Costa, 2005)

La exposición a niveles elevados de PM se ha asociado al impacto adverso a la salud humana, y es la mayor fuente de enfermedades y mortalidad alrededor del mundo. El PM con un diámetro igual o menor a 2,5 μm (PM2.5) es el más dañino para la salud humana, porque puede llegar fácilmente hasta los pulmones. Esta última proviene de la combustión residencial, los procesos industriales y los vehículos gasolineros y Diesel (pesados y livianos). La fracción gruesa proviene principalmente del polvo en suspensión generado por el tráfico, la construcción y quemas agrícolas y de leña. (Salini Calderón, 2014)

El Municipio del Distrito Metropolitano de Quito por medio de la Corporación para el Mejoramiento del Aire de Quito (CORPAIRE), menciona que se generaron 4 904 toneladas de PM10, el 56% de las cuales es producida por las fuentes de área, principalmente canteras, resuspensión de vías no asfaltadas, incendios y quemas y erosión eólica del suelo. Adicionalmente, las fuentes móviles generan el 32%, en su mayor parte producidas por los buses y vehículos pesados a diésel (20%). Las fuentes fijas generan el 12%; la mayor parte tiene su origen en procesos industriales de producción de agregados, cemento y pinturas (8%). (Corporación para el Mejoramiento del Aire de Quito, 2006)

Del mismo modo menciona que de las 2 409 toneladas de material particulado fino (PM2,5) producido en el DMQ en el año 2003, el 46% es producido por las fuentes móviles, principalmente a diésel (38%). Las fuentes de área aportan con el 40%, generado, al igual que el PM10, por las canteras, resuspensión de vías, incendios y quemas y erosión de suelo. Finalmente, las fuentes fijas generan el 14%, principalmente en procesos industriales de producción de agregados, cemento y pinturas (8%). (Corporación para el Mejoramiento del Aire de Quito, 2006)

2.3.4 Instrumentos de medida

La Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ) cuenta con equipos de marcas como Thermo Environmental Instruments (TEI) y TELEDYNE API, según el detalle que se presenta en la Tabla 1-2

Tabla 1-2: Detalle de los analizadores de gases y partículas de la REMMAQ

Contaminante	Número Equipos	Ubicación	Método de medida o principio de operación	Marca y modelo
Material particulado PM2.5	6	Bel, Cam, Cen, Cot, Car, Sap	Atenuación de rayos beta (Método equivalente para PM10 EPA No. EQPM-1102-150)	Thermo Andersen / FH62C14
Dióxido de azufre (SO2)	7	Bel, Cam, Cen, Tum, Cot, Car, Chi	Fluorescencia por pulsos de luz ultravioleta (Método equivalente EPA No. EQSA-0486-060)	THERMO 43C / 43i
Ozono (O3)	10	Bel, Cam, Cen, Tum, Chi, Cot, Car, Gua, Jip*, Lab. Electrónico	Absorción de luz ultravioleta (Método equivalente EPA No. EQQA-0880-047)	THERMO 49C / 49i
Óxidos de nitrógeno (NOX)	8	Bel, Cam, Cen, Cot, Car, Gua, Jip, Chi	Quimiluminiscencia Método de referencia EPA No. RFNA-1289-074)	THERMO 42C / 42i
Monóxido de carbono (CO)	8	Bel, Cam, Cen, Cot, Car, Gua, Jip, Lab. Electrónico	Absorción infrarroja no dispersiva (Método de referencia EPA No. RFCA-0981-054)	THERMO / 48C / 48i
Multicalibrador (SO2, NOX, CO, O3)	12	Bel, Cam, Cen, Tum, Chi, Cot, Car, Gua, Jip, Lab. Electrónico, E. móvil	Principio de operación: Dilución de gases, aire cero con un material de referencia certificado (contaminante de concentración conocida).	THERMO / 146C / 146i
Generador Aire Cero	12	Bel, Jip, Cam, Cen, Tum, Chi, Cot, Car, Gua, Lab. electrónico, E. móvil	Principio de operación: Filtración de aire comprimido por medio de carbón activado y purafill, y calentamiento para oxidación.	THERMO / 111

Fuente: (Municipio del Distrito Metropolitano de Quito, 2021)

Realizado por: Allauca Pancho, Fabián, 2021

2.3.5 Ubicación de las estaciones de la Red Metropolitana de Monitoreo Atmosférico de Quito

Las estaciones remotas de monitoreo con capacidad para analizar continua y automáticamente cada hora los contaminantes comunes del aire se ubican de acuerdo a la Tabla 2-2

Tabla 2-2: Ubicación de las estaciones de la REMMAQ

Código	Nombre	Altitud(msnm)	Dirección
COT	Cotocollao	277	Museo Cotocollao (Santa Teresa #70-121 entre Ignacio Loyola y Alfonso del Hierro)
CAR	Carapungo	2851	Edificio Andinatel Carapungo (Super Manzana B, el Verjel S/N)
BEL	Belisario	2835	Terraza del Edificio Administrativo del Colegio San Gabriel (Av. America 3541)
JIP	Jipijapa	2781	Patio de la Dirección Metropolitana Ambiental (Río Coca 1731 e Isla Fernandina)
CAM	El Camal	2840	Terraza del Hosp. Patronato Municipal San José del Sur (Adrián Navarro 1660 e Hinostroza)
CEN	Centro	2820	Terraza de la Radio Municipal (ex Hogar Javir, García Moreno 751 y Sucre)
GUA	Guamaní	2887	Tumbaco
TUM	Tumbaco	2331	Terraza de Andinatel (Gaspar de Carvajal)
LCH	Los Chillos	2453	Terraza de Andinatel (Av. Ilaló, Vía a El Tingo)

Fuente: (Municipio del Distrito Metropolitano de Quito, 2021)

Realizado por: Allauca Pancho, Fabián, 2021

2.3.6 Teoría del análisis de datos funcionales

El análisis de datos funcionales (ADF o FDA) es aquella parte de la estadística que trabaja con muestras de funciones aleatorias. Ramsay y Silverman (1997), entre otros, han introducido herramientas de análisis para este tipo de datos. Las medidas de tendencia central, de dispersión y de relación entre variables aplicadas a muestras de variables aleatorias se pueden definir de manera análoga para muestras de datos funcionales. Sólo hace falta considerar que ahora estaremos trabajando en un espacio vectorial distinto: el espacio L^2 (funciones de cuadrado integrable). En realidad, Ramsay y Silverman definen estas medidas de manera general para cualquier espacio que tenga definido un producto escalar, y luego deduce de ellas las definiciones para muestras de funciones. (Navarro Pérez, 2004)

El ADF presenta algunas ventajas sobre el análisis clásico de Series de Tiempo, ya que para este último es necesario que los intervalos de tiempo para los cuales fueron tomadas las mediciones de las observaciones deben ser equiespaciados, en otras palabras, los intervalos de tiempo iguales, además deben ser procesos estacionarios, que para el ADF no es condicional. (Guerrón Varela, 2015)

El análisis de datos funcionales tiene cuatro áreas de desarrollo, la descripción de la información, los modelos de regresión, el análisis de componentes principales y los modelos de clasificación, en éstas tres áreas se enfoca la mayoría de la investigación del ADF con aplicación en medicina (cardiología), finanzas, economía, meteorología, entre otras. (Guerrón Varela, 2015)

Con respecto a las referencias bibliográficas básicas tenemos los aportes de los profesores Ramsay y Silverman con su libro *Functional Data Analysis*, posteriormente un libro de los mismos autores *Applied Functional Data Analysis: Methods and Case Studies* en el cual se detallan algunos casos de estudio y el libro *Functional Data Analysis with R and MATLAB* de Ramsay J.O., y Hooker G. y Graves S un texto que se usan dos programas muy empleados en las carreras de ingeniería como son los citados, sin dejar atrás el aporte los profesores Ferraty y Vieu con su libro *Nonparametric Functional Data Analysis: Theory and Practice*, fuente indispensable para iniciar cualquier estudio de Análisis de Datos Funcionales. (Guerrón Varela, 2015)

2.3.6.1 Espacios de Hilbert

Para poder entender lo que es un espacio de Hilbert, se ha tomado textualmente lo que Erwin Kreyszig en su libro “*Introductory Functional Analysis with Applications*” menciona:

En un espacio normado podemos sumar vectores y multiplicar vectores por escalares, al igual que en el álgebra elemental de vectores. Además, la norma sobre tal espacio generaliza el concepto elemental de la longitud de un vector. Sin embargo, lo que todavía falta en un espacio normado general, y lo que nos gustaría tener si fuera posible, es un análogo del familiar producto escalar.

$$a \cdot b = \alpha_1 \beta_1 + \alpha_2 \beta_2 + \alpha_3 \beta_3$$

y fórmulas resultantes, en particular

$$|a| = \sqrt{a \cdot a}$$

y la condición de ortogonalidad (perpendicularidad)

$$a \cdot b = 0$$

que son herramientas importantes en muchas aplicaciones. Por tanto, surge la cuestión de si el producto escalar y la ortogonalidad se pueden generalizar a espacios vectoriales arbitrarios. De hecho, esto se puede hacer y conduce a *espacios de productos internos* y espacios de productos internos completos, llamados **espacios de Hilbert**.

Los espacios de producto interno son espacios normados especiales, como veremos. Históricamente son más antiguos que los espacios normados generales. Su teoría es más rica y conserva muchas características del espacio euclideo, siendo un concepto central la ortogonalidad. De hecho, los espacios de productos internos son probablemente la generalización más natural del espacio euclideo, y el lector debe notar la gran armonía y belleza de los conceptos y pruebas en este campo. Toda la teoría fue iniciada por el trabajo de D. Hilbert (1912) sobre ecuaciones integrables. La notación y la terminología geométricas utilizadas actualmente son análogas a las de la geometría euclidea y fueron acuñadas por E. Schmidt (1908), quien siguió una sugerencia de G. Kowalewski. Estos

espacios han sido, hasta ahora, los espacios más útiles en aplicaciones prácticas de **análisis funcional**.

Conceptos importantes, breve orientación sobre el contenido principal

Un espacio de producto interno X es un espacio vectorial con un producto interno $\langle x, y \rangle$ definido en él. Este último generaliza el producto escalar de los vectores en el espacio tridimensional y se utiliza para definir

- (I) Una norma $\| \cdot \|$ por $\|x\| = \langle x, x \rangle^{1/2}$
- (II) Ortogonalidad por $\langle x, y \rangle = 0$

Un espacio de Hilbert \mathbf{H} es un espacio de producto interno completo. La teoría del producto interno y los espacios de Hilbert es más rica que la de los espacios generales normados y de Banach. Las características distintivas son

- (i) representaciones de \mathbf{H} como una suma directa de un subespacio cerrado y su complemento ortogonal,
- (ii) conjuntos y secuencias ortonormales y representaciones correspondientes de elementos de \mathbf{H} ,
- (iii) la representación de Riesz de funcionales lineales acotados por productos internos,
- (iv) el operador adjunto de Hilbert T^* de un operador lineal acotado T . (Kreyszig, 1978)

Para la definición formal de un espacio de Hilbert se cita lo mencionado por el autor Benalcazar en su obra “Espacios Normados y Aplicaciones”

Definición

Sea V un espacio vectorial real provisto de un producto escalar notado $\langle \cdot, \cdot \rangle$ y norma asociada $\| \cdot \|$. Sea (x_n) una sucesión en V . Recordemos que (x_n) es una sucesión de Cauchy en V si y solo si (x_n) cumple con la siguiente propiedad:

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{Z}^+ \quad \text{con} \quad m, n \geq n_0 \rightarrow \|x_m - x_n\| < \varepsilon$$

Además, (x_n) es convergente en V si y solo si existe $x \in V$ tan que $\lim_{n \rightarrow \infty} x_n = x$, o lo que es lo mismo $x_n \rightarrow x$ y $n \rightarrow \infty$ y esto a su vez $\|x_n - x\| \rightarrow 0$. Toda sucesión convergente en V es una sucesión de Cauchy en V

Sea V un espacio prehilberiano. Se dice que V es un **espacio de Hilbert** si y solo si toda sucesión de Cauchy en V es convergente en V

El espacio vectorial V es un espacio de Hilbert si y solo si para cada sucesión de Cauchy (x_n) de V , existe $x \in V$ tal que $\lim_{n \rightarrow \infty} x_n = x$.

El espacio de las funciones continuas $C([a, b])$ en donde se define un producto escalar $\langle \cdot, \cdot \rangle$ dado por:

$$\langle f, g \rangle = \int_a^b f(t) g(t) dt \quad \forall f, g \in C([a, b])$$

Provisto de la norma $\|\cdot\|_2$ definida como:

$$\|f\|_2 = \left(\int_a^b |f(t)|^2 dt \right)^{1/2} \quad \forall f \in C([a, b])$$

O también llamado $L^2([a, b])$ (Benalcazar, 2005)

2.3.6.2 Configuración del sistema de bases para la construcción de funciones

Las funciones que queremos modelar suelen caer en dos categorías principales: periódica y no periódica. Para el caso de las funciones periódicas el sistema base que se usa por lo general es de Fourier, y en el caso de las funciones no periódicas la elección habitual es el sistema de base spline o bsplines o las wavelets. Así también podemos encontrar otras bases como: constante, monómica (polinómica), exponencial, poligonal y potencial. (Guerrón Varela, 2015)

Edgar Ramiro Gerrón Varela en su tesis titulada “Análisis de Datos Meteorológicos del Valle de los Chillos usando Datos Funcionales” menciona:

Utilizamos un conjunto de bloques de construcción de funciones $\phi_k, k = 1, \dots, K$ llamada funciones base, las cuales se combinan linealmente. Es decir, una función $x(t)$ definida de esta manera se expresa en notación matemática como:

$$x(t) = \sum_{k=1}^K c_k \phi_k(t) = c' \phi_k(t)$$

llamada expansión de la función base. Los parámetros; c_1, c_2, \dots, c_k son los coeficientes de la expansión. La expresión de la matriz en el último término de la ecuación anterior utiliza c para representar el vector de K coeficientes y ϕ denotar un vector de longitud K que contiene las funciones base.

Por lo general haremos la consideración que tenemos una muestra N de funciones, por lo tanto la expresión $x_i(t) = \sum_{k=1}^K c_{ik} \phi_k(t)$, donde $i = 1, 2, \dots, N$, que en notación matricial viene dada por:

$$x(t) = C\phi(t) \quad (\text{Guerrón Varela, 2015})$$

2.3.6.3 Suavizamiento mediante un operador diferencial lineal

Un operador diferencial lineal identificado con sus siglas Ldf, es una metodología de conseguir un modelo de datos funcionales que tengan la característica de poder calcular las derivadas de dicha función. La suavización de estas funciones se soporta usando la clase Lfd, que simplemente expresa un operador diferencial lineal. (Guerrón Varela, 2015)

La notación Lx se define como una aplicación de un operador diferencial lineal L a una función de x . Por ejemplo, tenemos la aceleración, $Lx = D^2x$, o algo más sofisticado como la aceleración armónica $L = \omega^2 D + D^3$. Una forma general se puede escribir de la forma: $Lx(t) = \beta_0(t)x(t) + \beta_1(t)Dx(t) + \dots + \beta_{m-1}(t)D^{m-1}x(t) + D^m x(t)$ donde los conocidos coeficientes función operador diferencial lineal $\beta_j(t), j = 0, \dots, m - 1$ son constantes o son funciones. (Guerrón Varela, 2015)

2.3.7 Herramientas del software libre R para análisis estadístico de datos

R es un software estadístico potente y flexible, este software compila sobre una gran variedad de plataformas UNIX, Windows y MacOS. Además, proporciona una amplia variedad de técnicas estadísticas y gráficas, entre las que se incluyen las metodologías básicas de análisis de datos funcionales (técnicas de reducción de dimensión y modelos lineales funcionales); es el paquete más usado por los investigadores en ADF, ya que es libre y permite el desarrollo de código propio para aquellos métodos no implementados directamente en sus librerías. (Aguilera del Pino & Aguilera Morillo, 2013)

El programa permite interactuar con el computador, paso a paso, según los resultados que son mostrados en la consola, es decir que puede uno monitorear los procesos. Una lista de ordenes pueden ser escritos en un archivo en ASCII con extensión “.R” y ser ejecutado mediante la orden source (nombre programa.R). Las órdenes de ejecución de los procedimientos son las mismas en cada plataforma, la diferencia está en la instalación y en algunas funciones de menú que presenta cada versión. (Escudero Villa, 2014)

El programa R dispone de una cantidad muy grande de paquetes lo necesario para realizar un análisis estadístico. R también provee toda la documentación mediante manuales y libros de todos los paquetes propuestos y está disponible en la dirección: <http://www.r-project.org/>. (Escudero Villa, 2014)

2.3.7.1 Librerías para Análisis de Datos Funcionales (ADF)

Las librerías para Análisis de Datos Funcionales en el software libre R son varias, por lo cual se describen algunas de ellas brevemente.

- **Librería fda:** implementa las técnicas de análisis de datos funcionales del libro de Ramsay y Silverman (2007).
- **Librería fda.usc:** integra y complementa la librería fda con medidas de profundidad, detección de datos atípicos (outliers) funcionales, modelos de regresión funcional y métodos de clasificación de un conjunto de datos funcionales.
- **Librería fds:** conjuntos de datos funcionales.
- **Librería far:** modelización de procesos autoregresivos hilbertianos.
- **Librería ftsa:** análisis de series de tiempo funcionales
- **Librería MFDF:** modelización funcional de datos de finanzas.
- **Librería refund:** regresión con datos funcionales.
- **Librería nfda:** análisis de datos funcionales no paramétrico.
- **Librería rainbow:** gráficos para análisis exploratorio de datos funcionales.

2.3.7.2 Funciones usadas en el tratamiento de datos

Comandos clave de R usadas para realizar el análisis exploratorio de los datos:

- `library(mgcv)`
- `library(fda.usc)`
- `read.table`
- `seq(a,b,c)`
- `fdata`
- `fdata.bootstrap`
- `func.mean`
- `outliers.depth.trim`

2.4 Hipótesis

No en todas las investigaciones cuantitativas se plantean hipótesis. El hecho de que formulemos o no hipótesis depende de un factor esencial: el alcance inicial del estudio. Las investigaciones cuantitativas que formulan hipótesis son aquellas cuyo planteamiento define que su alcance será correlacional o explicativo, o las que tienen un alcance descriptivo, pero que intentan pronosticar una cifra o un hecho. Esto se resume en la Tabla 3-2. (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2014)

Tabla 3-2: Formulación de hipótesis en estudios cuantitativos con diferentes alcances

Alcance del estudio	Formulación de hipótesis
Exploratorio	No se formulan hipótesis.
Descriptivo	Sólo se formulan hipótesis cuando se pronostica un hecho o dato.
Correlacional	Se formulan hipótesis correlacionales.
Explicativo	Se formulan hipótesis causales.

Fuente: (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2014)

Realizado por: Allauca Pancho, Fabián, 2021

El presente trabajo al ser una investigación descriptiva no presentará hipótesis, puesto que no se intenta pronosticar una cifra o un hecho de la contaminación del aire en el centro histórico de Quito para los años 2019 y 2020.

2.5 Identificación de variables

Por ser una investigación descriptiva solo se tiene una variable

2.5.1 Variable

Contaminación del aire; que monitorea la Red Metropolitana de Monitoreo Atmosférico de Quito en el centro histórico. Estos contaminantes monitoreados son; monóxido de carbono (CO), dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂), Ozono (O₃) y material particulado fino o de diámetro menor a 2,5 micrómetros (PM_{2,5}).

CAPÍTULO III

3. METODOLOGÍA DE INVESTIGACIÓN

3.1 Enfoque de investigación

La investigación se realiza mediante **enfoque cuantitativo**, el enfoque cuantitativo usa la recolección de datos para probar hipótesis con base a la medición numérica y el análisis estadístico, con el fin de establecer pautas de comportamiento y probar teorías.

“Las líneas de investigación que usan el enfoque cuantitativo tratan de ser lo más objetivas, es decir, la observación y la medición de los fenómenos no deben ser afectados por el investigador. Las decisiones críticas se toman antes de recolectar la información por lo cual sigue un patrón predecible y estructurado. Los datos encontrados son generalizables, esto es, se considera que los resultados obtenidos en la muestra son aplicables a la parte mayor a la cual representa: el universo o población.” (Granda, 2015)

3.2 Diseño de investigación

El diseño de investigación seleccionado para el desarrollo del presente estudio es el **diseño no experimental**, puesto que en este tipo de diseño de investigación no se altera, ni se crea una situación, sino que se observan las situaciones ya existentes, es decir, que las variables independientes no son controlables, porque ya sucedieron, al igual que sus efectos. (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2014)

El diseño no experimental es seleccionado para el presente trabajo ya que los datos de monóxido de carbono (CO), dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂), Ozono (O₃) y material particulado fino o de diámetro menor a 2,5 micrómetros (PM_{2,5}); monitoreados por la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ); no serán alterados, los datos observados por la estación ubicada en el centro histórico de Quito serán analizados por medio de análisis de datos funcionales para de esta manera encontrar el comportamiento de estos contaminantes.

De acuerdo al propósito de la presente investigación se ha seleccionado el tipo de **diseño no experimental transversal o transeccional**, puesto que este tipo de diseño no experimental se caracteriza por recolectar datos en un solo momento, en un tiempo único. Su propósito es describir

variables y analizar su incidencia e interrelación en un momento dado. Es como “tomar una fotografía” de algo que sucede. (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2014)

3.3 Tipo de investigación

Para el desarrollo de la investigación se utiliza la **investigación descriptiva**, esta investigación busca especificar las propiedades, las características y los perfiles de las personas, grupos, comunidades, objetos o cualquier otro fenómeno que se someta a un análisis. Es decir, consiste en la caracterización de un hecho, fenómeno o grupo con el fin de establecer su estructura o comportamiento. (Grajales, 2017)

La presente investigación al tener como objetivo general el analizar descriptivamente los datos funcionales de la contaminación del aire en el centro histórico de Quito. Periodo 2019 – 2020. Se encaja a este tipo de investigación descriptiva.

3.4 Modalidad de investigación

La modalidad que se utiliza es la modalidad de **investigación documental**, esta lleva un proceso basado en la búsqueda, recuperación, análisis, crítica e interpretación de datos primarios y secundarios, es decir, los obtenidos y registrados por otros investigadores en fuentes documentales impresas, audiovisuales y electrónicas. (Arias Odón, 2012)

Los datos de los contaminantes del aire en el centro histórico de Quito, utilizados para esta investigación son los registrados por la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ), para los años 2019 y 2020; los mismos que podemos encontrar en la página web de la Secretaría de Ambiente de Quito (www.quitoambiente.gob.ec), apartado, Red de Monitoreo Atmosférico, link, descarga datos históricos.

3.5 Población y muestra

La población motivo de estudio es el conjunto de los contaminantes del aire presentes en el centro histórico de Quito a lo largo de todo el tiempo.

La muestra de estudio son los **datos** de los contaminantes del aire que monitorea y registra la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ), a cada hora del día, en el centro histórico de la ciudad, para los años 2019 y 2020.

Los contaminantes del aire monitoreados son; monóxido de carbono (CO), dióxido de azufre (SO₂); dióxido de nitrógeno (NO₂), ozono (O₃), y, material particulado fino o de diámetro menor a 2.5 micrómetros (PM_{2.5}).

Cabe mencionar que el registro de estos cinco contaminantes para el centro histórico de Quito son mediciones realizadas por equipos ubicados en la terraza de la Radio Mundial (ex Hogar Javier, García Moreno 751 y Sucre).

3.6 Operacionalización de las variables

Tabla 4-3: Operacionalización de variables

Variable	Conceptualización	Dimensiones	Indicadores	Definición de los indicadores	Criterio de medición	Técnica	Instrumento	Escala
Contaminación del aire	Es la emisión al aire de sustancias peligrosas a una tasa que excede la capacidad de los procesos naturales de la atmósfera para transformarlos, precipitarlos y depositarlos o diluirlos por medio del viento y el movimiento del aire. (Yassi, Kjellstrom, de Kok, & Guidotti, 2002)	Tipos de contaminantes del aire	<ol style="list-style-type: none"> 1. Monóxido de carbono 2. Dióxido de azufre 3. Dióxido de nitrógeno 4. Ozono 5. Material particulado fino 	Los cinco contaminantes mencionados en los indicadores, son los contaminantes que son monitoreados por la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ) en el centro histórico de la ciudad	La medición tiene como finalidad producir datos confiables sobre la concentración de contaminantes atmosféricos en el territorio del Distrito Metropolitano de Quito	Observación; descarga de la base de datos de los cinco contaminantes mencionados en los indicadores desde la página web www.quitoambiente.gob.ec	Hojas de Microsoft Excel	<ol style="list-style-type: none"> 1. mg/m³ 2. ug/m³ 3. ug/m³ 4. ug/m³ 5. ug/m³

Realizado por: Allauca Pancho, Fabián, 2021

3.7 Matriz de consistencia

Tabla 5-3: Matriz de consistencia

Formulación del problema	Objetivo general	Hipótesis	VARIABLES	Indicadores	Técnicas	Instrumentos
¿Cuál es el análisis descriptivo para los datos funcionales de la contaminación del aire en el centro histórico de Quito, periodo 2019 - 2020?	Analizar descriptivamente los datos funcionales de la contaminación del aire en el centro histórico de Quito. Periodo 2019 - 2020.	No aplica por ser una investigación descriptiva	Contaminación del aire	1. Monóxido de carbono 2. Dióxido de azufre 3. Dióxido de nitrógenos 4. Ozono 5. Material particulado fino	Observación; descarga de la base de datos de los cinco contaminantes mencionados en los indicadores desde la página web www.quitoambiente.gob.ec	Hojas de Microsoft Excel

Realizado por: Allauca Pancho, Fabián, 2021

3.8 Recolección de información

La recolección de la información de los contaminantes presentes en el aire en el centro histórico de Quito para los años 2019 y 2020 fue obtenida de la página web de la Secretaría de Ambiente de Quito (www.quitoambiente.gob.ec), apartado, Red de Monitoreo Atmosférico, link, descarga datos históricos.

La base de datos que presenta la Secretaría de Ambiente de Quito para el centro histórico es en archivos de hojas de cálculo de Microsoft Excel, para cada tipo de contaminante (monóxido de carbono, dióxido de azufre, dióxido de nitrógeno, Ozono y material particulado fino menor a 2,5 micrómetros)

3.9 Procesamiento de la información

El procesamiento que se realiza a las bases de datos proporcionada por la Secretaría de Ambiente de Quito para el monóxido de carbono, dióxido de azufre, dióxido de nitrógeno, Ozono y material particulado fino menor a 2,5 micrómetros es el siguiente:

- Seleccionar los datos de los años 2019 y 2020 de cada contaminante.
- Crear una nueva hoja de cálculo de Microsoft Excel con los datos de los años 2019 y 2020 para cada contaminante.
- Ordenar los datos en filas de manera cronológica para los días y en columnas para las horas del día; cabe recalcar que los datos de los contaminantes son registrados cada hora del día.
- Completar datos faltantes por medio de interpolación lineal.
- Leer en el software libre R las hojas de cálculo de Microsoft Excel de cada contaminante.
- Suavizar cada base de datos en el software R.
- Hallar los datos atípicos de cada base de datos.
- Determinar la media y la desviación estándar.
- Establecer la banda de confianza.
- Plantear un modelo matemático para datos funcionales de la contaminación del aire.

CAPÍTULO IV

4. RESULTADOS Y DISCUSIÓN

A continuación se muestran los resultados del comportamiento de los datos registrados por la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ) en el centro histórico para los contaminantes de monóxido de carbono (CO), dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂), Ozono (O₃) y material particulado fino o de diámetro menor a 2,5 micrómetros (PM_{2,5}); en forma general, y luego por períodos, los mismos que serán período escolar, vacaciones escolares, de lunes a viernes, sábados y domingos, de pandemia y de no pandemia, que permitan describir de mejor manera el comportamiento de estos contaminantes.

4.1 Análisis general

4.1.1 Monóxido de carbono (CO)

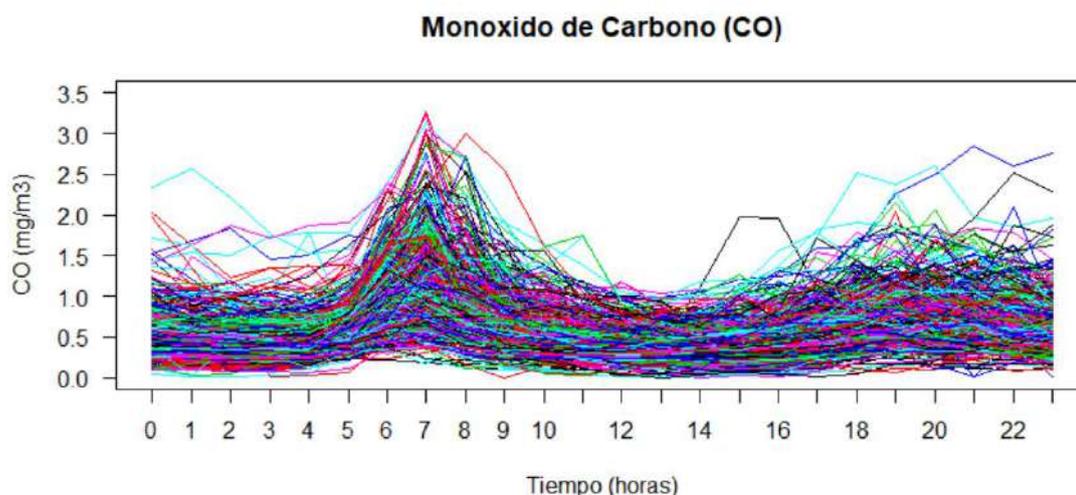


Figura 1-4: Curvas diarias del monóxido de carbono para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

En la Figura 1-4, se puede observar las 700 curvas diarias del monóxido de carbono correspondiente a los años 2019 y 2020, en la cual se aprecia unos picos alrededor de las siete de la mañana llegando a alcanzar aproximadamente los 3.26 mg/m³ y un pequeño incremento de la concentración de este contaminante alrededor de las siete de la noche alcanzando aproximadamente los 2.0 mg/m³; coincidiendo esta tendencia con las horas de máximo tráfico vehicular en el centro histórico de Quito. Cabe mencionar que el monóxido de carbono se forma por la combustión incompleta por déficit de oxígeno, cuya principal fuente son los vehículos automotores.

La concentración más elevada alrededor de las siete de la mañana se puede deber al proceso de fotosíntesis que realizan las plantas, absorbiendo carbono del ambiente durante el día y dejándolo de hacer durante la noche, puesto que para realizar este proceso las plantas necesitan de luz solar.

4.1.1.1 Detección de valores atípicos funcionales

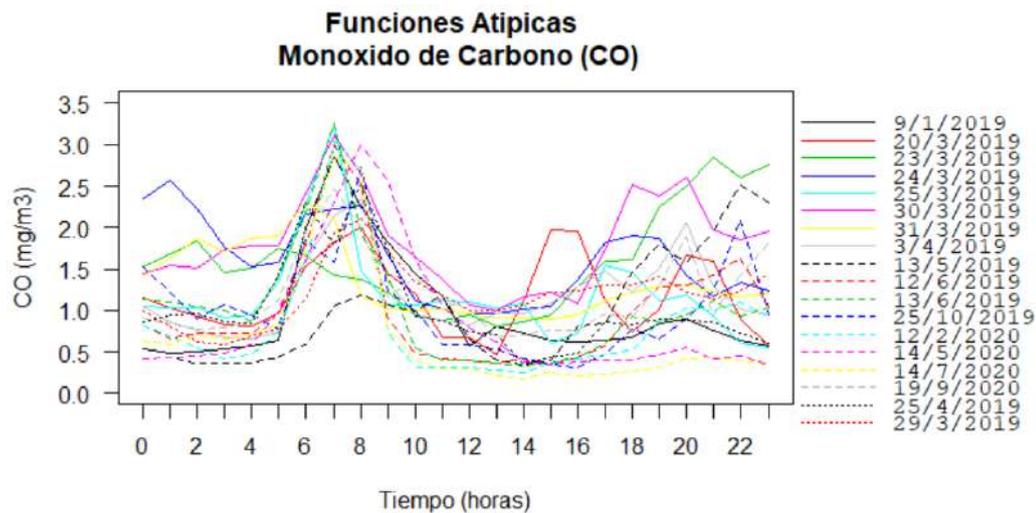


Figura 2-4: Funciones atípicas del monóxido de carbono para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

Las curvas de los días atípicos corresponden a los siguientes días:

- 9 de enero de 2019
- 20, 23, 24, 25, 29, 30 y 31 de marzo de 2019
- 3 y 25 de abril de 2019
- 13 de mayo de 2019
- 12 y 13 de junio de 2019
- 25 de noviembre de 2019
- 12 de febrero de 2020
- 14 de mayo de 2020
- 14 de julio de 2020
- 19 de septiembre de 2020

La Figura 2-4 muestra 18 curvas de los días atípicos, estas curvas no concuerdan con la tendencia de la distribución de los datos para el monóxido de carbono, ya que en algunas horas del día muestran valores más altos en comparación a las otras curvas. Se observa que el mes de marzo de

2019 posee 7 días atípicos, pudiéndose deber posiblemente alguna falla del equipo recolector de datos o algún motivo que incremento la concentración de este contaminante.

Para tener una mayor visualización de la tendencia de la distribución de los datos funcionales para el monóxido de carbono, se grafica sin estas 18 curvas atípicas, la Figura 3-4, muestra la gráfica exploratoria sin curvas atípicas; cabe recalcar que, para un posterior análisis, no se eliminaran estas curvas atípicas, puesto que se desconoce el origen de tal comportamiento.

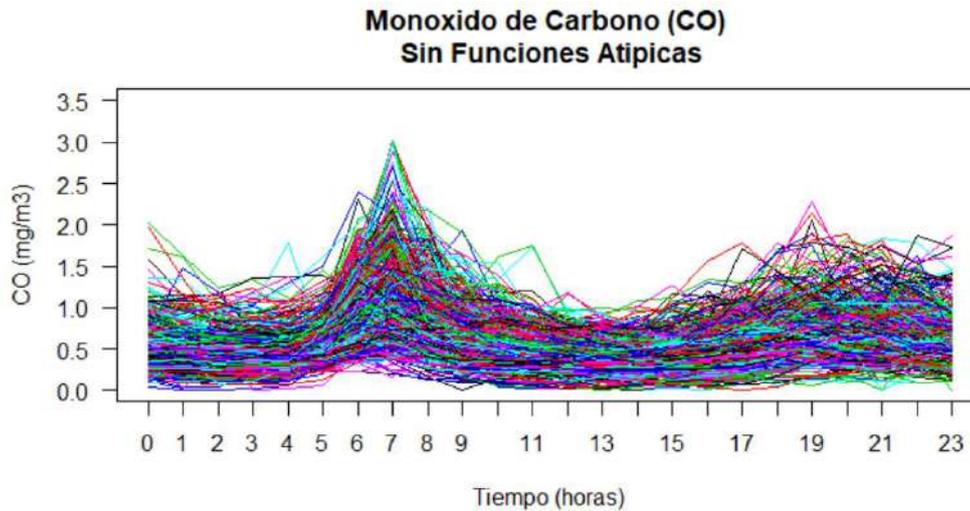


Figura 3-4: Gráfica exploratoria del monóxido de carbono sin curvas atípicas

Realizado por: Allauca Fabián, 2021

4.1.1.2 Curvas medias mensuales

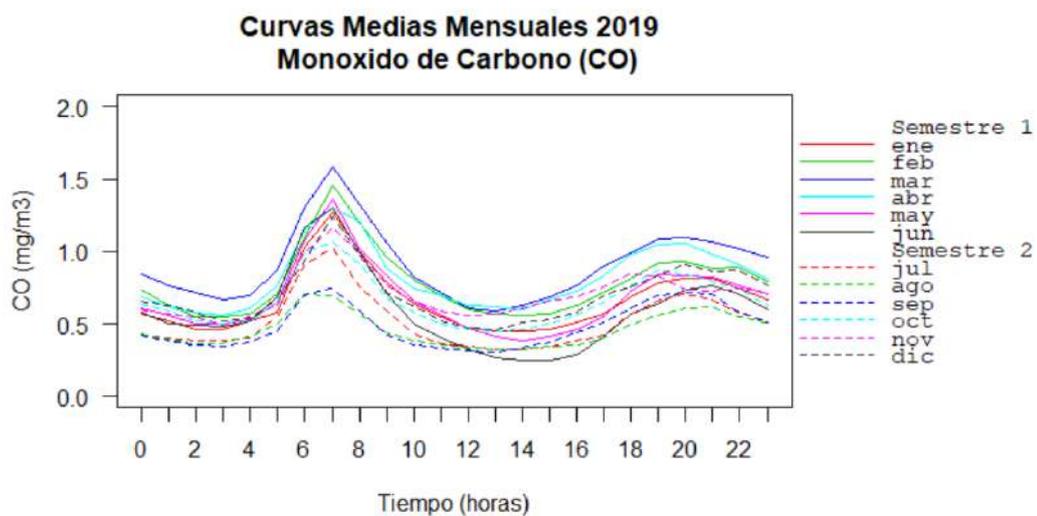


Figura 4-4: Curvas medias mensuales del monóxido de carbono para el año 2019

Realizado por: Allauca Fabián, 2021

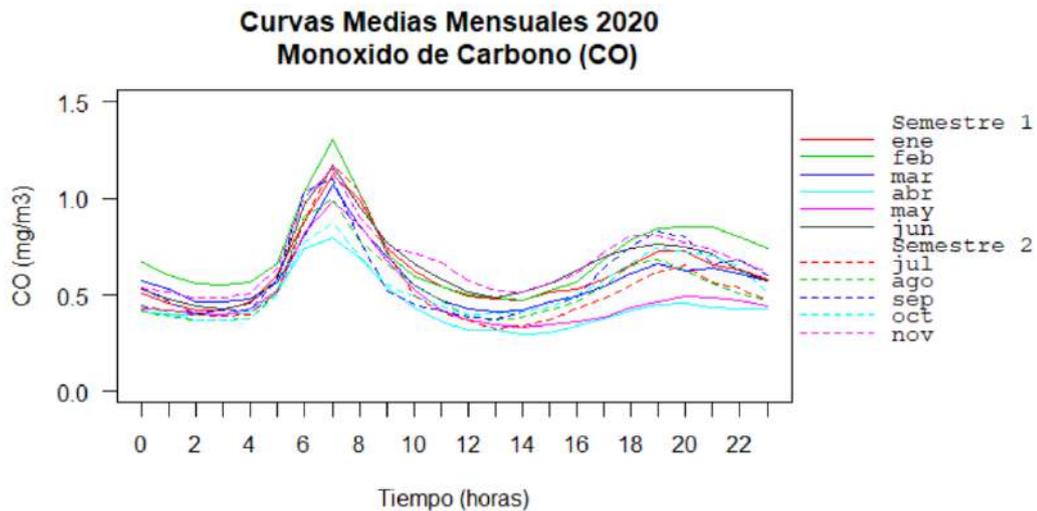


Figura 5-4: Curvas medias mensuales del monóxido de carbono para el año 2020

Realizado por: Allauca Fabián, 2021

Las Figuras 4-4 y 5-4 muestran las curvas de las medias mensuales del monóxido de carbono para los años 2019 y 2020 respectivamente. En la figura correspondiente al año 2019 se aprecia que las medias más altas corresponden a los meses de febrero, marzo y abril; mientras que los meses con medias más bajas corresponden a los meses de julio, agosto y septiembre; coincidiendo de alguna manera con los meses de vacaciones escolares. Por este indicio y para tener una mayor apreciación del comportamiento de este contaminante se realiza más adelante un análisis por períodos, siendo estos; período de clases escolares y periodo de vacaciones escolares.

En la Figura 5-4 se aprecia un comportamiento de las medias más parejo y menor en comparación con las curvas del 2019, esto se puede deber a que en el año 2020 se produjo la pandemia de covid-19, trayendo consigo un aislamiento y por ende a una menor circulación vehicular sobre todo en los meses de abril y mayo. Por tal razón se decide analizar también los períodos de pandemia y de no pandemia.

4.1.1.3 Curva media funcional con entorno de confianza bootstrap al 95%

La media funcional para los datos del monóxido de carbono se puede apreciar en la Figura 6-4, como también el entorno de confianza bootstrap al 95%. La curva media mostrada en color azul indica la tendencia vista anteriormente por los datos de este contaminante; donde se puede apreciar un mayor incremento alrededor de las siete de la mañana, como también un ligero incremento alrededor de las siete a ocho de la noche.

La curva media llega a un valor máximo de 1.1 mg/m^3 a las siete de la mañana y un valor aproximado de 0.75 mg/m^3 entre las siete y ocho de la noche. Cabe recalcar que estos valores,

incluso el máximo registrado (3.26 mg/m^3) a las siete de la mañana del día 13 de junio de 2019, tomados a una temperatura ambiente de $10.28 \text{ }^\circ\text{C}$ y a una presión atmosférica de 0.719 atmósferas; no conlleva ningún efecto perjudicial para la salud, ya que, de acuerdo a organismos internacionales, se ha fijado un valor límite de la media de ocho horas máxima en un día para la protección de la salud humana de 10 mg/m^3 . (Consejería de Salud de la Región de Murcia, 2021)

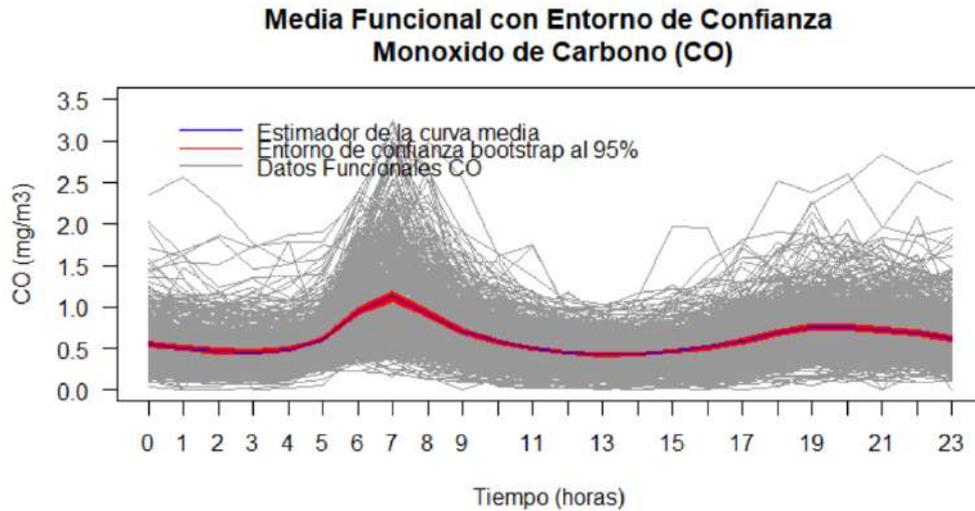


Figura 6-4: Estimador de curva media con entorno de confianza bootstrap al 95% del monóxido de carbono para los años 2019 y 2020.

Realizado por: Allauca Fabián, 2021

4.1.2 Dióxido de azufre (SO_2)

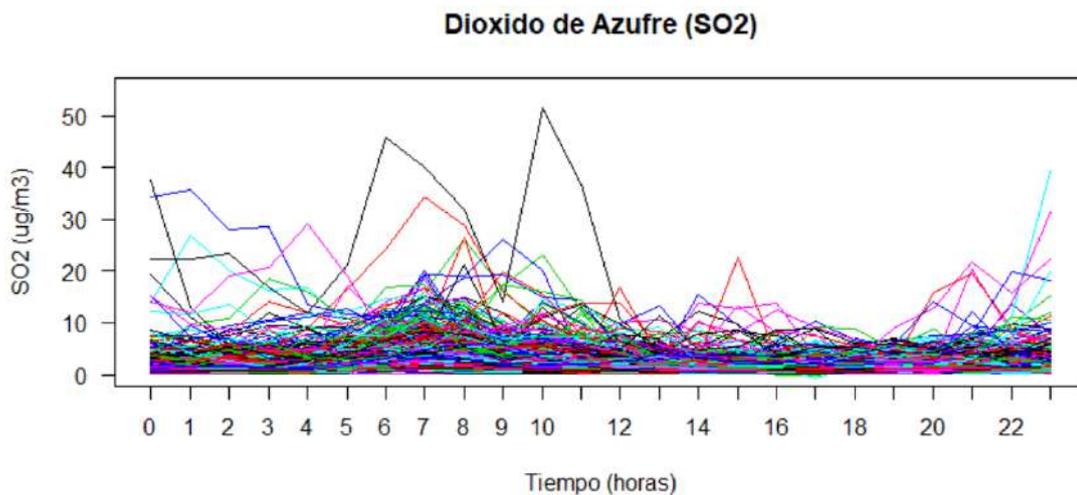


Figura 7-4: Curvas diarias del dióxido de azufre para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

En la Figura 7-4, se puede observar las 700 curvas diarias del dióxido de azufre correspondiente a los años 2019 y 2020, en la cual se aprecian algunas curvas que sobrepasan los $20 \mu\text{g}/\text{m}^3$ de SO_2 , posiblemente curvas atípicas que las estudiaremos más adelante. La Figura 7-4 no muestra claramente la tendencia de los datos de este contaminante, aunque si se puede apreciar un ligero pico a las siete de la mañana. Para tener una mejor observación de la tendencia de los datos, se hará necesario eliminar las curvas atípicas, con el único propósito de visualizar de mejor manera esta tendencia.

4.1.2.1 Detección de valores atípicos funcionales

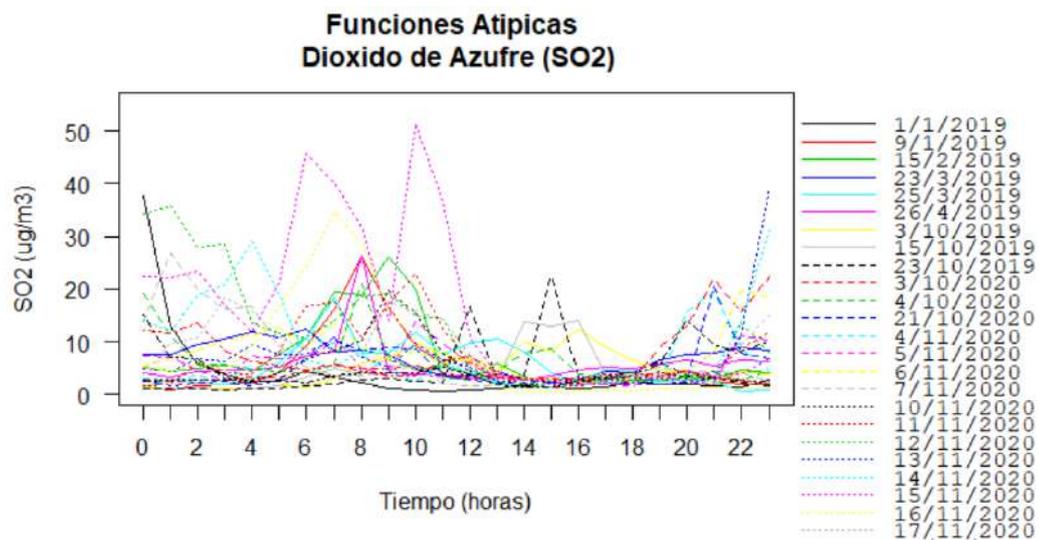


Figura 8-4: Funciones atípicas del dióxido de azufre para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

Las curvas de los días atípicos corresponden a los siguientes días:

- 1 y 9 de enero de 2019
- 15 de febrero de 2019
- 23 y 25 de marzo de 2019
- 26 de mayo de 2019
- 3, 8, 15 y 23 de octubre de 2019
- 3, 4, 21 de octubre de 2020
- 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 16, 17 y 18 de noviembre de 2020

La Figura 8-4 muestra 26 curvas de los días atípicos, estas curvas no concuerdan con la tendencia de la distribución de los datos para el dióxido de azufre, ya que estas sobrepasan los $20 \mu\text{g}/\text{m}^3$ de SO_2 , valor que la mayoría de las curvas no sobrepasan; y otras curvas muestran valores más altos en algunas horas del día en comparación con la gran mayoría de curvas. Se observa que el mes de

noviembre de 2020 posee 13 días atípicos, pudiéndose deber posiblemente alguna falla del equipo recolector de datos o algún motivo que incremento la concentración de este contaminante.

Para tener una mayor visualización de la tendencia de la distribución de los datos funcionales para el dióxido de azufre, se grafica sin estas 26 curvas atípicas, la Figura 9-4, muestra la gráfica exploratoria sin curvas atípicas; cabe recalcar que, para un posterior análisis, no se eliminaran estas curvas atípicas, puesto que se desconoce el origen de tal comportamiento.

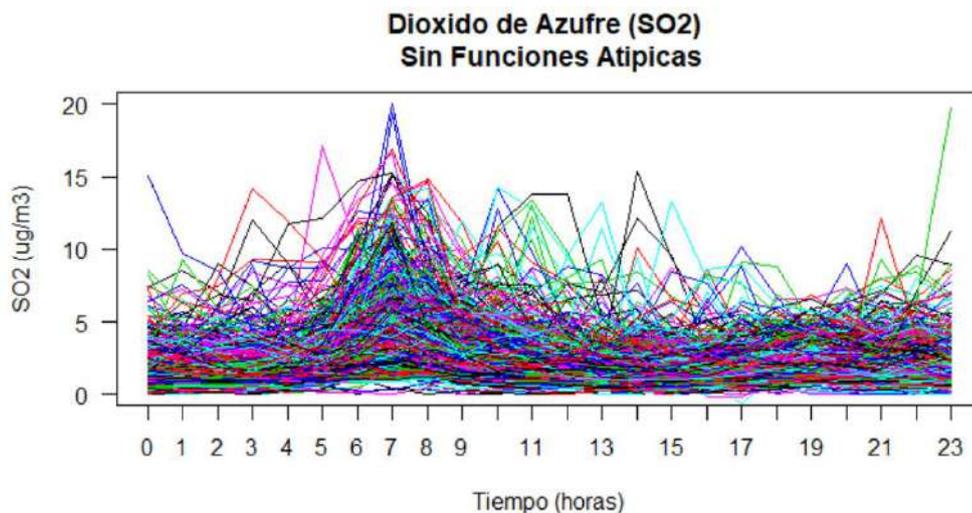


Figura 9-4: Gráfica exploratoria del dióxido de azufre sin curvas atípicas

Realizado por: Allauca Fabián, 2021

En la Figura 9-4, se observa de mejor manera la tendencia de los datos del dióxido de azufre, se visualiza que a las siete de la mañana existe un mayor incremento de este contaminante llegando alrededor de los 20 $\mu\text{g}/\text{m}^3$, esto se puede deber a que el dióxido de azufre es liberado en muchos procesos de combustión ya que los combustibles como el carbón, el petróleo, el diésel o el gas natural contienen ciertas cantidades de compuestos azufrados; coincidiendo de esta manera con las horas de máximo tráfico vehicular en el centro histórico de Quito.

4.1.2.2 Curvas medias mensuales

Las Figuras 10-4 y 11-4 muestran las curvas de las medias mensuales del dióxido de azufre para los años 2019 y 2020 respectivamente. En la figura correspondiente al año 2019 se aprecia que las medias más altas corresponden a los meses de enero, febrero, marzo, abril y mayo; mientras que el mes con media más baja corresponden al mes de agosto; coincidiendo de alguna manera con los meses de vacaciones escolares. Por este indicio y para tener una mayor apreciación del comportamiento de este contaminante se realiza más adelante un análisis por períodos, siendo estos; período de clases escolares y periodo de vacaciones escolares.

En la Figura 11-4 se aprecia un comportamiento de las medias más parejo y menor en comparación con las curvas del 2019, exceptuando el mes de noviembre de 2020, donde se pudo apreciar que ese mes contiene 13 días atípicos, por tal razón esta curva presenta valores más altos. Este comportamiento de las medias más parejas y menor en comparación con el año 2019 se puede deber a que en el año 2020 se produjo la pandemia de covid-19, trayendo consigo un aislamiento y por ende a una menor circulación vehicular sobre todo en los meses de abril y mayo. Por tal razón se decide analizar también los períodos de pandemia y de no pandemia.

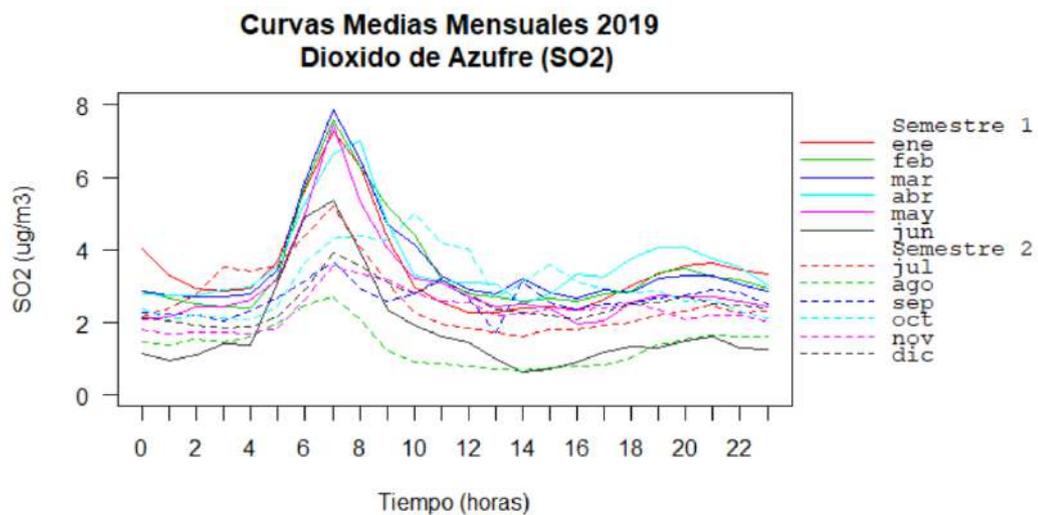


Figura 10-4: Curvas medias mensuales del dióxido de azufre para el año 2019

Realizado por: Allauca Fabián, 2021

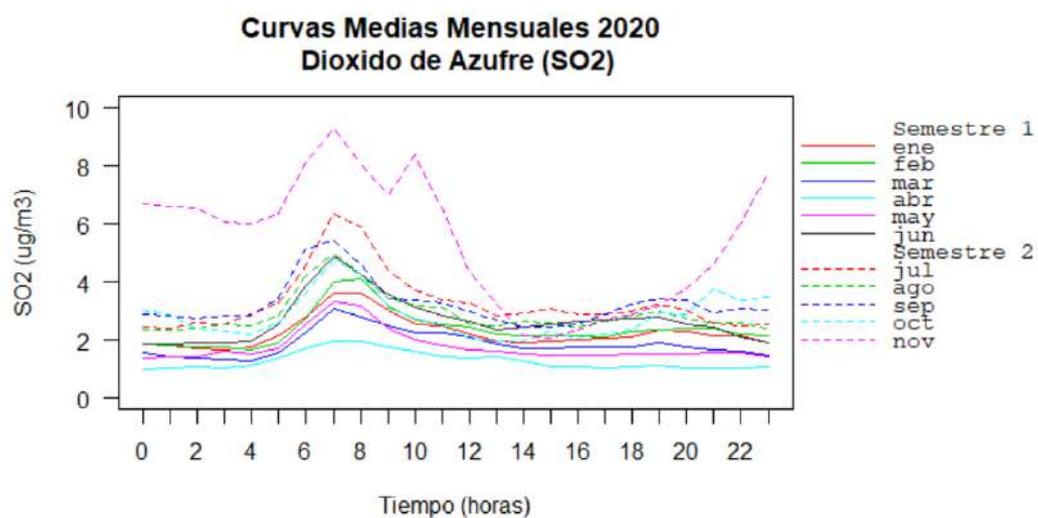


Figura 11-4: Curvas medias mensuales del dióxido de azufre para el año 2020

Realizado por: Allauca Fabián, 2021

4.1.2.3 Curva media funcional con entorno de confianza bootstrap al 95%

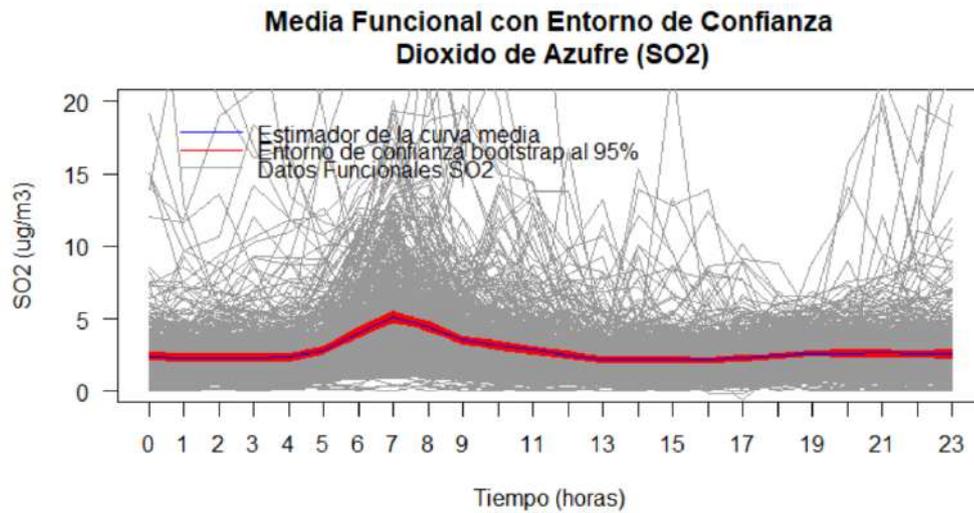


Figura 12-4: Estimador de curva media con entorno de confianza bootstrap al 95% del dióxido de azufre para los años 2019 y 2020.

Realizado por: Allauca Fabián, 2021

La media funcional para los datos del dióxido de azufre se puede apreciar en la Figura 12-4, como también el entorno de confianza bootstrap al 95%. La curva media mostrada en color azul indica más claramente la tendencia vista anteriormente por los datos de este contaminante; donde se puede apreciar un mayor incremento alrededor de las siete de la mañana, como también un ligero incremento alrededor de las siete a nueve de la noche.

La curva media llega a un valor máximo de $5 \mu\text{g}/\text{m}^3$ a las 07h00 y un valor aproximado de $2.6 \mu\text{g}/\text{m}^3$ entre las 19h00 y 21h00, la temperatura ambiente y la presión atmosférica promedio a la que fueron tomados los datos son de 14.86°C y 0.718 atm . respectivamente. La Organización Mundial de la Salud recomienda una concentración media de SO_2 de $20 \mu\text{g}/\text{m}^3$ por día, para que no exista daño a la salud de las personas. (Organización Mundial de la Salud, 2021). De acuerdo a la media funcional obtenida en esta investigación, la curva media no sobrepasa este valor.

4.1.3 Dióxido de nitrógeno (NO_2)

En la Figura 13-4, se puede observar las 700 curvas diarias del dióxido de nitrógeno correspondiente a los años 2019 y 2020, en la cual se aprecia una tendencia de doble montaña, la primera teniendo como cumbre un valor aproximado de $60 \mu\text{g}/\text{m}^3$ alrededor de las 7 a 8 de la mañana, y la segunda teniendo como cumbre el mismo valor aproximado de $60 \mu\text{g}/\text{m}^3$ alrededor

de las 5 a 6 de la tarde; correspondiendo estos picos o cumbres exactamente con las horas de máxima circulación vehicular. Cabe mencionar que el dióxido de nitrógeno tiene como una de sus fuentes de origen los procesos de combustión interna de motores diésel y gasolina de vehículos, representando el 80% para la ciudad de Quito en el año 2006. (Corporación para el Mejoramiento del Aire de Quito, 2006). Por este motivo se explicaría la tendencia que están tomando los datos de este contaminante.

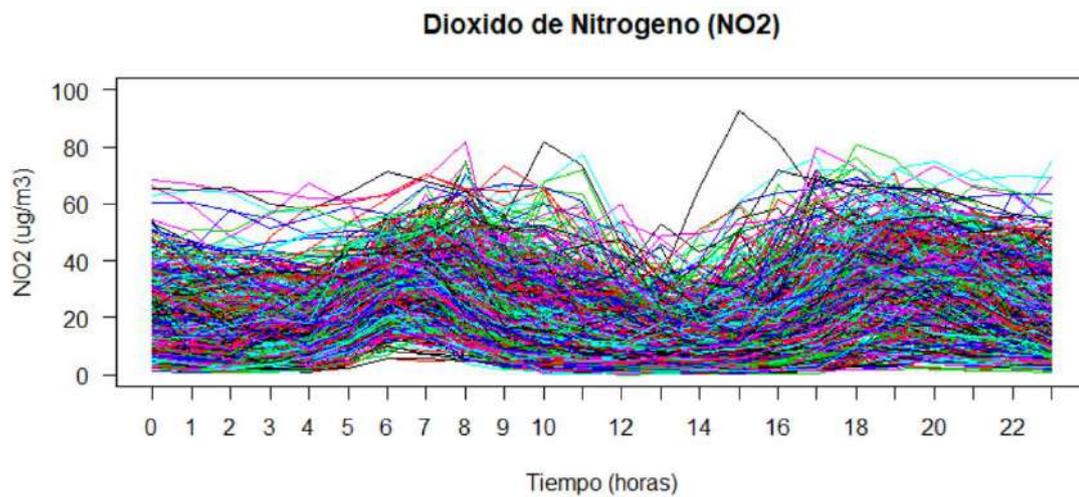


Figura 13-4: Curvas diarias del dióxido de nitrógeno para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

4.1.3.1 Detección de valores atípicos funcionales

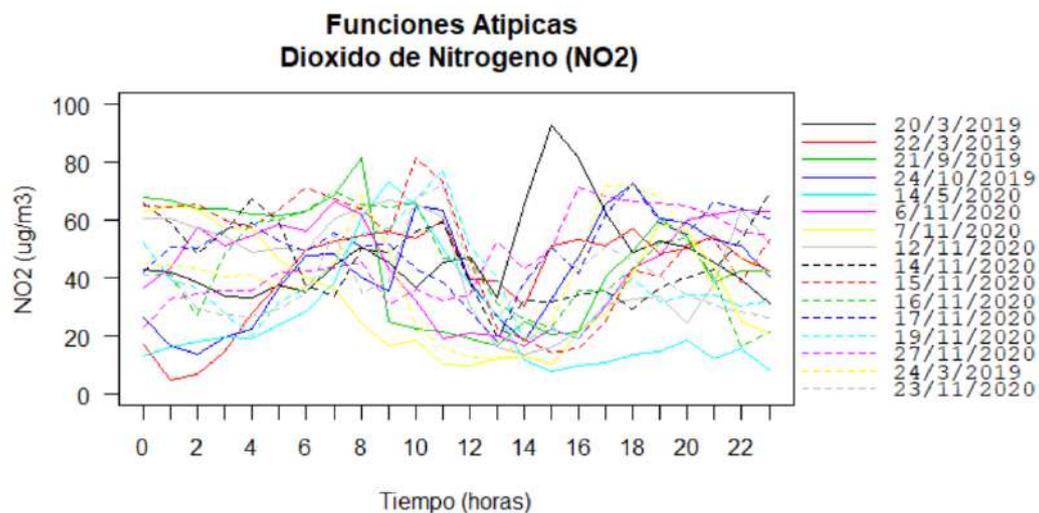


Figura 14-4: Funciones atípicas del dióxido de nitrógeno para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

Las curvas de los días atípicos corresponden a los siguientes días:

- 20, 22 y 24 de marzo de 2019

- 21 de septiembre de 2019
- 24 de octubre de 2019
- 14 de mayo de 2020
- 6, 7, 12, 14, 15, 16, 17, 19, 23 y 27 de noviembre de 2020

La Figura 14-4 muestra 16 curvas de los días atípicos, estas curvas no concuerdan con la tendencia de la distribución de los datos para el dióxido de nitrógeno, ya que en algunas horas del día muestran valores más altos en comparación a las otras curvas. Se observa que el mes de noviembre de 2020 posee 10 días atípicos, pudiéndose deber posiblemente alguna falla del equipo recolector de datos o algún motivo que incremento la concentración de este contaminante.

Para tener una mayor visualización de la tendencia de la distribución de los datos funcionales para el dióxido de nitrógeno, se grafica sin estas 16 curvas atípicas, la Figura 15-4, muestra la gráfica exploratoria sin curvas atípicas, pudiéndose apreciar de mejor manera la tendencia de doble montaña que están tomando los datos, presentando sus cumbres alrededor de las 7 a 8 de la mañana y alrededor de las 5 a 6 de la tarde; cabe recalcar que para un posterior análisis, no se eliminaran estas curvas atípicas, puesto que se desconoce el origen de tal comportamiento.

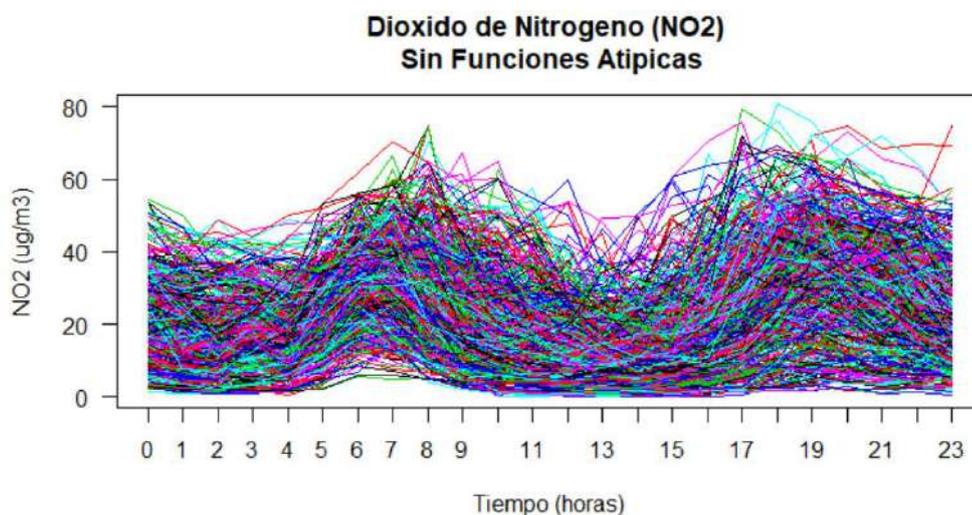


Figura 15-4: Gráfica exploratoria del dióxido de nitrógeno sin curvas atípicas

Realizado por: Allauca Fabián, 2021

4.1.3.2 Curvas medias mensuales

Las Figuras 16-4 y 17-4 muestran las curvas de las medias mensuales del dióxido de nitrógeno para los años 2019 y 2020 respectivamente. En la figura correspondiente al año 2019 se aprecia que las medias más altas corresponden a los meses de marzo y abril; mientras que los meses con

medias más bajas corresponden a los meses de julio y agosto; coincidiendo de alguna manera con los meses de vacaciones escolares. Por este indicio y para tener una mayor apreciación del comportamiento de este contaminante se realiza más adelante un análisis por períodos, siendo estos; período de clases escolares y periodo de vacaciones escolares.

En la Figura 17-4 se aprecia un comportamiento de las medias, menor, sobre todo para los meses de marzo, abril y mayo de 2020 en comparación con las curvas del 2019, exceptuando el mes de noviembre de 2020, donde se pudo apreciar 10 días atípicos. Este comportamiento de las medias, menor en comparación con el año 2019, se puede deber a que en el año 2020 se produjo la pandemia de covid-19, trayendo consigo un aislamiento y por ende a una menor circulación vehicular sobre todo en los meses de marzo, abril y mayo. Por tal razón se decide analizar también los períodos de pandemia y de no pandemia.

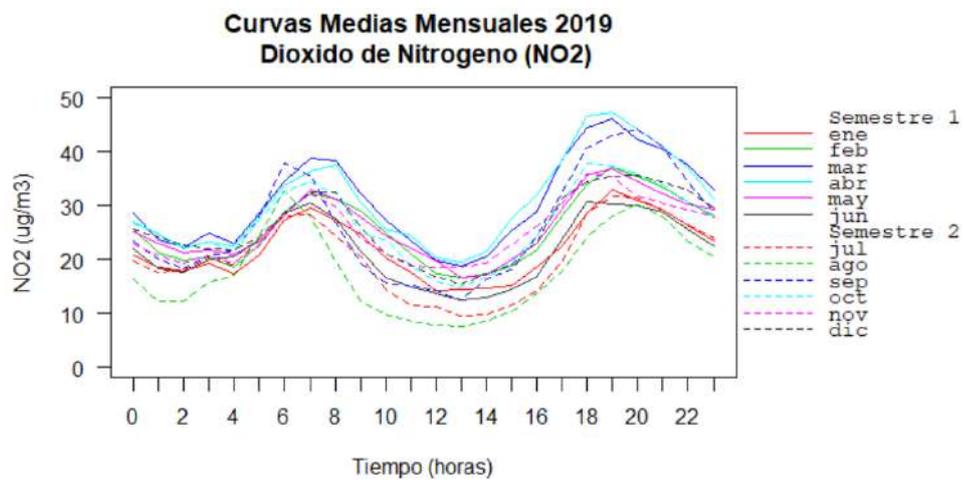


Figura 16-4: Curvas medias mensuales del dióxido de nitrógeno para el año 2019

Realizado por: Allauca Fabián, 2021

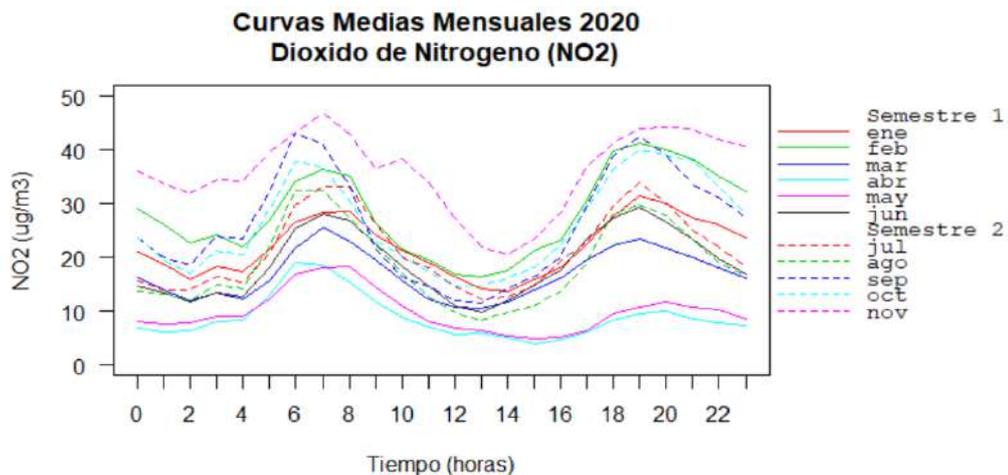


Figura 17-4: Curvas medias mensuales del dióxido de nitrógeno para el año 2020

Realizado por: Allauca Fabián, 2021

4.1.3.3 Curva media funcional con entorno de confianza bootstrap al 95%

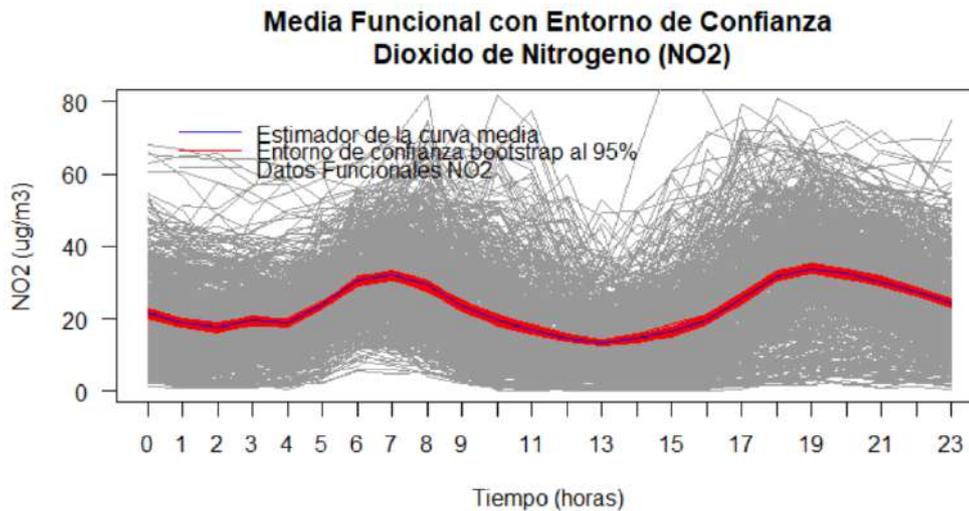


Figura 18-4: Estimador de curva media con entorno de confianza bootstrap al 95% del dióxido de nitrógeno para los años 2019 y 2020.

Realizado por: Allauca Fabián, 2021

La media funcional para los datos del dióxido de nitrógeno se puede apreciar en la Figura 18-4, como también el entorno de confianza bootstrap al 95%. La curva media mostrada en color azul indica la tendencia vista anteriormente de doble montaña, donde las cumbres llegan a un valor de $32 \mu\text{g}/\text{m}^3$ alrededor de las 6 a 8 de la mañana y un valor de $33 \mu\text{g}/\text{m}^3$ alrededor de las 6 a 8 de la noche, y tomando un valor de $12 \mu\text{g}/\text{m}^3$ en su punto más bajo.

La Organización Mundial de la Salud recomienda una concentración media anual de dióxido de nitrógeno de $40 \mu\text{g}/\text{m}^3$, para que no exista daño a la salud de las personas. (Organización Mundial de la Salud, 2021). De acuerdo a la media funcional obtenida en esta investigación, la curva media no sobrepasa este valor, pero se está muy cerca de este valor medio. La temperatura ambiente y la presión atmosférica promedio a la que fueron tomados los datos son de 14.86°C y 0.718 atm . respectivamente.

4.1.4 Ozono (O_3)

En la Figura 19-4, se puede observar las 700 curvas diarias del ozono correspondiente a los años 2019 y 2020, en la cual se aprecia un comportamiento parcialmente en forma acampanada, sobre todo a partir de las seis de la mañana hasta las seis de la tarde; llegando a tomar valores aproximados de $100 \mu\text{g}/\text{m}^3$ de O_3 en la zona más alta correspondiente a las 12 horas del mediodía.

Este comportamiento acampanado concuerda con la distribución en forma de campana que presentan los datos de la radiación solar en la ciudad de Riobamba para el año 2011. (Escudero Villa, 2014). Se asume que la radiación solar para la ciudad de Quito sea similar a los registrados en la ciudad de Riobamba por ser ciudades que están bajo similares condiciones de altura. Esta coincidencia se da puesto que el ozono se forma por la reacción con la luz solar (fotoquímica) de contaminantes como los óxidos de nitrógeno (NOx) procedentes de las emisiones de vehículos o la industria y los compuestos orgánicos volátiles (COV) emitidos por los vehículos, los disolventes y la industria. Los niveles de ozono más elevados se registran durante los períodos de tiempo soleado. (Organización Mundial de la Salud, 2021).

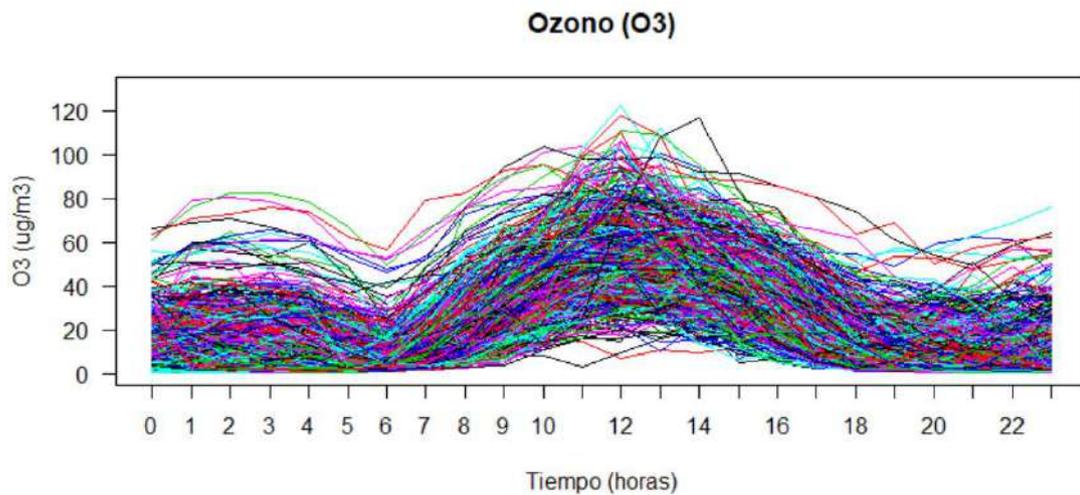


Figura 19-4: Curvas diarias del ozono para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

4.1.4.1 Detección de valores atípicos funcionales

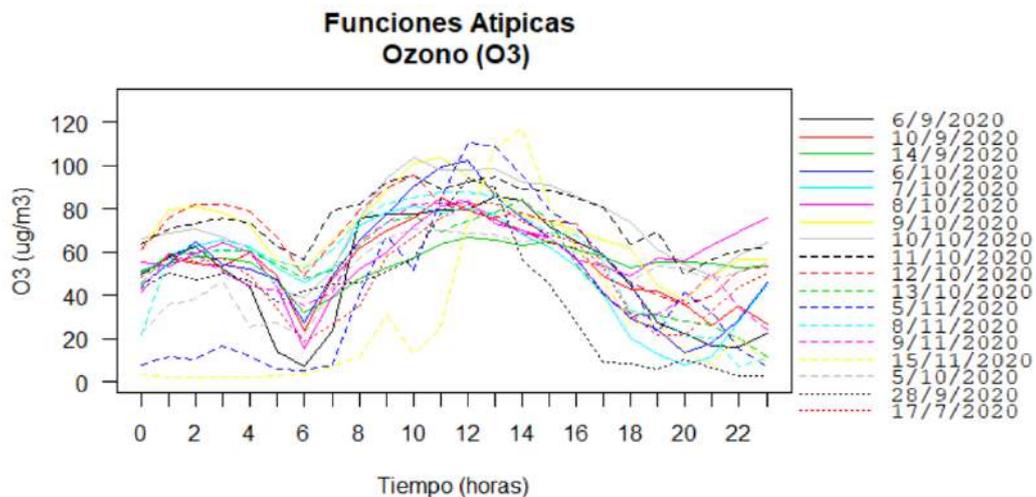


Figura 20-4: Funciones atípicas del ozono para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

Las curvas de los días atípicos corresponden a los siguientes días:

- 17 de julio de 2020
- 6, 10, 14 y 28 de septiembre de 2020
- 5, 6, 7, 8, 9, 10, 11, 12 y 13 de octubre de 2020
- 5, 8, 9 y 15 de noviembre de 2020

La Figura 20-4 muestra 18 curvas de los días atípicos, estas curvas no concuerdan con la tendencia de la distribución de los datos para el ozono, ya que en algunas horas del día muestran valores más altos en comparación a las otras curvas. Se observa que el mes de octubre de 2020 posee 9 días atípicos, al igual que septiembre y noviembre de 2020 presentan 4 días atípicos cada uno de estos meses; pudiéndose deber posiblemente alguna falla del equipo recolector de datos o algún motivo que incremento la concentración de este contaminante.

Para tener una mayor visualización de la tendencia de la distribución de los datos funcionales para el ozono, se grafica sin estas 18 curvas atípicas, la Figura 21-4, muestra la gráfica exploratoria sin curvas atípicas, pudiéndose observar de mejor manera el comportamiento acampanado desde las seis de la mañana hasta las seis de la tarde; cabe recalcar que, para un posterior análisis, no se eliminaran estas curvas atípicas, puesto que se desconoce el origen de tal comportamiento.

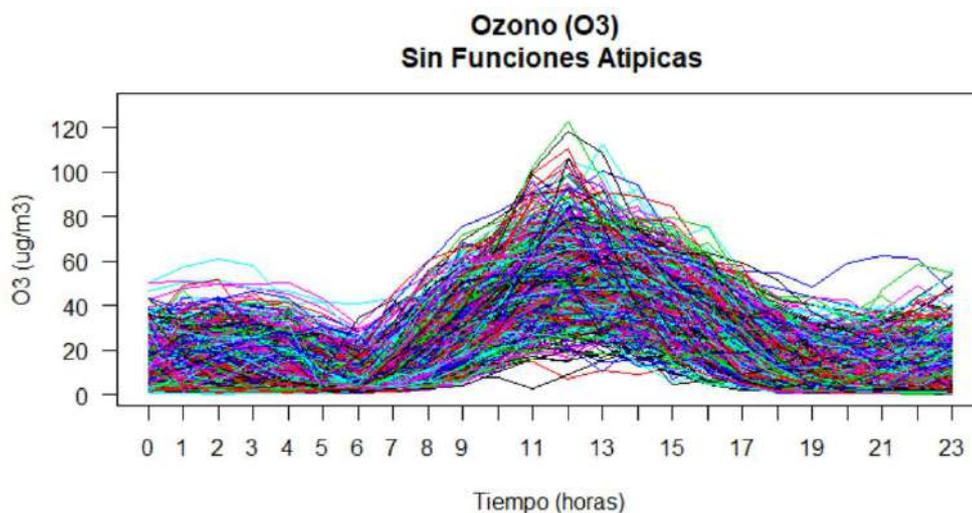


Figura 21-4: Gráfica exploratoria del ozono sin curvas atípicas

Realizado por: Allauca Fabián, 2021

4.1.4.2 Curvas medias mensuales

Las Figuras 22-4 y 23-4 muestran las curvas de las medias mensuales del ozono para los años 2019 y 2020 respectivamente. En la figura correspondiente al año 2019 se aprecia que las medias

más altas corresponden a los meses de agosto, septiembre y octubre; mientras que el mes con media más baja corresponden a enero; coincidiendo de alguna manera con los meses más soleados y menos soleados respectivamente.

En la Figura 23-4 se aprecia un comportamiento similar al de la Figura 22-4, presentando curvas medias altas para los meses de agosto, septiembre y una media más baja para el mes de enero; las curvas medias en general del año 2020 para el ozono son más altas en comparación con las del año 2019, dándonos un indicio que para el año 2020 la radiación solar fue mayor, esto se puede decir, ya que el ozono se forma por la reacción de la luz solar.

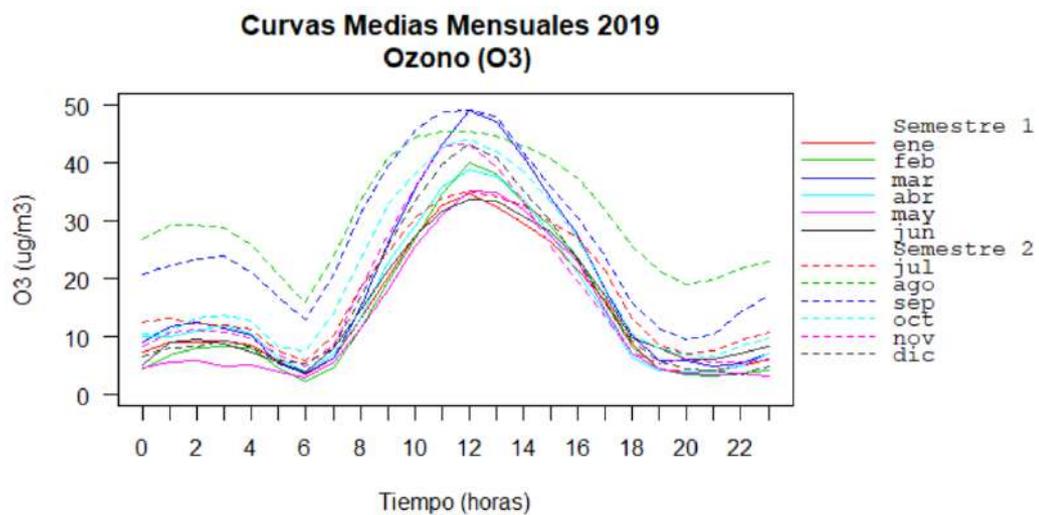


Figura 22-4: Curvas medias mensuales del ozono para el año 2019

Realizado por: Allauca Fabián, 2021

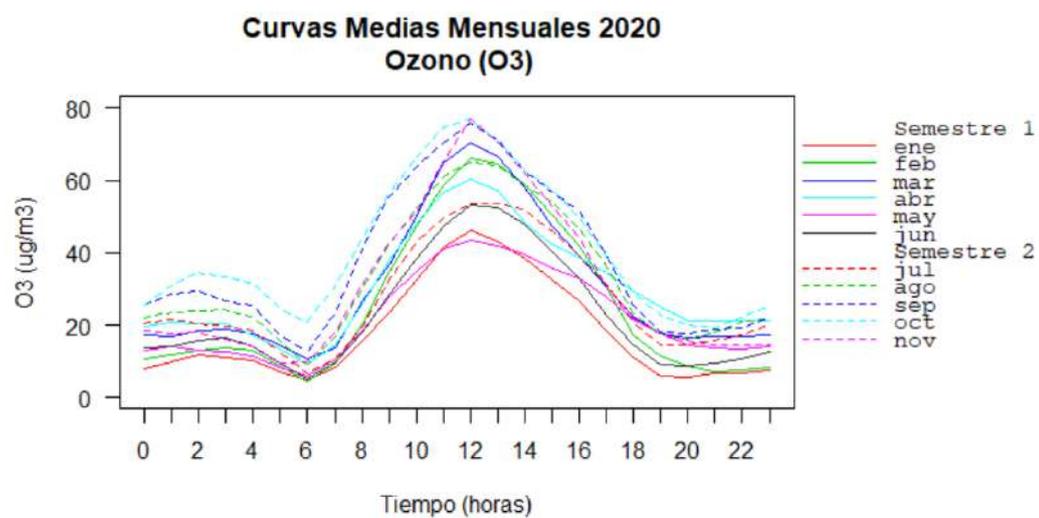


Figura 23-4: Curvas medias mensuales del ozono para el año 2020

Realizado por: Allauca Fabián, 2021

4.1.4.3 Curva media funcional con entorno de confianza bootstrap al 95%

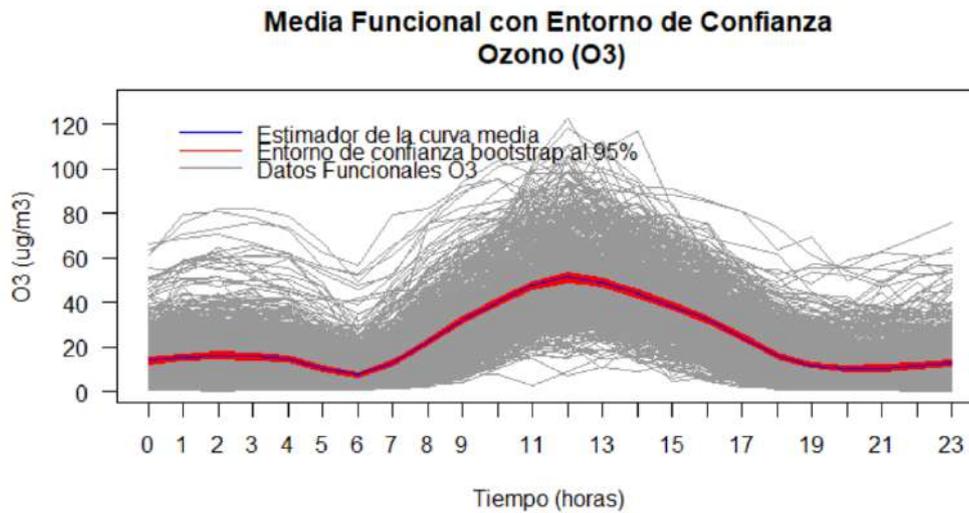


Figura 24-4: Estimador de curva media con entorno de confianza bootstrap al 95% del ozono para los años 2019 y 2020.

Realizado por: Allauca Fabián, 2021

La media funcional para los datos del ozono se puede apreciar en la Figura 24-4, como también el entorno de confianza bootstrap al 95%. La curva media mostrada en color azul indica la tendencia vista anteriormente por los datos de este contaminante, siendo esta de forma acampanada, de seis de la mañana a seis de la tarde; donde se puede apreciar un valor máximo de la campana al medio día, llegando este a un valor de $55 \mu\text{g}/\text{m}^3$ de O_3 .

La Organización Mundial de la Salud recomienda una concentración media en 8 horas de exposición de $100 \mu\text{g}/\text{m}^3$ de O_3 , para que no exista daño a la salud de las personas. (Organización Mundial de la Salud, 2021). De acuerdo a la media funcional obtenida en esta investigación, la curva media no sobrepasa este valor. La temperatura ambiente y la presión atmosférica promedio a la que fueron tomados los datos son de 14.86°C y 0.718 atm . respectivamente.

4.1.5 Material particulado fino (PM2.5)

En la Figura 25-4, se puede observar las 700 curvas diarias del material particulado fino correspondiente a los años 2019 y 2020, en la cual se aprecian algunas curvas que sobrepasan los $60 \mu\text{g}/\text{m}^3$ de $\text{PM}_{2.5}$, posiblemente curvas atípicas que las estudiaremos más adelante. La Figura 25-4 muestra claramente una tendencia ligeramente constante alrededor de los $20 \mu\text{g}/\text{m}^3$ de $\text{PM}_{2.5}$ a lo largo del día; la cual se verificará cuando obtengamos la media funcional para estos datos.

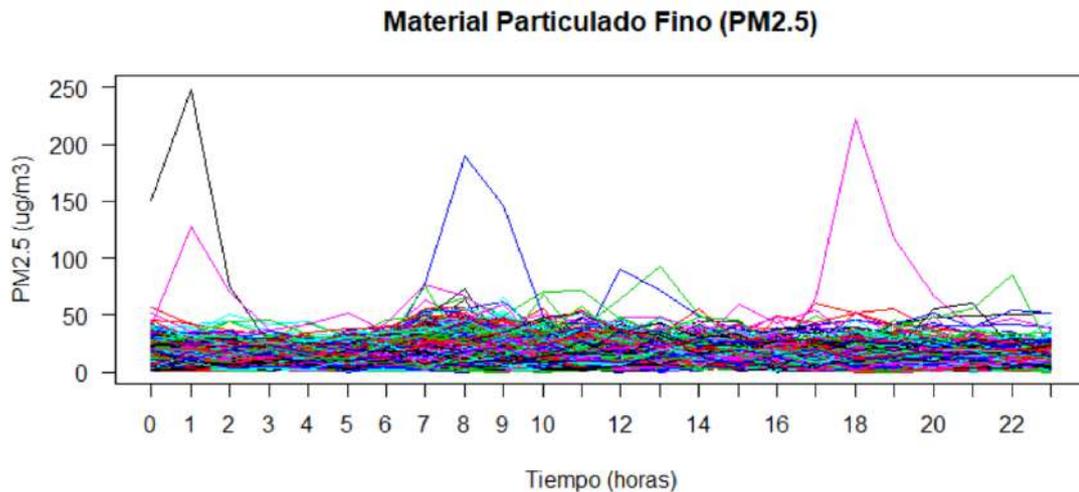


Figura 25-4: Curvas diarias del material particulado fino para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

4.1.5.1 Detección de valores atípicos funcionales

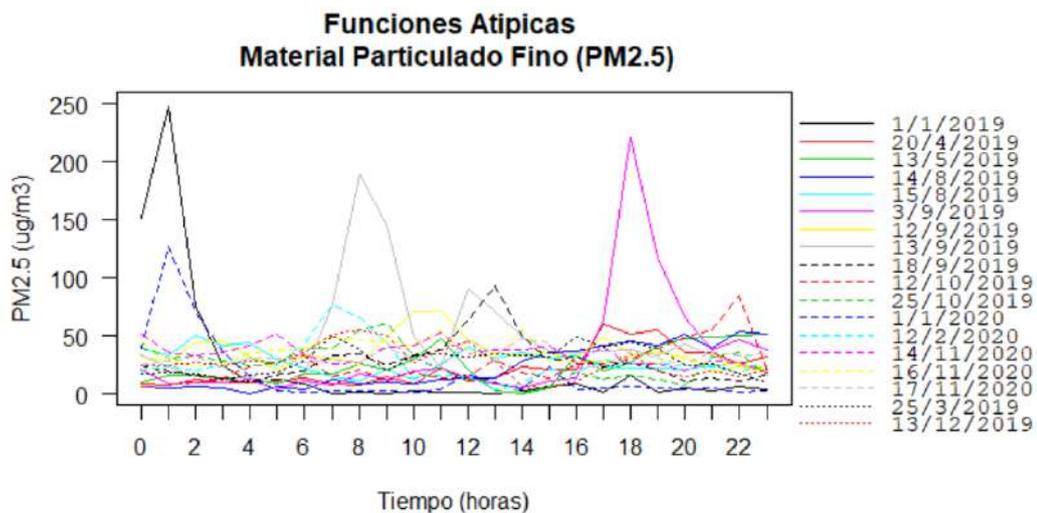


Figura 26-4: Funciones atípicas del material particulado fino para los años 2019 y 2020

Realizado por: Allauca Fabián, 2021

Las curvas de los días atípicos corresponden a los siguientes días:

- 1 de enero de 2019
- 25 de marzo de 2019
- 20 de abril de 2019
- 13 de mayo de 2019
- 14 y 15 de agosto de 2019
- 3, 12, 13 y 18 de septiembre de 2019
- 12 y 25 de octubre de 2019

- 13 de diciembre de 2019
- 1 de enero de 2020
- 12 de febrero de 2020
- 14, 16 y 17 de noviembre de 2020

La Figura 26-4 muestra 18 curvas de los días atípicos, estas curvas no concuerdan con la tendencia de la distribución de los datos para el material particulado fino, ya que en algunas horas del día muestran valores más altos en comparación a las otras curvas. Se observa que el mes de septiembre de 2019 posee 4 días atípicos y el mes de noviembre de 2020 posee 3 días atípicos; pudiéndose deber posiblemente alguna falla del equipo recolector de datos o algún motivo que incremento la concentración de este contaminante.

Para tener una mayor visualización de la tendencia de la distribución de los datos funcionales para el material particulado fino, se grafica sin estas 18 curvas atípicas, la Figura 27-4, muestra la gráfica exploratoria sin curvas atípicas; cabe recalcar que, para un posterior análisis, no se eliminaran estas curvas atípicas, puesto que se desconoce el origen de tal comportamiento.

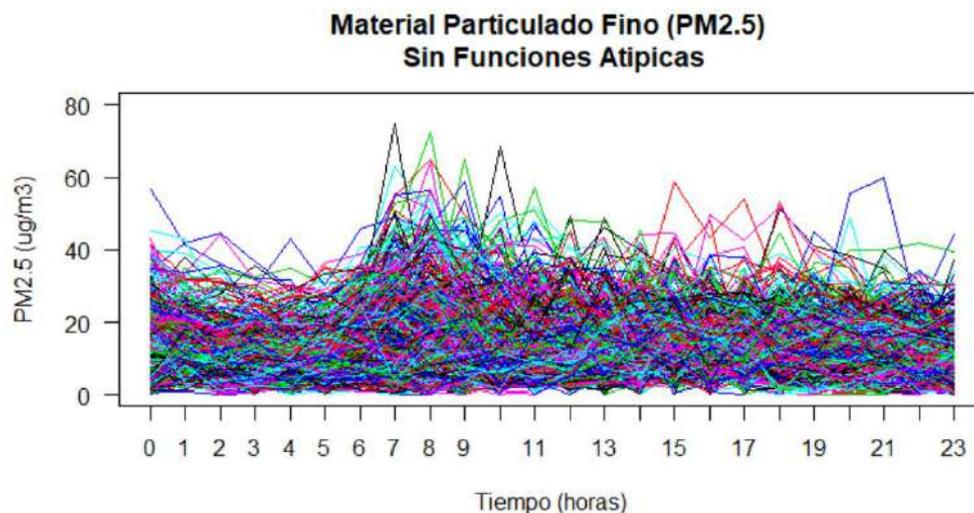


Figura 27-4: Gráfica exploratoria del material particulado fino sin curvas atípicas

Realizado por: Allauca Fabián, 2021

4.1.5.2 Curvas medias mensuales

Las Figuras 28-4 y 29-4 muestran las curvas de las medias mensuales del material particulado fino para los años 2019 y 2020 respectivamente. En la figura correspondiente al año 2019 se aprecia que las medias más altas corresponden a los meses de febrero y marzo; mientras que el mes con media más baja corresponde al mes de agosto; coincidiendo de alguna manera con los

meses de vacaciones escolares. Por este indicio y para tener una mayor apreciación del comportamiento de este contaminante se realiza más adelante un análisis por períodos, siendo estos; período de clases escolares y periodo de vacaciones escolares.

En la Figura 29-4 se aprecia un comportamiento de las medias, menor, en comparación con las curvas del 2019, exceptuando el mes de noviembre de 2020, donde se pudo apreciar 3 días atípicos. Este comportamiento de las medias, menor en comparación con el año 2019, se puede deber a que en el año 2020 se produjo la pandemia de covid-19, trayendo consigo un aislamiento y por ende a una menor circulación vehicular sobre todo en los meses de abril y mayo. Por tal razón se decide analizar también los períodos de pandemia y de no pandemia.

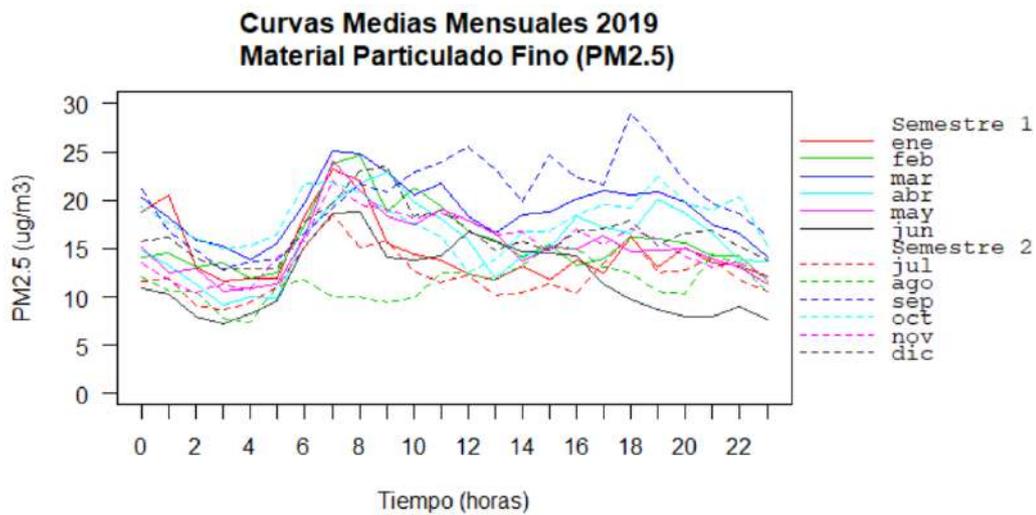


Figura 28-4: Curvas medias mensuales del material particulado fino para el año 2019

Realizado por: Allauca Fabián, 2021

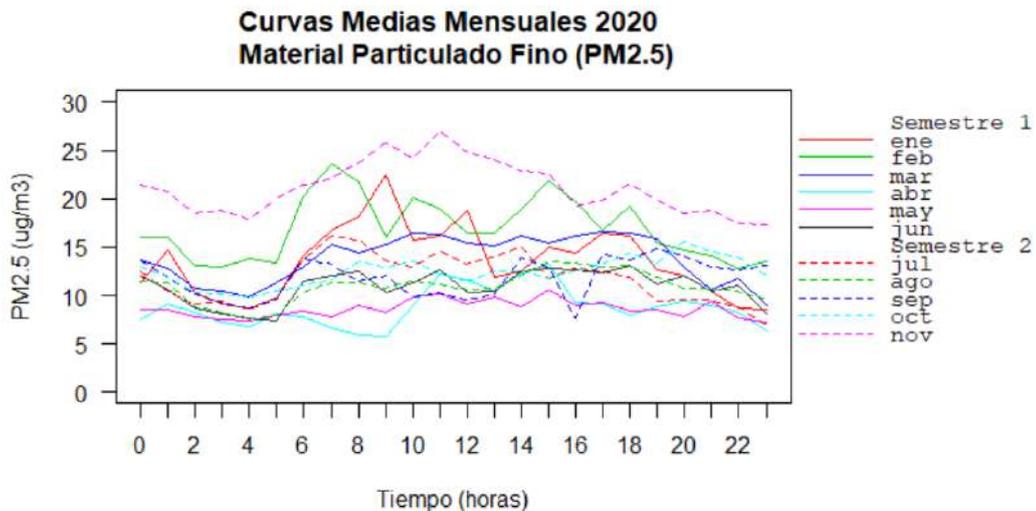


Figura 29-4: Curvas medias mensuales del material particulado fino para el año 2020

Realizado por: Allauca Fabián, 2021

4.1.5.3 Curva media funcional con entorno de confianza bootstrap al 95%

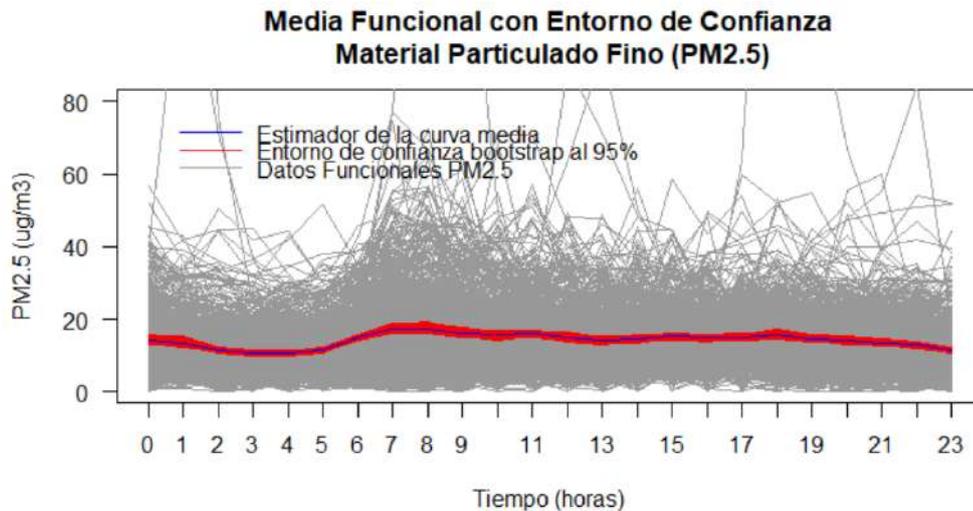


Figura 30-4: Estimador de curva media con entorno de confianza bootstrap al 95% del material particulado fino para los años 2019 y 2020.

Realizado por: Allauca Fabián, 2021

La media funcional para los datos del material particulado fino se puede apreciar en la Figura 30-4, como también el entorno de confianza bootstrap al 95%. La curva media mostrada en color azul indica una tendencia ligeramente constante entre los 18 a 20 $\mu\text{g}/\text{m}^3$ de PM2.5; tomando un pequeño incremento entre las 07h00 a 08h00 y una pequeña disminución entre la 01h00 a 05h00.

La Organización Mundial de la Salud recomienda una concentración media diaria de PM2.5 de 25 $\mu\text{g}/\text{m}^3$, para que no exista daño a la salud de las personas. (Organización Mundial de la Salud, 2021). De acuerdo a la media funcional obtenida en esta investigación, la curva media no sobrepasa este valor. La temperatura ambiente y la presión atmosférica promedio a la que fueron tomados los datos son de 14.86°C y 0.718 atm. respectivamente.

4.2 Análisis por períodos

Para tener un mejor entendimiento del comportamiento de los datos de los cinco contaminantes descritos anteriormente se procede a realizar un análisis por períodos, los mismos que serán; períodos escolares y vacaciones escolares, estos períodos se los ha dividido de acuerdo al Ministerio de Educación, quienes nos indican las fechas de los periodos escolares, siendo estas para el año lectivo 2018-2019 del 3 de septiembre de 2018 al 5 de julio de 2019 y para el año lectivo 2019-2020 del 2 de septiembre de 2019 al 3 de julio de 2020 (Ministerio de Educación, 2021).

También se ha decidido realizar una división de los datos de los cinco contaminante por los períodos pandemia y no pandemia, empezando el periodo de pandemia desde el 12 de marzo hasta el 13 de septiembre de 2020, fecha en la cual el Gobierno de la República del Ecuador realizó el confinamiento por covid-19 (Ministerio de Salud Pública, 2021)

Y finalmente se decide realizar una división más de los datos por períodos de lunes-viernes y sábados-domingos, con la finalidad de observar el comportamiento de los datos de los contaminantes en estos días, que corresponden a días laborables y días de descanso de las jornadas laborales normales.

4.2.1 Monóxido de carbono (CO)

4.2.1.1 Períodos escolares y vacaciones escolares

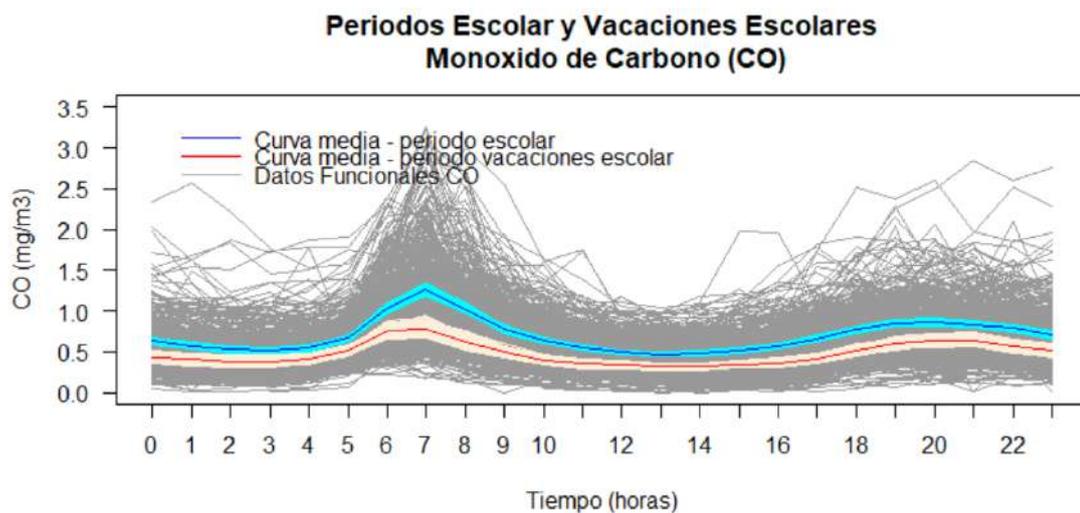


Figura 31-4: Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el monóxido de carbono

Realizado por: Allauca Fabián, 2021

La Figura 31-4 muestra la comparación entre medias funcionales para los períodos escolares y vacaciones escolares, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período vacaciones escolares es menor que la media del periodo escolar a lo largo de todo el día; esto nos indica que existe menor concentración de monóxido de carbono para los meses de julio y agosto que son los meses del periodo de vacaciones escolares; las dos curvas medias tienen la misma tendencia de distribución, siendo sus valores más altos alrededor de las siete de la mañana; alcanzando un valor máximo de 1.25 mg/m³ para el periodo escolar y de 0.8 mg/m³ para el periodo vacaciones escolares.

Este valor más bajo en el periodo vacaciones escolares nos da un indicio de que existe menor circulación vehicular en los meses de julio y agosto, esto se puede indicar ya que el monóxido de carbono se forma por la combustión incompleta por déficit de oxígeno, cuya principal fuente son los vehículos automotores. (Organización Panamericana de la Salud, 2005)

4.2.1.2 Períodos pandemia y no pandemia

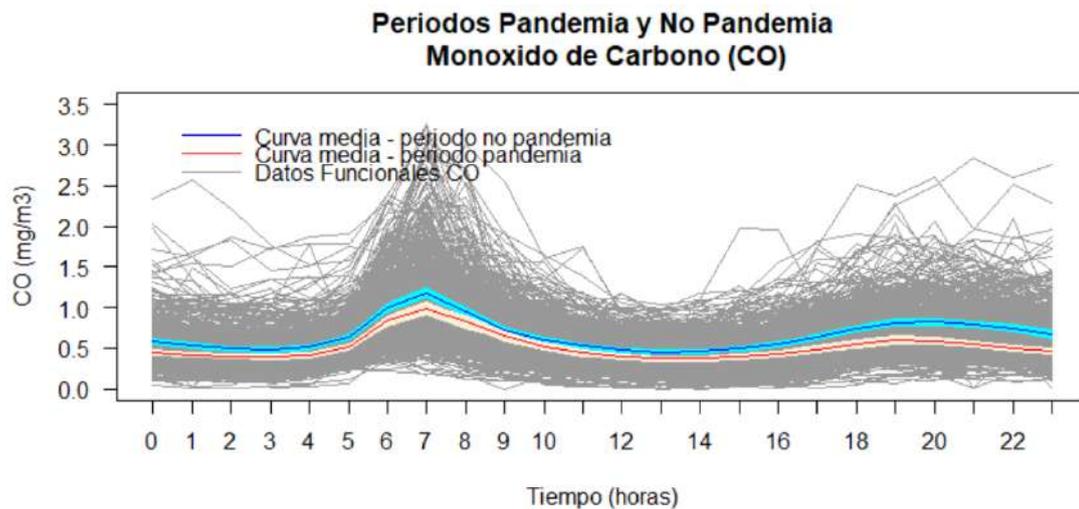


Figura 32-4: Comparación de medias funcionales entre los periodos pandemia y no pandemia para el monóxido de carbono

Realizado por: Allauca Fabián, 2021

La Figura 32-4 muestra la comparación entre medias funcionales para los periodos pandemia y no pandemia, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del periodo pandemia es menor que la media del periodo no pandemia a lo largo de todo el día; esto nos indica que existe menor concentración de monóxido de carbono para los meses de marzo, abril, mayo, junio, julio, agosto y septiembre de 2020, meses que duró el confinamiento por covid-19; las dos curvas medias tienen la misma tendencia de distribución, siendo sus valores más altos alrededor de las siete de la mañana; alcanzando un valor máximo de 1.15 mg/m³ para el periodo no pandemia y de 1.0 mg/m³ para el periodo pandemia.

Este valor más bajo en el periodo pandemia se debe a que hubo restricción vehicular para los meses en que duró el confinamiento por covid-19, resultando concentraciones menores de monóxido de carbono, lo que se verifica en la Figura 32-4.

4.2.1.3 Períodos lunes-viernes y sábados-domingos

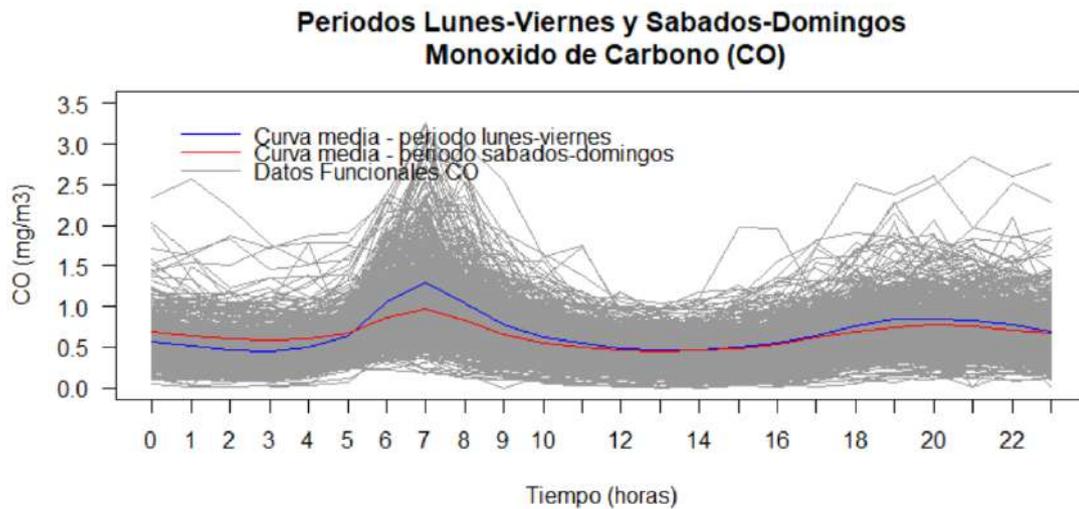


Figura 33-4: Comparación de medias funcionales entre los períodos lunes-viernes y sábados-domingos para el monóxido de carbono

Realizado por: Allauca Fabián, 2021

La Figura 33-4 muestra la comparación entre medias funcionales para los períodos lunes-viernes y sábados-domingos, en la Figura 33-4 no se muestra la banda de confianza bootstrap al 95% puesto que las curvas medias se traslapan entre las 12:00 a 17:00, impidiéndonos tener una buena observación de las curvas medias; en la gráfica se puede apreciar que las curvas medias para los dos períodos son casi similares, existiendo una diferencia significativa en las primeras horas de la mañana de 00:00 a 10:00. Cabe mencionar que las curvas medias para los dos períodos tienen un pico más alto alrededor de las 7 de la mañana, alcanzando un valor de 1.4 mg/m^3 para el período de lunes-viernes y de 0.98 mg/m^3 para el período sábados-domingos.

También se puede observar en la gráfica que la curva media para el período sábados-domingos es mayor en comparación con el período lunes-viernes en las horas de 00:00 a 05:00 dándonos un indicio de que existe mayor circulación vehicular los fines de semana a esas horas del día, esto se puede indicar ya que el monóxido de carbono se forma por la combustión incompleta por déficit de oxígeno, cuya principal fuente son los vehículos automotores. (Organización Panamericana de la Salud, 2005)

4.2.2 Dióxido de azufre (SO₂)

4.2.1.1 Períodos escolares y vacaciones escolares

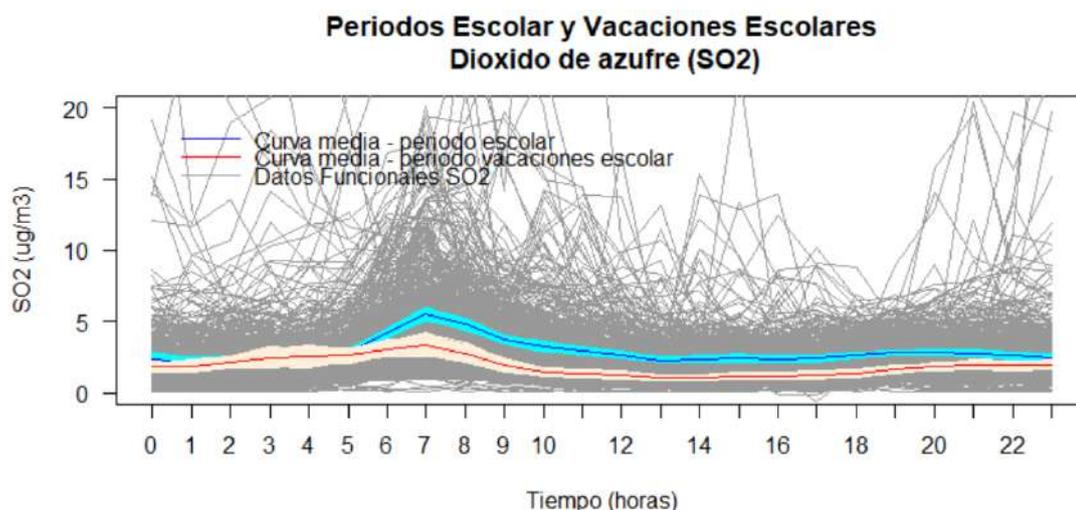


Figura 34-4: Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el dióxido de azufre

Realizado por: Allauca Fabián, 2021

La Figura 34-4 muestra la comparación entre medias funcionales para los períodos escolares y vacaciones escolares, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período vacaciones escolares es menor que la media del periodo escolar en gran parte del día, con excepción de 01:00 a 05:00, donde se aprecia que las dos curvas medias toman valores similares; esto nos indica que existe menor concentración de dióxido de azufre para los meses de julio y agosto que son los meses del periodo de vacaciones escolares; las dos curvas medias tienen la misma tendencia de distribución, siendo sus valores más altos alrededor de las siete de la mañana; alcanzando un valor máximo de $5.5 \mu\text{g}/\text{m}^3$ para el periodo escolar y de $3.5 \mu\text{g}/\text{m}^3$ para el periodo vacaciones escolares.

Este valor más bajo en el periodo vacaciones escolares se puede deber a la menor circulación de vehículos propulsados por motores diésel, como es el caso de buses y busetas escolares, principales fuentes generadoras de dióxido de azufre según estudios realizados por la CORPAIRE (Corporación para el Mejoramiento del Aire de Quito, 2006)

4.2.1.2 Períodos pandemia y no pandemia

La Figura 35-4 muestra la comparación entre medias funcionales para los períodos pandemia y no pandemia, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período pandemia es menor que la media del periodo no pandemia a lo largo de todo el día; esto nos indica que existe menor concentración de dióxido de azufre para los meses de marzo, abril, mayo, junio, julio, agosto y septiembre de 2020, meses que duró el confinamiento por covid-19; las dos curvas medias tienen la misma tendencia

de distribución, siendo sus valores más altos alrededor de las siete de la mañana; alcanzando un valor máximo de $5.5 \mu\text{g}/\text{m}^3$ para el periodo no pandemia y de $4.0 \mu\text{g}/\text{m}^3$ para el periodo pandemia.

Este valor más bajo en el periodo pandemia se debe a que hubo restricción vehicular para los meses en que duró el confinamiento por covid-19; los buses no circularon la gran mayoría de meses de confinamiento, mientras que las busetas escolares no circularon durante todo el confinamiento, cabe recalcar que los vehículos propulsados por motores diésel son los principales generadores de este contaminante. (Corporación para el Mejoramiento del Aire de Quito, 2006)

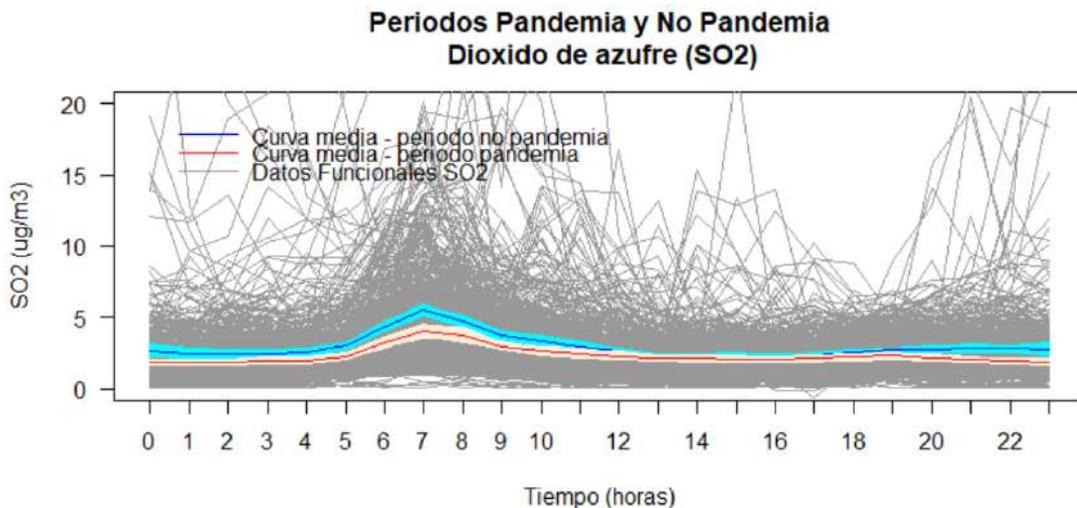


Figura 35-4: Comparación de medias funcionales entre los periodos pandemia y no pandemia para el dióxido de azufre

Realizado por: Allauca Fabián, 2021

4.2.1.3 Períodos lunes-viernes y sábados-domingos

La Figura 36-4 muestra la comparación entre medias funcionales para los períodos lunes-viernes y sábados-domingos, en la Figura 36-4 no se muestra la banda de confianza bootstrap al 95% puesto que las curvas medias se traslapan entre las 00:00 a 05:00 y entre las 12:00 a 15:00, impidiéndonos tener una buena observación de las curvas medias; en la gráfica se puede apreciar que las curvas medias para los dos períodos son casi similares, existiendo una diferencia significativa alrededor de las 05:00 a 12:00. Cabe mencionar que las curvas medias para los dos períodos tienen un pico más alto alrededor de las 7 de la mañana, alcanzando un valor de $6.0 \mu\text{g}/\text{m}^3$ para el período de lunes-viernes y de $4.0 \mu\text{g}/\text{m}^3$ para el período sábados-domingos.

También se puede observar en la gráfica que la curva media para el período sábados-domingos es menor en comparación con el período lunes-viernes en las horas de 05:00 a 12:00 dándonos un

indicio de que existe menor circulación vehicular los fines de semana a esas horas del día, sobre todo para buses y busetas escolares; principales generadoras de este contaminante.

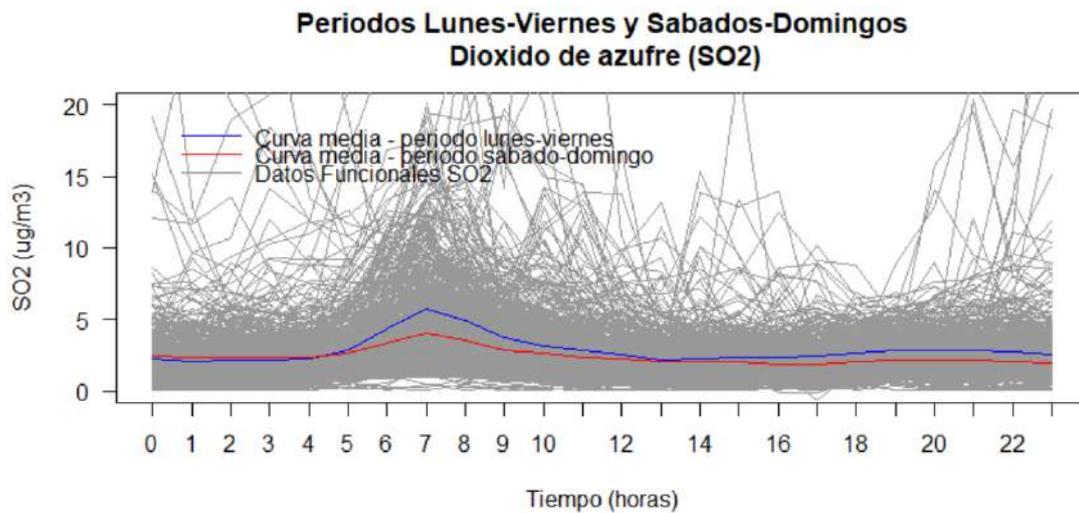


Figura 36-4: Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el dióxido de azufre

Realizado por: Allauca Fabián, 2021

4.2.3 Dióxido de nitrógeno (NO2)

4.2.1.1 Períodos escolares y vacaciones escolares

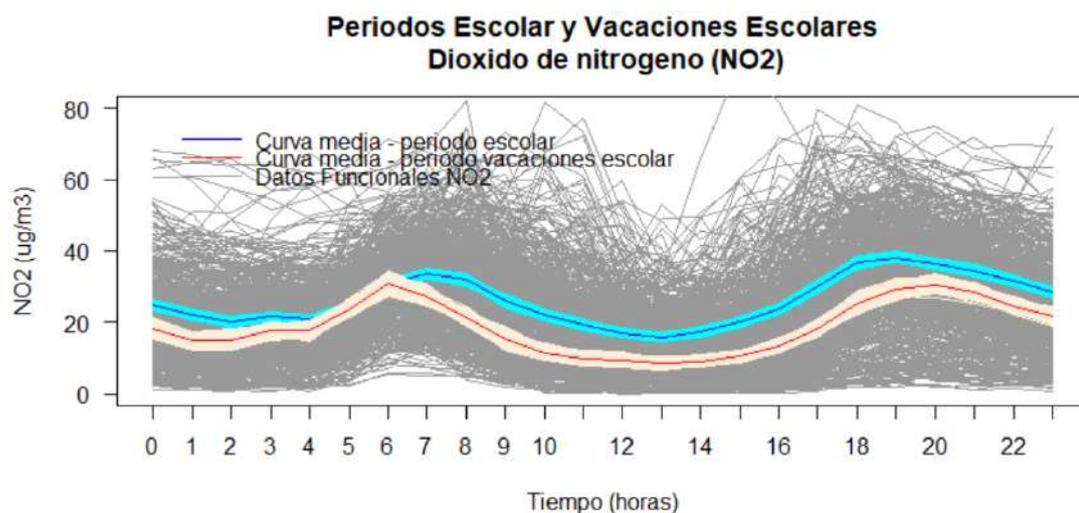


Figura 37-4: Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el dióxido de nitrógeno

Realizado por: Allauca Fabián, 2021

La Figura 37-4 muestra la comparación entre medias funcionales para los períodos escolares y vacaciones escolares, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período vacaciones escolares es menor que la media del periodo escolar, con excepción de 05:00 a 06:00 donde las curvas toman valores similares; esto nos indica que existe menor concentración de dióxido de nitrógeno para los meses de julio y agosto que son los meses del periodo de vacaciones escolares; las dos curvas medias tienen la misma tendencia de distribución, siendo sus valores más altos alrededor de las siete de la mañana y siete de la noche.

Este valor más bajo en el periodo vacaciones escolares nos da un indicio de que existe menor circulación vehicular en los meses de julio y agosto, sobre todo para vehículos con motor diésel, ya que ellos aportan con el 42% de NO₂, mientras que los vehículos livianos aportan el 35% de este contaminante (Corporación para el Mejoramiento del Aire de Quito, 2006).

4.2.1.2 Períodos pandemia y no pandemia

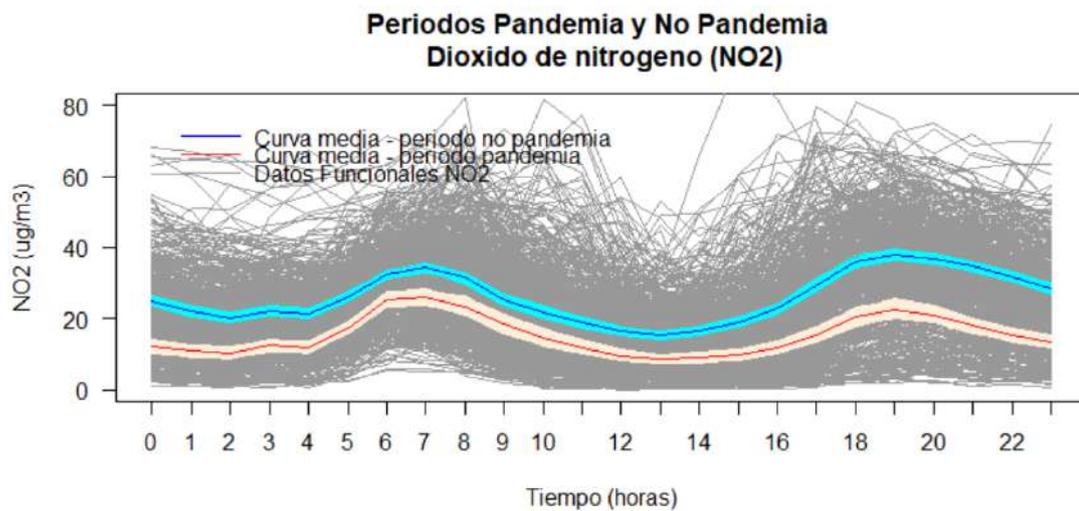


Figura 38-4: Comparación de medias funcionales entre los periodos pandemia y no pandemia para el dióxido de nitrógeno

Realizado por: Allauca Fabián, 2021

La Figura 38-4 muestra la comparación entre medias funcionales para los períodos pandemia y no pandemia, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período pandemia es mucho menor que la media del periodo no pandemia a lo largo de todo el día; esto nos indica que existe menor concentración de dióxido de nitrógeno para los meses de marzo, abril, mayo, junio, julio, agosto y septiembre de 2020, meses que duró el confinamiento por covid-19; las dos curvas medias

tienen la misma tendencia de distribución, siendo sus valores más altos alrededor de las siete de la mañana y siete de la noche.

Este valor más bajo en el periodo pandemia se debe a que hubo restricción vehicular para los meses en que duró el confinamiento por covid-19, resultando concentraciones menores de dióxido de nitrógeno, lo que se verifica en la Figura 38-4.

4.2.1.3 Períodos lunes-viernes y sábados-domingos

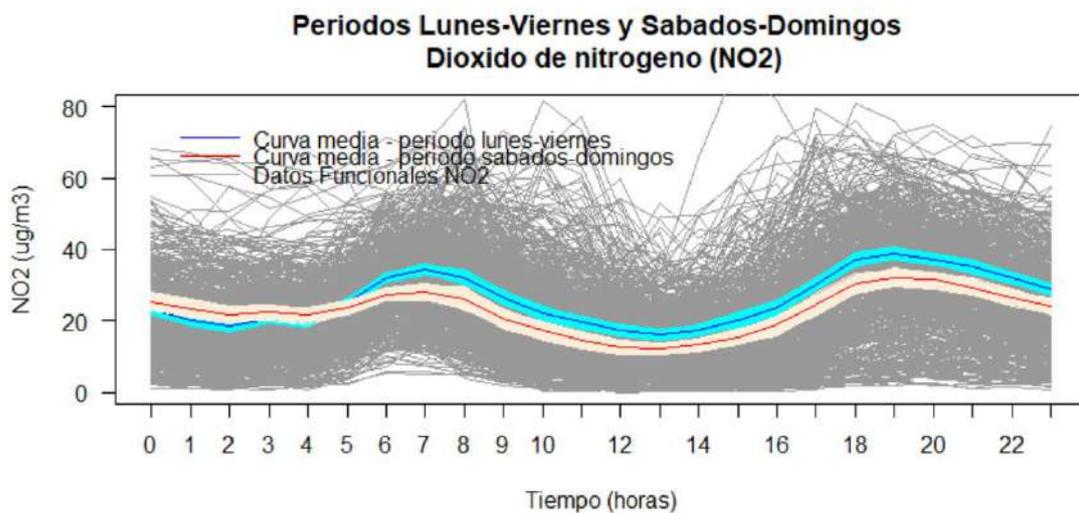


Figura 39-4: Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el dióxido de nitrógeno

Realizado por: Allauca Fabián, 2021

La Figura 39-4 muestra la comparación entre medias funcionales para los periodos lunes-viernes y sábados-domingos, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período sábados-domingos es menor que la media del periodo lunes-viernes, con excepción de 00:00 a 05:00; esto nos indica que existe menor concentración de dióxido de nitrógeno los fines de semana de 05:00 a 23:00; las dos curvas medias tienen la misma tendencia de distribución, siendo sus valores más altos alrededor de las siete de la mañana y siete de la noche.

De 00:00 a 05:00 la curva media del período sábados-domingos es mayor que la curva media del período lunes-viernes dándonos un indicio de que existe mayor circulación vehicular los fines de semana a esas horas del día, esto se puede indicar ya que el dióxido de nitrógeno, la mayor parte tiene su origen está en la oxidación del NO que se produce en la combustión de los motores de los vehículos, fundamentalmente los diésel. (Gobierno de España, 2021)

4.2.4 Ozono (O3)

4.2.1.1 Períodos escolares y vacaciones escolares

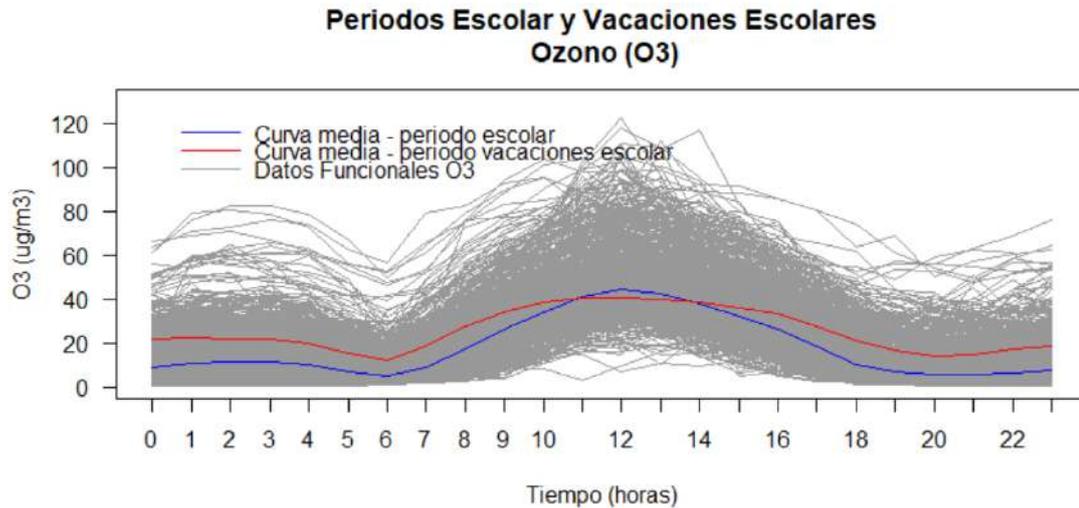


Figura 40-4: Comparación de medias funcionales entre los períodos escolares y vacaciones escolares para el ozono

Realizado por: Allauca Fabián, 2021

La Figura 40-4 muestra la comparación entre medias funcionales para los períodos escolares y vacaciones escolares, se puede apreciar que las dos medias muestran una tendencia similar para sus datos, esto se puede deber a que el ozono se forma por la reacción con la luz solar (fotoquímica) de contaminantes como los óxidos de nitrógeno (NO_x) procedentes de las emisiones de vehículos o la industria y los compuestos orgánicos volátiles (COV) emitidos por los vehículos, los disolventes y la industria. Los niveles de ozono más elevados se registran durante los períodos de tiempo soleado. (Organización Mundial de la Salud, 2021).

4.2.1.2 Períodos pandemia y no pandemia

La Figura 41-4 muestra la comparación entre medias funcionales para los períodos pandemia y no pandemia, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período pandemia es mayor que la media del periodo no pandemia a lo largo de todo el día; esto nos indica que existe mayor concentración de ozono para los meses de marzo, abril, mayo, junio, julio, agosto y septiembre de 2020, meses que duró el confinamiento por covid-19.

Este valor más alto en el periodo pandemia, nos da un indicio de que existió mayor radiación solar para los meses en que duró el confinamiento por covid-19, ya que el ozono se forma por la reacción con la luz solar.

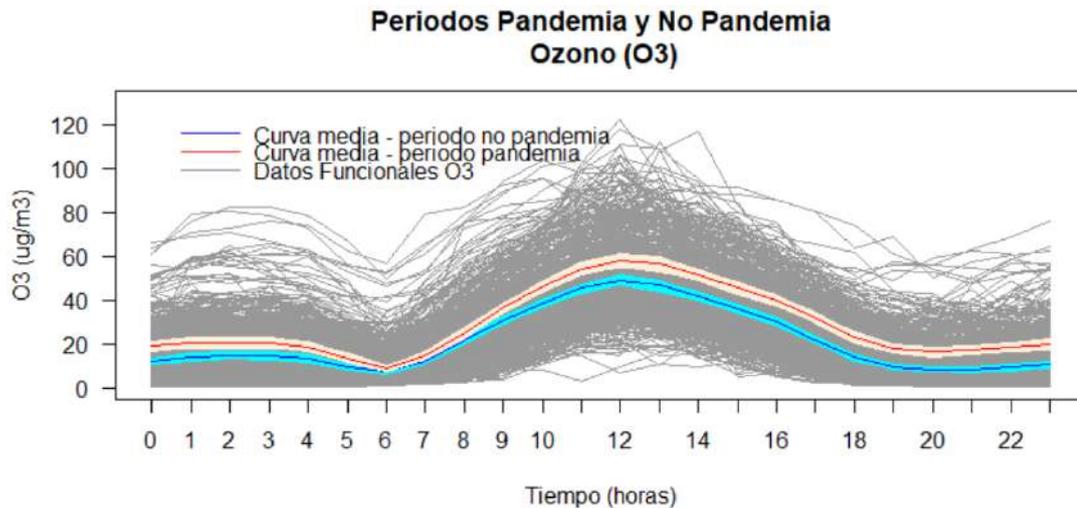


Figura 41-4: Comparación de medias funcionales entre los periodos pandemia y no pandemia para el ozono

Realizado por: Allauca Fabián, 2021

4.2.1.3 Períodos lunes-viernes y sábados-domingos

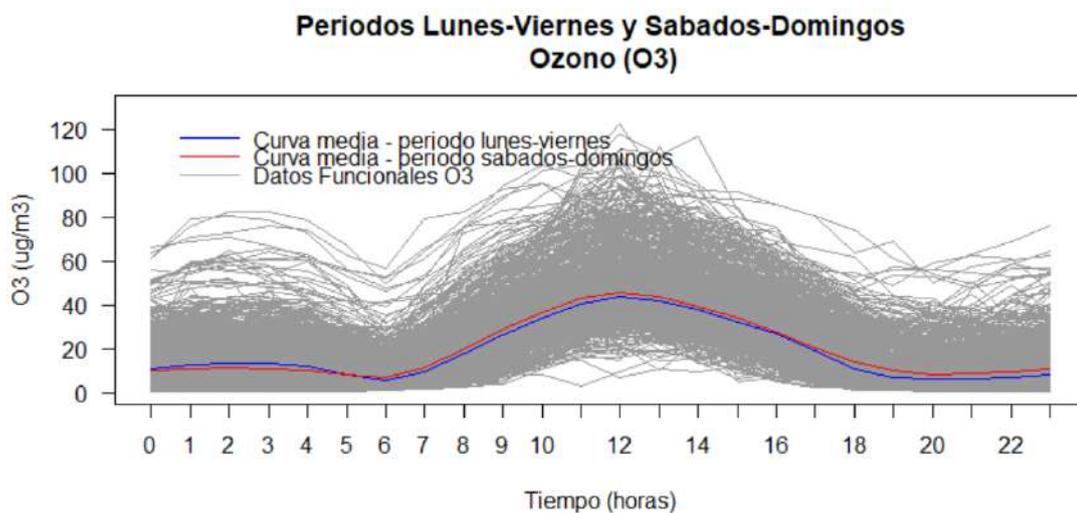


Figura 42-4: Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el ozono

Realizado por: Allauca Fabián, 2021

La Figura 42-4 muestra la comparación entre medias funcionales para los períodos lunes-viernes y sábados-domingos, en la Figura 42-4 no se muestra la banda de confianza bootstrap al 95% puesto que las curvas medias se traslapan a lo largo de todo el día, impidiéndonos tener una buena

observación de las curvas medias; en la gráfica se puede apreciar que no existe diferencia significativa entre las curvas de los dos períodos, esto nos da como indicio de que existe similar radiación solar de lunes a viernes, comparada con sábados y domingos.

4.2.5 Material particulado fino (PM2.5)

4.2.1.1 Períodos escolares y vacaciones escolares

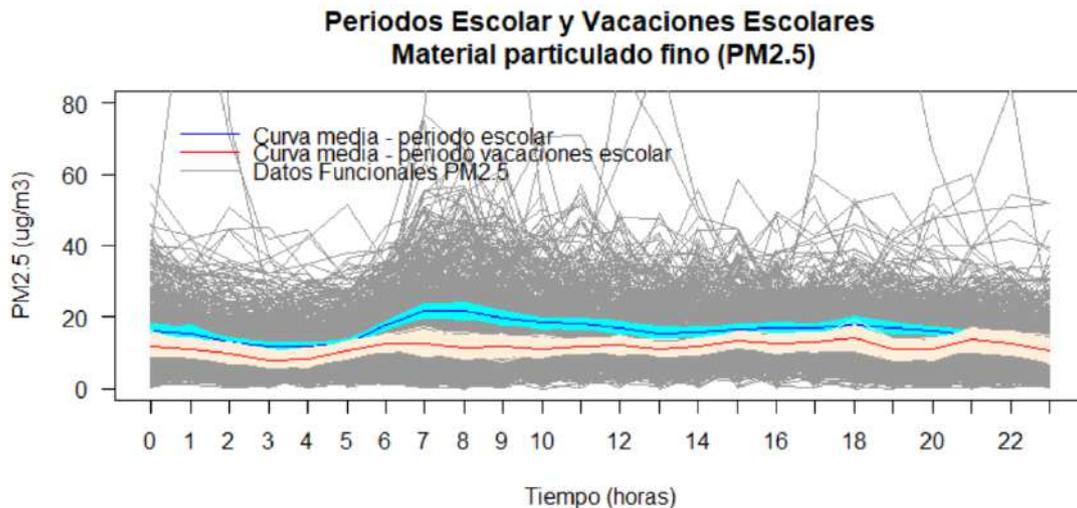


Figura 43-4: Comparación de medias funcionales entre los periodos escolares y vacaciones escolares para el material particulado fino

Realizado por: Allauca Fabián, 2021

La Figura 43-4 muestra la comparación entre medias funcionales para los períodos escolares y vacaciones escolares, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período vacaciones escolares es menor que la media del periodo escolar a lo largo de todo el día; esto nos indica que existe menor concentración de material particulado fino para los meses de julio y agosto que son los meses del periodo de vacaciones escolares; las dos curvas medias tienen similar tendencia de distribución, siendo sus valores más altos alrededor de las siete de la mañana; alcanzando un valor máximo de $22 \mu\text{g}/\text{m}^3$ para el periodo escolar y de $14 \mu\text{g}/\text{m}^3$ para el periodo vacaciones escolares.

Este valor más bajo en el periodo vacaciones escolares nos da un indicio de que existe menor circulación vehicular en los meses de julio y agosto, sobre todo para vehículos con motor diésel, ya que ellos aportan con el 38% de PM2.5, mientras que los vehículos livianos aportan el 8% de este contaminante, de acuerdo a estudios de la CORPAIRE para 2006 (Corporación para el Mejoramiento del Aire de Quito, 2006).

4.2.1.2 Períodos pandemia y no pandemia

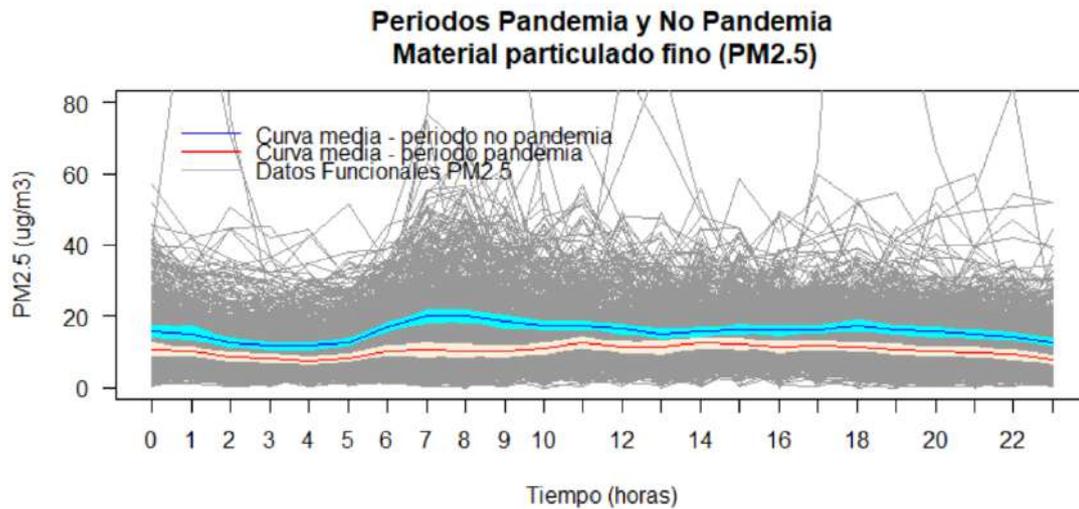


Figura 44-4: Comparación de medias funcionales entre los periodos pandemia y no pandemia para el material particulado fino

Realizado por: Allauca Fabián, 2021

La Figura 44-4 muestra la comparación entre medias funcionales para los períodos pandemia y no pandemia, cada media funcional está acompañada con su entorno de confianza bootstrap al 95%; se puede apreciar que la media funcional del período pandemia es menor que la media del periodo no pandemia a lo largo de todo el día; esto nos indica que existe menor concentración de material particulado fino para los meses de marzo, abril, mayo, junio, julio, agosto y septiembre de 2020, meses que duró el confinamiento por covid-19; las dos curvas medias tienen similar tendencia de distribución.

Este valor más bajo en el periodo pandemia se debe a que hubo restricción vehicular para los meses en que duró el confinamiento por covid-19, resultando concentraciones menores de material particulado fino, lo que se verifica en la Figura 44-4.

4.2.1.3 Períodos lunes-viernes y sábados-domingos

La Figura 45-4 muestra la comparación entre medias funcionales para los períodos lunes-viernes y sábados-domingos, en la Figura 45-4 no se muestra la banda de confianza bootstrap al 95% puesto que las curvas medias se traslapan entre las 00:00 a 05:00 y entre las 11:00 a 23:00, impidiéndonos tener una buena observación de las curvas medias; en la gráfica se puede apreciar que las curvas medias para los dos períodos son casi similares, existiendo una diferencia significativa entre las 05:00 a 11:00. Cabe mencionar que la curva media para el período lunes-viernes tienen un pico alrededor de las 7 de la mañana, alcanzando un valor de $22 \mu\text{g}/\text{m}^3$, mientras

que para el período sábados-domingos la curva media es casi horizontal. Esta diferencia alrededor de las 05:00 a 11:00, nos da un indicio de que existe menor circulación vehicular los fines de semana en el transcurso de esas horas.

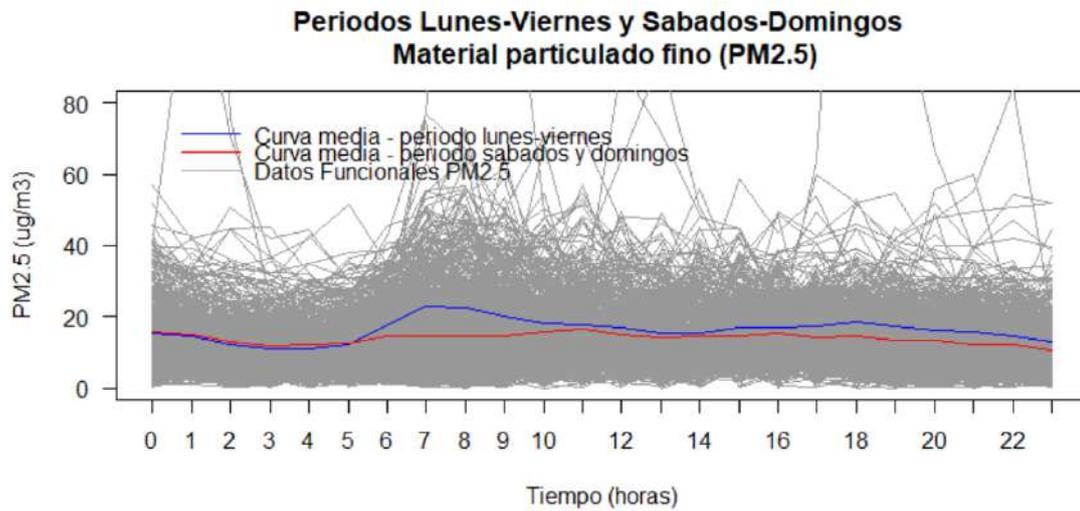


Figura 45-4: Comparación de medias funcionales entre los periodos lunes-viernes y sábados-domingos para el material particulado fino

Realizado por: Allauca Fabián, 2021

CAPÍTULO V

5. PROPUESTA

Cuatro de los cinco contaminantes analizados descriptivamente en el Capítulo IV tienen como su principal fuente de emisión a los vehículos automotores, estos contaminantes son: monóxido de carbono (CO), dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂) y material particulado fino (PM_{2.5}). (Córdoba & Ramos, 2001)

El material particulado fino tiene como componentes principales los sulfuros, los nitratos, el amoníaco, el cloruro de sodio, el hollín, los polvos minerales y el agua. (Organización Panamericana de la Salud, 2005). Para observar la relación que existe entre el PM_{2.5} con el CO, SO₂ y NO₂ se realiza un modelo lineal funcional con respuesta escalar para estas variables.

También se realiza un modelo lineal funcional con respuesta escalar de la variable media diaria del ozono en función de la variable funcional dióxido de nitrógeno, con el objetivo de observar la relación que existe entre estos dos contaminantes; debido a que el ozono se forma por la reacción con la luz solar (fotoquímica) de contaminantes como los óxidos de nitrógeno (NO_x) procedentes de las emisiones de vehículos o la industria y los compuestos orgánicos volátiles (COV) emitidos por los vehículos, los disolventes y la industria. (Organización Mundial de la Salud, 2021)

5.1 Ajuste de modelos lineales funcionales con respuesta escalar de la variable: media diaria del material particulado fino (variable escalar “y”), en función de las variables funcionales: monóxido de carbono, dióxido de azufre y dióxido de nitrógeno.

Modelo de regresión lineal funcional: $y = \langle X, \beta \rangle + \varepsilon$, donde $X \in \mathcal{L}_2(T)$ e $y \in \mathbb{R}$

5.1.1 Modelo lineal de la variable escalar “y” en función de la variable funcional “monóxido de carbono”

Se realiza la gráfica del monóxido de carbono (CO) y su representación en bases de Fourier, por las peculiaridades de periodicidad que representan las curvas, es decir aproximadamente a las 00:00 horas tiene valores similares que a las 23:00 horas. Este criterio se toma para las demás variables funcionales como es el dióxido de azufre y dióxido de nitrógeno.

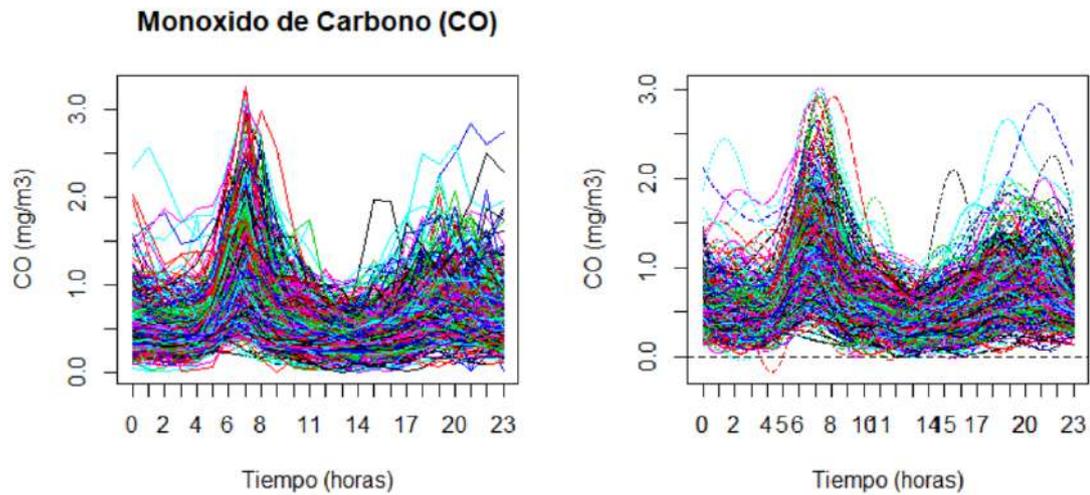


Figura 1-5: Curvas diarias del monóxido de carbono para los años 2019 y 2020 (izquierda) y su representación en bases de Fourier (derecha)

Realizado por: Allauca Fabián, 2021

Modelo \rightarrow y: media diaria del material particulado fino \sim monóxido de carbono

Bases de representación: monóxido de carbono y beta se utilizan bases de Fourier

Resultados: los resultados se muestran en la Tabla 1-5

Tabla 1-5: Resultados del modelo lineal de la variable escalar “y” en función de la variable funcional “monóxido de carbono”

Coefficients	Estimate	Std. Error	t value	Pr(> t)	Multiple R-squared	p-value
Intercept	14.2136	0.1758	80.838	< 2e-16	0.3161	< 2.2e-16
CO.const	2.4914	0.2390	10.424	< 2e-16		
CO.sin1	-0.1275	0.4499	-0.283	0.77690		
CO.cos1	1.0354	0.6002	1.725	0.08494		
CO.sin2	-1.8298	0.6433	-2.845	0.00458		
CO.cos2	-0.7081	0.7586	-0.934	0.35088		
CO.sin3	1.4787	0.9726	1.520	0.12889		
CO.cos3	2.0893	1.0650	1.962	0.05020		

Realizado por: Allauca Pancho, Fabián, 2021

El porcentaje de variabilidad de la “media diaria del material particulado fino” explicada por el monóxido de carbono es de 31,61%, con un p-valor <0.001

5.1.2 Modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de azufre”

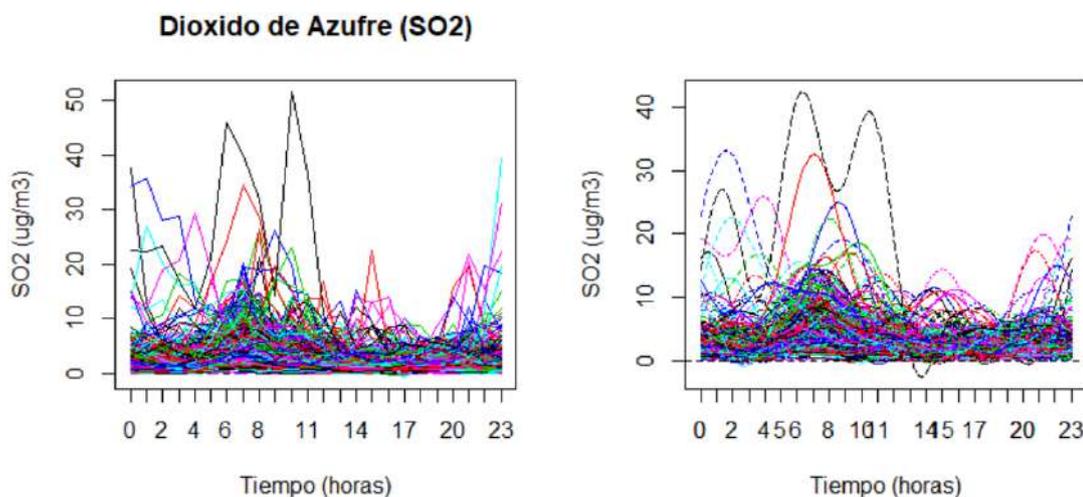


Figura 2-5: Curvas diarias del dióxido de azufre para los años 2019 y 2020 (izquierda) y su representación en bases de Fourier (derecha)

Realizado por: Allauca Fabián, 2021

Modelo \rightarrow y: media diaria del material particulado fino \sim dióxido de azufre

Bases de representación: para las variables dióxido de azufre y beta se utilizan bases de Fourier

Resultados: los resultados se muestran en la Tabla 2-5

Tabla 2-5: Resultados del modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de azufre”

Coefficients	Estimate	Std. Error	t value	Pr(> t)	Multiple R-squared	p-value
Intercept	14.21364	0.17131	82.972	<2e-16	0.3508	< 2.2e-16
SO2.const	0.47868	0.03751	12.760	<2e-16		
SO2.sin1	-0.10473	0.05809	-1.803	0.0719		
SO2.cos1	-0.14489	0.05820	-2.490	0.0130		
SO2.sin2	-0.16186	0.08079	-2.003	0.0455		
SO2.cos2	0.21434	0.09214	2.326	0.0203		
SO2.sin3	-0.01902	0.10834	-0.176	0.8607		
SO2.cos3	0.24265	0.12525	1.937	0.0531		

Realizado por: Allauca Pancho, Fabián, 2021

El porcentaje de variabilidad de la “media diaria del material particulado fino” explicada por el dióxido de azufre es de 35,08%, con un p-valor <0.001

5.1.3 Modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de nitrógeno”

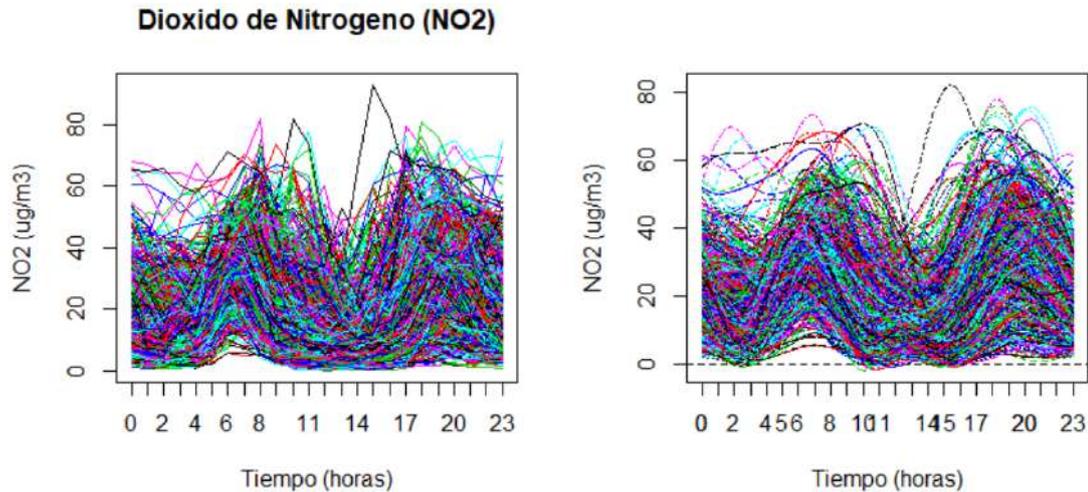


Figura 3-5: Curvas diarias del dióxido de nitrógeno para los años 2019 y 2020 (izquierda) y su representación en bases de Fourier (derecha)

Realizado por: Allauca Fabián, 2021

Modelo \rightarrow y: media diaria del material particulado fino \sim dióxido de nitrógeno

Bases de representación: para las variables dióxido de nitrógeno y beta se utilizan bases de Fourier

Resultados: los resultados se muestran en la Tabla 3-5

Tabla 3-5: Resultados del modelo lineal de la variable escalar “y” en función de la variable funcional “dióxido de nitrógeno”

Coefficients	Estimate	Std. Error	t value	Pr(> t)	Multiple R-squared	p-value
Intercept	14.213636	0.138679	102.493	< 2e-16	0.5745	< 2.2e-16
NO2.const	0.092047	0.004274	21.538	< 2e-16		
NO2.sin1	0.014465	0.008747	1.654	0.098635		
NO2.cos1	0.032305	0.009272	3.484	0.000525		
NO2.sin2	-0.008000	0.012406	-0.645	0.519269		
NO2.cos2	0.047038	0.013460	3.495	0.000505		
NO2.sin3	0.081503	0.016192	5.033	6.15e-07		
NO2.cos3	-0.003100	0.020832	-0.149	0.881749		

Realizado por: Allauca Pancho, Fabián, 2021

El porcentaje de variabilidad de la “media diaria del material particulado fino” explicada por el dióxido de nitrógeno es de 57,45%, con un p-valor <0.001

5.1.4 Modelo lineal múltiple de la variable escalar “y” en función de las variables funcionales: monóxido de carbono, dióxido de azufre y dióxido de nitrógeno

Para el modelo lineal múltiple: $y \sim$ monóxido de carbono + dióxido de azufre + dióxido de nitrógeno

Bases de representación: para beta y las variables funcionales se utilizan bases de Fourier

Resultados: los resultados se muestran en la Tabla 4-5

Tabla 4-5: Resultados del modelo lineal múltiple de la variable escalar “y” en función de las variables funcionales: monóxido de carbono, dióxido de azufre y dióxido de nitrógeno

Coefficients	Estimate	Std. Error	t value	Pr(> t)	Multiple R-squared	p-value
Intercept	14.213636	0.134424	105.737	< 2e-16	0.6083	< 2.2e-16
CO.const	-0.019288	0.236666	-0.082	0.935068		
CO.sin1	0.371501	0.519027	0.716	0.474383		
CO.cos1	1.909037	0.653201	2.923	0.003587		
CO.sin2	1.131316	0.718535	1.574	0.115844		
CO.cos2	-1.111319	0.851607	-1.305	0.192347		
CO.sin3	-1.977002	1.045004	-1.892	0.058935		
CO.cos3	1.711389	1.077724	1.588	0.112760		
SO2.const	0.048539	0.037634	1.290	0.197576		
SO2.sin1	0.064468	0.053838	1.197	0.231552		
SO2.cos1	-0.120421	0.051636	-2.332	0.019987		
SO2.sin2	0.052282	0.075896	0.689	0.491145		
SO2.cos2	0.296361	0.088161	3.362	0.000818		
SO2.sin3	-0.048125	0.093822	-0.513	0.608166		
SO2.cos3	0.272813	0.109701	2.487	0.013127		
NO2.const	0.080248	0.005975	13.430	< 2e-16		
NO2.sin1	-0.018238	0.013491	-1.352	0.176881		
NO2.cos1	0.001323	0.014213	0.093	0.925890		
NO2.sin2	-0.010344	0.017401	-0.594	0.552405		
NO2.cos2	0.048440	0.019097	2.537	0.011418		
NO2.sin3	0.085932	0.023181	3.707	0.000227		
NO2.cos3	-0.055678	0.027349	-2.036	0.042154		

Realizado por: Allauca Pancho, Fabián, 2021

Con un p-valor <0.001 , el porcentaje de variabilidad de la “media diaria del material particulado fino” explicada por las variables funcionales monóxido de carbono, dióxido de azufre y dióxido de nitrógeno es de un 60,83%.

5.2 Ajuste del modelo lineal funcional con respuesta escalar de la variable: media diaria del ozono (variable escalar “y”), en función de la variable funcional: dióxido de nitrógeno.

Modelo de regresión lineal funcional: $y = \langle X, \beta \rangle + \varepsilon$, donde $X \in \mathcal{L}_2(T)$ e $y \in \mathbb{R}$

Modelo \rightarrow y: media diaria del ozono \sim dióxido de nitrógeno

Bases de representación: para las variables dióxido de nitrógeno y beta se utilizan bases de Fourier

Resultados: los resultados se muestran en la Tabla 5-5

Tabla 5-5: Resultados del modelo lineal de la variable escalar “media diaria del ozono” en función de la variable funcional “dióxido de nitrógeno”

Coefficients	Estimate	Std. Error	t value	Pr(> t)	Multiple R-squared	p-value
Intercept	23.46956	0.31736	73.952	$< 2e-16$	0.3042	$< 2.2e-16$
NO2.const	-0.04557	0.00978	-4.659	$3.80e-06$		
NO2.sin1	0.15471	0.02002	7.729	$3.83e-14$		
NO2.cos1	-0.04392	0.02122	-2.070	0.038847		
NO2.sin2	-0.06124	0.02839	-2.157	0.031346		
NO2.cos2	-0.24400	0.03080	-7.921	$9.38e-15$		
NO2.sin3	-0.27114	0.03706	-7.317	$7.05e-13$		
NO2.cos3	-0.16322	0.04767	-3.424	0.000654		

Realizado por: Allauca Pancho, Fabián, 2021

El porcentaje de variabilidad de la “media diaria del ozono” explicada por el dióxido de nitrógeno es de 30,42%, con un p-valor <0.001

CONCLUSIONES

- En esta investigación se examinó los datos de la contaminación del aire proporcionados por la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ) para los años 2019 y 2020, los mismos que fueron ordenados en filas de manera cronológica para los días y en columnas para las horas del día y los datos faltantes fueron completados por medio de interpolación lineal, trabajo realizado en Microsoft Excel y guardados en extensión .txt para posteriormente ser leídos en el software estadístico R por medio de la función `read.table`
- Los datos de los contaminantes describieron un comportamiento irregular a lo largo del día; se observó picos alrededor de las 07:00 y 19:00 llegando a alcanzar aproximadamente los 3.26 mg/m^3 y 2.0 mg/m^3 respectivamente para el monóxido de carbono (CO). El dióxido de azufre (SO₂) describió un comportamiento casi regular a lo largo del día, se visualizó picos alrededor de las 07:00 llegando a alcanzar aproximadamente los $20 \text{ } \mu\text{g/m}^3$. El dióxido de nitrógeno (NO₂) describió una tendencia de doble montaña, alcanzando valores aproximados de $60 \text{ } \mu\text{g/m}^3$ para las dos cumbres, alrededor de las 07:00 a 08:00 y de 17:00 a 18:00. El ozono (O₃) describió un comportamiento en forma de campana, llegando a tomar valores aproximados de $100 \text{ } \mu\text{g/m}^3$ en la zona más alta. El material particulado fino (PM_{2.5}) describió una tendencia ligeramente constante alrededor de los $20 \text{ } \mu\text{g/m}^3$ a lo largo del día.
- La curva media para los datos del CO llegó a un valor máximo de 1.1 mg/m^3 a las 07:00 y un valor aproximado de 0.75 mg/m^3 entre las 19:00 y 20:00. La curva media para los datos del SO₂ llegó a un valor máximo de $5 \text{ } \mu\text{g/m}^3$ a las 07:00 y un valor aproximado de $2.6 \text{ } \mu\text{g/m}^3$ entre las 19:00 y 21:00. La curva media para los datos del NO₂ llegó a un valor máximo de $32 \text{ } \mu\text{g/m}^3$ alrededor de las 06:00 a 08:00 y un valor de $33 \text{ } \mu\text{g/m}^3$ alrededor de las 18:00 a 20:00. La curva media para los datos del O₃, llegó a un valor máximo de $55 \text{ } \mu\text{g/m}^3$ alrededor del mediodía. La curva media para los datos del PM_{2.5} llegó a un valor máximo de $20 \text{ } \mu\text{g/m}^3$ entre las 07:00 a 08:00. La temperatura ambiente y la presión atmosférica promedio a la que fueron tomados los datos de los cinco contaminantes fueron de 14.86°C y 0.718 atm . respectivamente.
- El análisis por periodos arrojó que existen diferencias significativas entre las medias funcionales de los cinco contaminantes, para los períodos de clases escolares y vacaciones escolares; periodos de pandemia y no pandemia y para los periodos de lunes-viernes y sábados domingos. Se observó que existe mayor concentración de estos contaminantes en los períodos de clases escolares, no pandemia y lunes-viernes, dándonos un indicio de que existió mayor circulación vehicular para estos períodos.

- Los modelos lineales funcionales con respuesta escalar de la variable: media diaria del PM_{2.5} en función de las variables funcionales: CO, SO₂ y NO₂, mostraron un porcentaje de variabilidad del 31.61%, 35.08% y del 57.45% respectivamente, lo que indica que el dióxido de nitrógeno es el contaminante que aporta mayormente a la formación de material particulado fino. Mientras que el porcentaje de variabilidad de la media diaria del PM_{2.5} explicada por las variables de los tres contaminantes mostró un porcentaje del 60,83%. El modelo lineal funcional con respuesta escalar de la variable: O₃ en función de la variable funcional NO₂, mostró un porcentaje de variabilidad del 30,42% con un p-valor < 0.001.

RECOMENDACIONES

- Se recomienda realizar análisis descriptivos funcionales para los contaminantes de las otras estaciones que monitorea la Red Metropolitana de Monitoreo Atmosférico de Quito (REMMAQ) como es el caso de Cotocollao, Carapungo, Belisario, Jipijapa, El Carmen, Guamaní, Tumbaco y Los Chillos.
- Realizar un análisis descriptivo funcional a la variable radiación solar, con la finalidad de realizar un modelo lineal funcional con respuesta escalar de la variable: ozono en función de la variable funcional radiación solar y observar la variabilidad que presentan estas dos variables, debido a que el ozono se forma por medio de la interacción de la luz solar.
- Elaborar una base de datos de la cantidad de vehículos que circulan por el centro histórico de Quito, con la finalidad de observar el grado de incidencia que tienen los vehículos en la contaminación del aire en las diferentes horas del día.
- Se recomienda impartir cursos de análisis funcional en carreras donde se manejen grandes cantidades de datos para que estos puedan ser analizados y entendidos de mejor manera.

BIBLIOGRAFÍA

- LUMITOS AG. (19 de Febrero de 2021). *QUIMICA.ES*. Obtenido de https://www.quimica.es/enciclopedia/%C3%93xido_de_azufre_%28IV%29.html
- Aguilera del Pino, M., & Aguilera Morillo, C. (2013). *Introducción al Análisis de Datos Funcionales con R*. Cádiz: Universidad de Cádiz. Dpto. Estadística e Investigación Operativa.
- Arias Odón, F. (2012). *El proyecto de investigación. Introducción a la metodología científica*. Caracas: Editorias Episteme.
- Benalcazar, G. (2005). *Espacios Normados y Aplicaciones*. Quito.
- Consejería de Salud de la Región de Murcia. (21 de Junio de 2021). *Murcia Salud*. Obtenido de <http://www.murciasalud.es/pagina.php?id=180398#>
- Córdoba, D., & Ramos, J. (2001). *Monóxido de Carbono*. Bogotá: Editorial el Manual Moderno.
- Corporación para el Mejoramiento del Aire de Quito. (2006). *Inventario de emisiones del Distrito Metropolitano de Quito - 2003*. Quito: PPL Impresores. 2529 762-Quito.
- Escudero Villa, A. I. (2014). *Análisis exploratorio funcional de los datos de radiación solar 2011 como soporte para la simulación de su comportamiento*. Riobamba: Escuela Superior Politécnica de Chimborazo.
- Espín Mayorga, M. E., & Veloz Cuichán, S. M. (2013). *Evaluación de los efectos de la Calidad del Aire en la productividad de los cultivos en los barrios La Morita, La Tola, El Arenal, La Esperanza y Collaquí ubicados en la parroquia de Tumbaco, cantón Quito, provincia de Pichincha*. Quito: Universidad Central del Ecuador.
- Foster, M., & Costa, D. (2005). Air Pollutants and the Respiratory Tract. *Lung Biology in Health and Disease*.
- Fundación Aequae. (3 de Febrero de 2021). *AQUAE FUNDACIÓN*. Obtenido de www.fundacionaequae.org/causas-contaminacion-ambiental/
- Gobierno de España. (19 de Febrero de 2021). *Registro Estatal de Emisiones y Fuentes Contaminantes*. Obtenido de <http://www.prtr-es.es/NOx-oxidos-de-nitrogeno,15595,11,2007.html>

- Grajales, T. (25 de Septiembre de 2017). *Tipos de Investigación*. Recuperado el 27 de 3 de 2000, de <http://tgrajales.net/investigipos.pdf>
- Granda, J. B. (2015). *Manual de metodología de la investigación científica*. Chimbote - Perú: Universidad Católica los Ángeles Chimbote.
- Guerrón Varela, E. R. (2015). *Análisis de datos meteorológicos del Valle de los Chillos usando datos funcionales*. Quito: Universidad San Francisco de Quito.
- Gutiérrez, J., Romieu, I., Corey, G., & Fortoul, T. (1997). *Contaminación del aire, riesgos para la salud*. México DF: UNAM/El Manual Moderno.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación*. México D.F.: McGraw-Hill / Interamericana Editores, S.A. de C.V.
- Instituto para la Salud Geoambiental. (19 de Febrero de 2021). *Instituto para la Salud Geoambiental - Un ser vivo, un entorno sano*. Obtenido de <https://www.saludgeoambiental.org/dioxido-nitrogeno-no2>
- Kreyszig, E. (1978). *Introductory Functional Analysis with Applications*. New York: University of Windsor.
- Meneses, E., Turtós, L., & Molina, E. (2004). Mejoras en la estimación de las externalidades de la generación eléctrica en Cuba. *Revista electrónica Ecosolar*, 7. Obtenido de <http://www.cubasolar.org/biblioteca/ecosolar.htm>
- Ministerio de Ambiente de Colombia . (19 de Febrero de 2021). *IDEAM - Instituto de Hidrología, Meteorología y Estudios Ambientales*. Obtenido de <http://www.ideam.gov.co/web/tiempo-y-clima/los-agentes-precursores-de-la-lluvia-acida>
- Ministerio de Educación. (24 de Junio de 2021). *Ministerio de Educación*. Obtenido de <https://educacion.gob.ec/calendario-escolar/>
- Ministerio de Salud Pública. (24 de Junio de 2021). *Ministerio de Salud Pública*. Obtenido de <https://www.salud.gob.ec/noticias/>
- Municipio del Distrito Metropolitano de Quito. (3 de Febrero de 2021). *Secretaría de Ambiente*. Obtenido de <http://www.quitoambiente.gob.ec/ambiente/index.php/politicas-y-planeacion-ambiental/red-de-monitoreo>
- Navarro Pérez, V. (2004). *Análisis de Datos Funcionales. Implementación y Aplicaciones*. Catalunya: Universidad Politécnica de Catalunya.

- Organización Mundial de la Salud. (22 de Junio de 2021). *Organización Mundial de la Salud*.
Obtenido de [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)
- Organización Panamericana de la Salud. (3 de Enero de 2005). *Contaminación ambiental*.
Obtenido de <http://www.paho.org>
- Pérez Montilla, A. (2018). *Métodos avanzados de análisis de Datos Funcionales*. Puerto Real: Universidad de Cádiz.
- Pérez Plaza, S. M. (2020). *Estudio y aplicación de técnicas de Análisis de Datos Funcionales de geoposicionamiento*. Puerto Real: Universidad de Cádiz.
- Picón Llamas, D. M. (2019). *Análisis de Datos Funcionales aplicado a datos de temperatura en España*. Valladolid: Universidad de Valladolid.
- Ricoy Lorenzo, C. (2006). Contribución sobre los paradigmas de investigación. *Revista do Centro de Educação*, 11-22.
- Romero Placeres, M., Diego Olite, F., & Álvarez Toste, M. (2006). La contaminación del aire: su repercusión como problema de salud. *Revista Cubana de Higiene y Epidemiología*, 44.
- Salini Calderón, G. A. (2014). Estudio Acerca del Material Particulado Emitido en Ciudades de Tamaño Medio al Sur de Santiago de Chile. *Revista INGE CUC*, 97-108.
- Sanchiz, J. (2014). El ozono atmosférico. *Revista Tiempo y Clima*, 15.
- Torres, W. (2002). *Biología de las especies de oxígeno reactivas*. México D.F.: Universidad Nacional Autónoma de México.
- U.S Environmental Protection Agency. (2001). *National air quality 2001 status and trends. Carbon Monoxide*. Washington, D.C: EPA.
- Yassi, A., Kjellstrom, T., de Kok, T., & Guidotti. (2002). *Salud Ambiental Básica*. Maxico DF: PNUMA.

