

UNIVERSIDAD NACIONAL INTERCULTURAL DE LA SELVA CENTRAL JUAN SANTOS ATAHUALPA



FACULTAD DE INGENIERÍA ESCUELA PROFESIONAL DE INGENIERÍA AMBIENTAL

Variabilidad temporal de los gases contaminantes atmosféricos (NO_2 ,
 O_3 y CO) y la relación con los parámetros meteorológicos en Lima
Metropolitana, 2022-2023

TESIS

Para optar el título profesional de Ingeniero Ambiental

AUTORES

Deyvis Bryan Barboza De La Cruz

David Steven Huamani Quiroz

ASESOR

Dr. Alex Rubén Huamán De La Cruz

Chanchamayo, Perú

2024

AGRADECIMIENTOS

Queremos expresar nuestro más sincero agradecimiento a todas las personas que nos acompañaron y apoyaron durante la realización de esta tesis. Este trabajo no habría sido posible sin su guía, soporte, apoyo y compañía en cada etapa del proceso.

En primer lugar, agradecemos profundamente a nuestro asesor, Dr. Alex Rubén Huamán De La Cruz, por su orientación, paciencia y compromiso. Sus consejos y observaciones fueron claves para el desarrollo de esta investigación.

Agradecemos también a nuestros docentes de la facultad de ingeniería, de la carrera profesional de ingeniería ambiental, quienes contribuyeron a nuestra formación académica y nos inspiraron a seguir adelante con entusiasmo y responsabilidad.

A nuestras familias, por su amor incondicional, por creer en nosotros y por brindarnos su apoyo constante en los momentos más exigentes. Gracias por ser nuestro motor y motivo.

A nuestros amigos y compañeros, gracias por las conversaciones, los ánimos, y la buena energía que compartieron con nosotros durante este proceso.

Finalmente, a todas las personas que, de una forma u otra, fueron parte de este camino: gracias por su ayuda y por dejar una huella en esta etapa de nuestras vidas.

DEDICATORIA

A mi mama, Luz Mery de la Cruz Aguilar por la guía e inspiración en mi vida diaria así mismo a mi abuela Isabel Aguilar Herrera porque a pesar que son años difíciles nunca dejo de velar por el bienestar mío.

Barboza de la Cruz, Deyvis Bryan

Dedico esta tesis a mi madre, Bertilda Quiroz Sihuín, por su amor incondicional, sus sacrificios silenciosos y por enseñarme el valor del esfuerzo desde pequeño. A mis amigos más cercanos, por sus palabras de aliento y por hacer más liviano este camino. Y a mí mismo, por haber tenido la fortaleza de seguir adelante, incluso cuando las dudas parecían ganar.

Huamani Quiroz, David Steven

RESUMEN

Estudiar la variabilidad de los contaminantes atmosféricos es vital para conocer como estos se comportan durante el día, la noche, los fines de semana, las estaciones del año, entre otros. Por ello, este trabajo tuvo como objetivo evaluar la variabilidad temporal de los gases contaminantes atmosféricos (NO₂, O₃ y CO) y la relación con los parámetros meteorológicos en Lima Metropolitana, 2022-2023. Datos horarios de los contaminantes gaseosos (2022-2023) fueron proporcionados por el Servicio Nacional de Meteorología e Hidrología (SENAMHI). Estos datos fueron ordenados y alimentados al Software R, en donde se hizo imputación de datos para los datos faltantes y tratamiento de datos estadísticos. Se graficó la variabilidad horaria, diaria, semanal, mensual, estacional, anual, y la relación entre las variables meteorológicas y contaminantes atmosféricos gaseosos. Los resultados revelaron que el CO presenta un similar comportamiento en el horario diurno y nocturno para todos los días de la semana, pero con pequeños incrementos de concentración en la parte del día de 7:00 horas hasta las 17:00 horas, y con el mes de mayo que presentó las mayores concentraciones. El NO₂ mostró un incremento de su concentración en el horario de 07:00 horas a 12:30 horas, y después una reducción de la concentración hasta alrededor las 18:00 horas, febrero presentó las concentraciones promedio más altas de NO₂ (variando de 16 a 38 µg/m³). El O₃ se formó principalmente en el horario con presencia de luz solar entre las 08:00 hr a 19:00 hr, con mayor pico aproximadamente a las 14:00 horas. Además, hubo mayor generación de O₃ en los meses de diciembre a febrero (estación de verano). Se encontraron correlación positiva del O₃ con RS, Temperatura y velocidad del viento, y correlación negativa con la HR, NO₂, y CO. El NO₂ presentó correlación positiva con la RS, y velocidad del viento, y negativa con la HR, CO y temperatura. El CO mostró correlación positiva con la RS y temperatura y negativa con la HR, NO₂, y velocidad del viento. En ambos años muestreados, el viento sopló desde el suroeste al noreste. Las concentraciones de NO₂, CO y O₃ no excedieron los límites del estándar de calidad ambiental (ECA) de la normativa peruana.

Palabras clave: Contaminantes atmosféricos gaseosos, variabilidad temporal, relación.

ABSTRACT

Studying the variability of atmospheric pollutants is vital to know how they behave during the day, night, weekends, seasons of the year, among others. Therefore, this work aimed to evaluate the temporal variability of atmospheric polluting gases (NO_2 , O_3 and CO) and the relationship with meteorological parameters in Metropolitan Lima, 2022-2023. Hourly data on gaseous pollutants (2022-2023) were provided by the National Meteorology and Hydrology Service (SENAMHI). These data were ordered and fed to Software R, where data imputation was made for the missing data and statistical data processing. The hourly, daily, weekly, monthly, seasonal, and annual variability and the relationship between meteorological variables and gaseous air pollutants were graphed. The results revealed that O_3 presents a similar behavior in the day and night hours for all days of the week, but with small increases in concentration in the part of the day from 7:00 a.m. to 5:00 p.m., and with the month of May that presented the highest concentrations. NO_2 showed an increase in its concentration from 07:00 hours to 12:30 hours, and then a reduction in concentration until around 18:00 hours, February presented the highest average concentrations of NO_2 (ranging from 16 to 38 $\mu\text{g}/\text{m}^3$). O_3 formed mainly in the time with the presence of sunlight between 08:00 hr to 19:00 hr, with the highest peak approximately at 14:00 hours. In addition, there was greater generation of O_3 in the months of December to February (summer season). Positive correlation of O_3 with RS, temperature and wind speed, and negative correlation with RH, NO_2 , and CO were found. NO_2 showed a positive correlation with SR and wind speed, and negative correlation with RH, CO and temperature. CO showed a positive correlation with SR and temperature and negative correlation with RH, NO_2 , and wind speed. In both sampled years, the wind blew from the southwest to the northeast. The concentrations of NO_2 , CO and O_3 did not exceed the limits of the environmental quality standard (ECA) of Peruvian regulations.

Keywords: Gaseous air pollutants, temporal variability, relationship.

ÍNDICE

AGRADECIMIENTOS.....	ii
DEDICATORIAS.....	iii
RESUMEN.....	iv
ABSTRACT.....	v
INTRODUCCIÓN.....	xii
CAPÍTULO I.....	13
PLANTEAMIENTO DEL PROBLEMA.....	13
1.1. Fundamentación del problema.....	13
1.2. Formulación del problema.....	14
1.2.1. Problema general.....	14
1.2.2. Problemas específicos.....	14
1.3. Objetivos.....	15
1.3.1. Objetivo general.....	15
1.3.2. Objetivos específicos.....	15
1.4. Definición y operacionalización de variables.....	16
1.5. Hipótesis.....	17
1.5.1. Hipótesis general.....	17
1.5.2. Hipótesis específicas.....	17
1.6. Justificación.....	18
1.6.1. Ambiental.....	18
1.6.2. Social.....	18
1.6.3. Económica.....	18
1.6.4. Teórica.....	19
1.6.5. Práctica.....	19
CAPÍTULO II.....	20
MARCO TEÓRICO.....	20

2.1.	Antecedentes.....	20
2.1.1.	A nivel internacional	20
2.1.2.	A nivel nacional.....	20
2.1.3.	A nivel local.....	21
2.2.	Bases teóricas	22
2.2.1.	Tipos de contaminantes gaseosos	22
2.2.2.	Parámetros meteorológicos.....	23
2.2.3.	Imputación de datos.....	26
2.2.4.	Datos faltantes	26
CAPÍTULO III		27
METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN.....		27
3.1.	Ámbito de estudio.....	27
3.2.	Tipo, Nivel, y diseño de investigación	28
3.2.1.	Tipo de investigación.....	28
3.2.2.	Nivel de investigación	28
3.2.3.	Diseño de investigación.....	29
3.3.	Población y muestra.....	29
3.3.1.	Población	29
3.3.2.	Muestra	29
3.4.	Procedimientos, técnicas e instrumentos de recolección de datos.....	30
3.4.1.	Procedimientos para la recolección de los datos	30
3.4.2.	Técnicas	30
3.4.3.	Instrumento de recolección de datos	31
3.5.	Análisis de datos.....	32
3.6.	Consideraciones éticas.....	33
CAPÍTULO IV		35
RESULTADOS Y DISCUSIONES		35
4.1.	Presentación de resultados y discusiones	35

4.1.1. Concentración anual de los contaminantes gaseosos y parámetros meteorológicos	35
4.1.2. Rosa de vientos de la estación de monitoreo de Carabayllo para el año 2022 y 2023.	38
4.1.3. Comportamiento estacional y temporal promedio de CO, NO ₂ y O ₃ (µg/m ³) para el año 2022-2023.	40
4.1.4. Correlación	43
CONCLUSIONES	47
RECOMENDACIONES	48
REFERENCIAS BIBLIOGRÁFICAS	49
Anexo 1.....	54
Anexo 2.....	56
Anexo 3.....	57
Anexo 4.....	58
Anexo 6.....	63
Anexo 7.....	72

ÍNDICE DE FIGURAS

Figura 1 Piranómetro.....	25
Figura 2 Anemómetro y veleta para medir la velocidad y dirección del viento, respectivamente.	25
Figura 3 Áreas que contienen estaciones de monitoreo de contaminantes gaseosos y parámetros meteorológicos.....	28
Figura 4. Flujograma del proceso para la recolecta de los datos y su respectivo análisis.....	32
Figura 5. Distribución de datos en una serie temporal para los contaminantes atmosféricos gaseosos (CO, NO ₂ , y O ₃ en µg / m ³) y parámetros meteorológicos (T, HR y RS) durante el periodo de 2022-2023. Son mostrados el valor máximo, valor mínimo, media y mediana para cada parámetro.....	38
Figura 6. Rosa de vientos de la estación de monitoreo de Carabayllo durante el 2022 y 2023.	40
Figura 7. <i>Variación de la concentración promedio de CO (µg/m³) en función del tiempo (horaria por día, por hora, por mes y días de la semana) para el año 2022 – 2023 de la estación Carabayllo.</i>	40
Figura 8. Variación de la concentración promedio de NO ₂ (µg.m ³) en diferentes tiempos de medidas (horaria por día, por hora, por mes y días de la semana) para el año 2022 – 2023....	42
Figura 9. Variación de la concentración promedio de O ₃ (µg.m ₃) en diferentes tiempos de medidas (horaria por día, por hora, por mes y días de la semana) para el año 2022 – 2023....	43
Figura 10. Análisis de correlación de Spearman de los contaminantes atmosféricos gaseosos y parámetros meteorológicos para el año 2022-2023.....	46
Figura 11 Evaluación para la imputación de datos del parámetro meteorológico humedad relativa (PPM-imputed distribution).....	58
Figura 12 Evaluación para la imputación de datos del parámetro meteorológico temperatura (CART-imputed distribution).....	58
Figura 13 Evaluación para la imputación de datos del parámetro meteorológico radiación solar (CART-imputed distribution).....	59
Figura 14 Evaluación para la imputación de datos del parámetro contaminante CO (FOREST)	59
Figura 15 Evaluación para la imputación de datos del parámetro contaminante NO ₂ (ppm – imputed distribution)	60
Figura 16 Evaluación para la imputación de datos del parámetro contaminante O ₃ (CART – imputed distribution)	60

Figura 17 Distribución de datos en una serie temporal para los contaminantes atmosféricos gaseosos (CO, NO₂, y O₃ en µg / m³) y parámetros meteorológicos (T, HR y RS) durante el periodo de 2022. Son mostrados el valor máximo, valor mínimo, media y mediana para cada parámetro.....61

Figura 18 Distribución de datos en una serie temporal para los contaminantes atmosféricos gaseosos (CO, NO₂, y O₃ en µg / m³) y parámetros meteorológicos (T, HR y RS) durante el periodo de 2023. Son mostrados el valor máximo, valor mínimo, media y mediana para cada parámetro..... 61

Figura 19 Comportamiento horario, mensual y diario de CO, NO₂, y O₃ para el periodo 2022-2023 de la estación de monitoreo de Carabayllo..... 62

Figura 20 Visita técnica a la Estación de Monitoreo de Carabayllo 62

ÍNDICE DE TABLAS

Tabla 1 Operacionalización de variables	16
Tabla 2 Nivel de investigación	28
Tabla 3. Diseño y técnicas de acopio de información en campo.....	30
Tabla 4. Estándares nacionales de la calidad ambiental del aire y métodos de análisis.....	31
Tabla 5. Estadística descriptiva de las concentraciones anuales de los contaminantes gaseosos y parámetros meteorológicos.....	36

INTRODUCCIÓN

La contaminación del aire por gases contaminantes es un problema general a nivel mundial. Esto porque la mayoría de estos gases liberados entran en contacto con el medio ambiente y las personas de manera directa (NO_x, SO_x, CO, material particulado, entre otros), mientras otros se forman por las reacciones de estos contaminantes con otros compuestos presentes (O₃). Estos gases dependiendo de su concentración y tiempo de exposición (en especial personas mayores, niños y mujeres embarazadas) podrían generar una diversidad de efectos sobre el medio ambiente (por ejemplo; lluvia ácida, eutrofización, daño a la infraestructura, modificar el hábitat de los animales y dañar a las plantas) y la salud humana (por ejemplo; problemas respiratorios y cardiovasculares, reducción de visibilidad, etc.)

Lima es la ciudad más grande, poblada, urbana, y comercial del Perú, y por ende la ciudad más contaminada (Moretti-Villegas et al., 2023). Esta contaminación está relacionada a la presencia de inúmeras industrias y tráfico vehicular que la ciudad viene sufriendo desde hace décadas (Romero et al., 2020). Por ejemplo, Romero et al., (2020) reportó que más altas concentraciones de NO_x, CO, y SO₂ fueron observados en el área metropolitana de Lima en las horas punta de la mañana (8:00 a 10: 00 h) y tarde (18:00 a 19:00 h) de lunes a sábado. Gonzales & Alfaro, (2023) determinaron los niveles de CO del aire de las principales calles de Lima durante cinco días de la semana, siendo que los viernes se midieron las más altas concentraciones de CO (promedio de 38.3 $\mu\text{g}\cdot\text{m}^{-3}$) y que los niveles de CO están entre 29.8 $\mu\text{g}\cdot\text{m}^{-3}$, muy cercano al máximo límite permisible (LMP). Asimismo, Carbo-Bustinza et al., (2022) reportó altas concentraciones de O₃ entre las 00:00 a 14:00 horas en cuatro estaciones de monitoreo del área Metropolitana de Lima, siendo la estación de monitoreo de Ate, quien presentó las más altas concentraciones de O₃.

Así, el objetivo de este trabajo se centró en evaluar la variabilidad temporal de los contaminantes gaseosos atmosféricos y ver la relación que estos presentan frente a las variables meteorológicas de la estación de monitoreo Carabayllo, del Área Metropolitana de Lima, 2022-2023. Este trabajo es importante para conocer como las variables meteorológicas influyen para la reducción o incremento de los contaminantes atmosféricos, como también cual es el comportamiento de estos contaminantes a diferentes temporalidades (horario, diurno, días de la semana, mensual, estacional y

anual).

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1. Fundamentación del problema

La contaminación atmosférica es un problema que aqueja a todo el mundo. Esta contaminación es producto de la emisión de contaminantes (en estado sólido, líquido o gaseoso) provenientes de fuentes naturales (volcanes, aerosol marino, resuspensión del suelo, etc) o antrópicas (industrias, actividades domésticas, etc), los cuales podrían tener efectos negativos sobre la salud humana y el medio ambiente. Los principales contaminantes del aire de criterio se encuentran: el monóxido de carbono (CO), dióxido de nitrógeno (NO₂), ozono (O₃), dióxido de azufre (SO₂), plomo (Pb), y material particulado en sus clasificaciones de PM_{2.5} (diámetro aerodinámico < 2.5 micrómetros), PM₁₀ (diámetro aerodinámico < 10 micrómetros) y TSP (todos los tamaños de partícula) (EPA, 2024). Los contaminantes que son emitidos directamente de la fuente a la atmósfera son conocidos como contaminantes primarios (CO, NO_x, SO_x), mientras aquellos que se forman de la reacción entre los contaminantes primarios y otros compuestos son conocidos como contaminantes secundarios (O₃).

Las emisiones de gases contaminantes, no solo representan una crisis ambiental, sino un obstáculo multidimensional para el desarrollo sostenible. Afecta la salud, la seguridad alimentaria, la disponibilidad de agua y la biodiversidad, e incrementa las desigualdades entre regiones y países. Compromete gravemente el cumplimiento de los **Objetivos de Desarrollo Sostenible (ODS)**, el cumplimiento del **ODS 13** y sus sinergias con otros ODS exige un cambio estructural urgente en los patrones de producción y consumo, además de una cooperación internacional efectiva. (Sachs, 2015)

Los efectos de exposición por parte de las personas a los contaminantes dependen básicamente de la concentración del contaminante, la duración de la exposición, y de la susceptibilidad de las personas expuestas (Manisalidis et al., 2020). El monóxido de carbono (CO) es un gas incoloro, tóxico e inodoro producido por la combustión incompleta de

combustibles fósiles (petróleo, gasolina, madera, etc), y que en contacto con las personas podría ocasionar diversos problemas a la salud (hipoxia, náuseas, fatiga, dificultad para concentrarse, daño cerebral permanente, fallo cardíaco, etc) que está en función a su concentración y tiempo de exposición. Ya en el medio ambiente podría reducir la calidad del aire (menor concentración de oxígeno), alterar la dinámica de los ciclos químicos atmosféricos, y contribuir a que se formen contaminantes secundarios como el ozono troposférico que daña los ecosistemas y salud humana. El dióxido de nitrógeno (NO_2), gas altamente reactivo es generado básicamente por la quema de combustibles fósiles (industrias, vehículos, centrales eléctricas, entre otros). Entre los efectos a la exposición al NO_2 por ejemplo a corto plazo están afecciones como obstrucción bronquial, incremento de síntomas respiratorios y exacerbaciones del asma y a largo plazo posible decremento y obstrucción crónica del desarrollo pulmonar (Saki et al., 2020). La exposición al CO podría acarrear en una reducción de la capacidad de ejercicio, mientras la exposición al O_3 podría generar reducción de la frecuencia cardíaca y respiratoria y alteración y daño del epitelio alveolar (Oyarzún, 2010).

Cichowicz & Stelegowski (2019) reportaron una compleja dinámica entre los contaminantes gaseosos atmosféricos el CO, NO_2 y O_3 , y que toman relevancia sobre todo en áreas urbanizadas. Asimismo, Monks et al. (2015) reveló que el incremento del O_3 troposférico responde al uso de precursores como por ejemplo el CO, metano (CH_4) e hidrocarburos (HCNM) en presencia del NO_x ($\text{NO}_x = \text{NO} + \text{NO}_2$).

1.2. Formulación del problema

1.2.1. Problema general

- ¿Cuál es la variabilidad temporal de los gases contaminantes atmosféricos (NO_2 , O_3 , y CO) y su relación con los parámetros meteorológicos en Lima Metropolitana, 2022-2023?

1.2.2. Problemas específicos

- ¿Qué variabilidad presentarán los gases contaminantes atmosféricos (NO_2 , O_3 y CO) y parámetros meteorológicos (temperatura, radiación solar, dirección y velocidad del viento) en Lima Metropolitana, 2023-2024?
- ¿Será que los gases contaminantes atmosféricos (NO_2 , O_3 y CO) superaron los estándares de calidad ambiental (ECAs)?

- ¿Habrá algún tipo de relación entre las variables meteorológicas y concentración de los gases contaminantes atmosféricos (NO₂, CO y O₃)?

1.3. Objetivos

1.3.1. Objetivo general

- Determinar la variabilidad temporal de los gases contaminantes atmosféricos (NO₂, O₃, y CO) y su relación con los parámetros meteorológicos en Lima Metropolitana

1.3.2. Objetivos específicos

- Estudiar la variabilidad temporal de las variables meteorológicas y concentración de los gases contaminantes atmosféricos (NO₂, O₃ y CO).
- Comparar la concentración de los gases contaminantes atmosféricos con el estándar de calidad ambiental, usando la normativa nacional vigente.
- Determinar la relación entre las variables meteorológicas y concentración de los gases contaminantes atmosféricos (NO₂, O₃ y CO) mediante la correlación respectiva.

1.4. Definición y operacionalización de variables

Tabla 1

Operacionalización de variables

VARIABLE	DEFINICIÓN CONCEPTUAL	DIMENSIONES	INDICADORES	UNIDAD MEDIDA
Variables Independiente: Parámetros meteorológicos	Característica o propiedad del aire a través de la cual se puede determinar cuantitativamente el estado o las condiciones de la atmósfera en un momento y lugar dados (Ventura & Pinto, 2013)	Temperatura	Temperaturas máximas Temperaturas mínimas	°C
		Radiación solar	Radiación solar máxima Radiación solar mínima	W/m ²
		Velocidad del viento	Calma (<1.39 m/s) Flojo (1.7 – 5.6 ms/s) Moderado (5.8 – 11.1 m/s) Fuertes (11.4 – 19.4 m/s) Muy flertes (19.7 – 33,3 m/s) Huracanados (>33.3 m/s)	m/s
		Dirección del viento	Norte (N), Noreste (NE), Este (E) Sureste (SE), Sur (S), Suroeste (SO) Oeste (O), Noroeste (NO)	o
Variable dependiente: Concentración de NO ₂ , O ₃ y CO	La concentración, o abundancia, es la cantidad de un gas específico en el aire. Las emisiones más grandes de gases de efecto invernadero generan concentraciones más altas en la atmósfera (Roberts–Semple et al., 2012)	NO ₂ O ₃ CO	Concentraciones máximas y mínimas Concentraciones máximas y mínimas Concentraciones máximas y mínimas	µg/m ³ µg/m ³ ppm

1.5. Hipótesis

1.5.1. *Hipótesis general*

H₀ = Existe relación de los gases contaminantes atmosféricos (NO₂, O₃ y CO) en el análisis temporal con los parámetros meteorológicos en Lima Metropolitana

1.5.2. *Hipótesis específicas*

- Se presentó la variabilidad temporal de las variables meteorológicas y concentración de los gases contaminantes atmosféricos (NO₂, O₃ y CO).
- La concentración de los gases contaminantes atmosféricos superó los estándares de calidad ambiental, de la normativa nacional vigente.
- Hay relación entre las variables meteorológicas y concentración de los gases contaminantes atmosféricos (NO₂, O₃ y CO) mediante la correlación respectiva.

1.6. Justificación

1.6.1. Ambiental

Nuestra investigación es fundamental para entender mejor como la contaminación atmosférica tiene profundos y múltiples impactos sobre los ecosistemas, la biodiversidad, el clima, recursos naturales y medio ambiente. Así, es vital obtener información sobre la concentración de los contaminantes gaseosos y ver como estos afectan al medio ambiente y la salud humana para que se puedan implementar medidas correctivas o de mitigación. Asimismo, como el estudio presentará el comportamiento temporal de los contaminantes, las personas podrán saber en qué momento del tiempo, estos contaminantes presentan mayores concentraciones, para que estos puedan evitar su exposición.

El Objetivo de Desarrollo Sostenible 13 (ODS 13) establece la necesidad urgente de “adoptar medidas urgentes para combatir el cambio climático y sus efectos”, lo cual implica no solo el reconocimiento del calentamiento global como una amenaza global, sino también la necesidad de una acción inmediata y coordinada. Este objetivo demanda la reducción significativa de las emisiones de gases de efecto invernadero (GEI), como el dióxido de carbono (CO₂), el metano (CH₄) y el óxido nitroso (N₂O), los cuales son los principales responsables del incremento de la temperatura media global.

1.6.2. Social

Es vital estudiar y ver formas de mermar o reducir la contaminación atmosférica, así como los efectos directos e indirectos que estos ocasionan sobre la calidad del aire y por ende sobre la salud, el bienestar y la equidad en las comunidades. Estos impactos generan profundos cambios en la estructura social y en la calidad de vida de las personas, especialmente en las poblaciones más vulnerables. Así, el saber en qué instante del día, en que área, estos contaminantes se presentan en más altas concentraciones o cuáles son sus fuentes de emisión es importante para que la sociedad este prevenida.

1.6.3. Económica

Este trabajo aborda los contaminantes atmosféricos, principalmente los gases, quienes tienen impacto negativo sobre la salud humana, así, su evaluación será útil para el medio ambiente y la economía en general. Esto porque en relación a la salud humana genera aumento en los gastos de salud por enfermedades respiratorias, cardiovasculares y otros problemas relacionados. Se estima que millones de personas mueren cada año a nivel mundial debido a la exposición prolongada a contaminantes como las partículas en suspensión (PM_{2.5}), el ozono

troposférico y otros gases tóxicos como el NO_x y CO. Además, es importante ver como los contaminantes se comportan en función a las variables meteorológicas, para poder predecir si tendrá efecto por ejemplo sobre el cultivo y así, evitar gastos innecesarios.

1.6.4. Teórica

La presente investigación proporcionará información de la concentración de contaminantes gaseosos, así como, de parámetros meteorológicos dentro del área de Lima Metropolitana. Asimismo, será evaluado el comportamiento estos contaminantes en relación de los parámetros meteorológicos dentro de áreas urbanas. Esta información servirá para poder tomar decisiones que permitan cuidar la calidad de aire de las diversas áreas urbanas y ver si estos mejoran o empeoran con relación a los diversos parámetros meteorológicos.

1.6.5. Práctica

Los resultados encontrados y reportados de este trabajo proporcionan datos de tres diferentes contaminantes gaseosos atmosférico, de su variabilidad temporal y de la relación con los parámetros meteorológicos que ayudará a que se implementen políticas ambientales y regulaciones locales, para proteger ecosistemas y la salud humana.

CAPÍTULO II

MARCO TEÓRICO

2.1. Antecedentes

2.1.1. A nivel internacional

Oujidi et al., (2022) en su artículo científico titulado “*Gaseous Pollutants and Particulate Matter in Ambient Air: First Field Experiment in an-Urban Mediterranean Area (Nador, Morocco)*”, tuvieron el objetivo de medir los contaminantes gaseosos (NO₂, CO y O₃) y material particulado de la ciudad de Nador, Marruecos. Para esto tres campañas de muestreo fueron realizadas entre 2016 y 2018. Los gases monitoreados estuvieron por debajo de los límites estándar nacionales. Concluyeron que se necesita establecer un plan de gestión de la calidad del aire implementando una red de estaciones de monitoreo para que se identifique y controlen actividades que causan la contaminación del aire.

Li et al., (2020) en su artículo científico “*Spatial distribution characteristics of gaseous pollutants and particulate matter inside a city in the heating season of Northeast China*” se propusieron analizar las características de los contaminantes gaseosos (SO₂ y NO₂) dentro las áreas urbanas de la ciudad de Shengan, Noreste de China. Para ello, estos contaminantes fueron medidos en dos principales carreras de la ciudad que se instalaron sobre vehículos durante tres épocas determinadas. Los resultados demostraron que el promedio de SO₂, fueron similares en las tres épocas, mientras que el NO₂ fue reportado en altas concentraciones en los tres eventos. Concluyeron que este estudio aporta para un mejor entendimiento de la distribución padrón de los contaminantes y que factores influyen para así poder tomar decisiones adecuadas en base a base científica.

Chlebnikovas et al. (2023) en su artículo científico “*Short - term field research on air pollution within the boundaries of the large city in the Baltic region*” tuvieron como objetivo determinar la contaminación de NO₂, O₃, CO y SO₂, dentro de las áreas de la región Báltica (contiene áreas residenciales, de recreación y áreas cercanas a calles de intenso flujo vehicular) en diferentes estaciones del año. La primera fase inició el 2016-2017, mientras el segundo de 2019-2020. Fue encontrado que en las áreas cercanas al intenso transporte la concentración del NO superó los 48 µg/m³.

2.1.2. A nivel nacional

Córdova-Mendoza et al., (2021) en su artículo de título “Primera caracterización de emisiones contaminantes y la calidad del aire en Ica, Perú” tuvieron como objetivo caracterizar las emisiones contaminantes (SO_2 y NO_2) y evaluar la calidad del aire de Ica, Perú, a través de dos estaciones de monitoreo instalados (E1 y E2). Los resultados reportaron menores concentraciones comparado a la normativa peruana en las dos estaciones para los contaminantes gaseosos.

Rebaza et al. (2021) en su artículo científico de título “Análisis y evaluación de los niveles de CO y CO_2 en la Ciudad de Tacna en relación con el parque automotor y los estándares de calidad ambiental del aire”, tuvieron como objetivo determinar en qué horario, lugar de monitoreo y cómo influye la concentración de CO y CO_2 en la ciudad de Tacna. Asimismo, se compararon los niveles de CO con los estándares de calidad Ambiental del aire (ECA) peruano y se determinó el grado de asociación entre el número de vehículos y la concentración de los contaminantes. Concluyeron que hay un efecto significativo del horario, lugar de monitoreo. La concentración de CO no superó los ECA, pero se asoció de manera positiva y baja al número de vehículos, mientras el CO_2 no se asoció al parque automotor.

2.1.3. A nivel local

Benites-morales y Pacsi-Valdivia (2023) en su trabajo titulado “Análisis temporal de los contaminantes atmosféricos (NO_2 , O_3 troposférico y CO) y su relación con la temperatura del aire y la radiación solar en Lima Metropolitana (LM) tuvieron como objetivo evaluar de manera temporal los niveles de NO_2 , CO, y O_3 troposférico en Lima Metropolitana entre los años 2015 al 2018 y su relación con la temperatura y radiación solar. Se realizaron análisis temporales descriptivos, correlacionales y de regresión usando los datos horarios registrados en diferentes áreas de LM. Los gases encontrados estuvieron debajo de la normativa peruana y la OMS, indicando que no es perjudicial sobre la salud humana. El comportamiento de los gases es bimodal el cual fue influenciado por el tráfico vehicular y desplazamiento nocturno. El comportamiento diario evidenció mayores concentraciones en fin de semana. Fue observado una clara correlación negativa del CO con el O_3 , y correlación positiva entre el NO_2 y CO.

Medina (2023) en su artículo científico “Análisis temporal y espacial del ozono en tres distritos de Lima y su relación con el parque automotor” tuvo como objetivo en estudiar la distribución y cambio de concentración del O_3 y sus precursores (NO_x , NO, y NO_2), así como, de los parámetros meteorológicos (velocidad del viento, temperatura, y humedad) en tres distritos (Puente Piedra, Santa Anita y Jesús María) de Lima Metropolitana en el periodo del 2017. Mayores concentraciones de O_3 ($7.01 \mu\text{g}/\text{m}^3$) fue reportado en Puente Piedra los cuales estuvieron relacionados a las emisiones de NO del parque automotor, y en menor proporción

por el clima cálido, vientos del sur a norte y la poca dispersión del contaminante. Los resultados indican un efecto negativo de las condiciones meteorológicas.

2.2. Bases teóricas

2.2.1. Tipos de contaminantes gaseosos

Entre los tipos de contaminantes gaseosos más comunes se encuentran los hidrocarburos, monóxido de carbono (CO), óxidos de nitrógeno (NO_x), óxidos de azufre (SO_x), y ozono (O₃). Estos usualmente se encuentran a diferentes concentraciones y son principalmente emitidos por la quema de combustibles fósiles, industrias y volcanes.

- a) Monóxido de carbono:** Un producto de la combustión incompleta y tiende a ser peligroso si las personas o animales se exponen, debido a que este contaminante se fija en la hemoglobina, lo que impide se transporte en oxígeno en el organismo. Es inodoro, y cuando se siente un ligero dolor de cabeza ya es demasiado tarde. Tiende a diluirse en áreas abiertas como el medio ambiente, pero en ambientes cerrados, se torna tóxico e inclusive podría ser mortal. Fuentes: Aparatos de calefacción y cocinas, Automóviles, Chimeneas y Hornos, Generadores portátiles, Barbacoas y parrillas, Hogueras y calentadores a base de leña y Cigarrillos y fumar en interiores. Entre las consecuencias el CO es peligroso porque se combina con la hemoglobina en la sangre, impidiendo que el cuerpo transporte oxígeno adecuadamente, lo que puede causar intoxicación severa o incluso la muerte. Por eso es vital contar con detectores de CO y asegurar una ventilación adecuada en los lugares donde se utilizan fuentes de combustión (Andreae et al., 2012).
- b) Dióxido de nitrógeno:** es un gas tóxico de color marrón-rojizo, con un olor picante característico, que se forma como resultado de la combustión de combustibles fósiles. Es uno de los principales contaminantes del aire y un componente clave de la formación del smog. Entre sus principales fuentes de emisión se encuentran: los vehículos a motor, centrales eléctricas, calefacción en hogares y edificios, procesos industriales, Aviación y navegación, Incendios forestales y quema de biomasa. Entre las consecuencias que trae su exposición son los problemas respiratorios, infecciones respiratorias, riesgo cardiovascular. Ya en el medio ambiente, hay formación de lluvia ácida, Contribución al smog y ozono troposférico, y Eutrofización de cuerpos de agua (D'Amato et al., 2014).
- c) Dióxido de azufre:** la principal fuente de emisión de SO₂ está relacionado a la combustión del carbón el cual contiene azufre. El SO₂ que resulta de la combustión tiende a oxidarse formando aso el ácido sulfúrico (H₂SO₄) que es un componente de la lluvia acida, nocivo para las plantas, el cual hace que aparezca manchas amarillas en las hojas si tienen contacto con algunas gotitas. Asimismo, este compuesto tiende a atacar a los materiales de

construcción que usualmente lo conforman los compuestos carbonatados, como por ejemplo la piedra caliza o mármol, haciendo que se formen sustancias solubles en el agua, el cual podría afectar la integridad y vida de esculturas y edificios. Entre las consecuencias para la Salud Humana están los problemas respiratorios, agravamiento de enfermedades preexistentes, inflamación de los pulmones, inflamación del sistema respiratorio, disminución de la función pulmonar y al aumento de la sensibilidad a las infecciones. Entre los problemas para el medio ambiente están la formación de lluvia ácida, acidificación de cuerpos de agua, daños a los cultivos, efectos sobre el Clima, formación de aerosoles de sulfato (Peláez et al., 2020).

d) Ozono: es un constituyente natural presente en la atmósfera, no en tanto, si su concentración supera la concentración normal es considerado como un contaminante. Su concentración usualmente se incrementa debido a los gases que se generan en el tubo de escape de los automóviles es alta, así como la radiación solar, la concentración del ozono también se incrementa. La presencia de ozono en bajas concentraciones podría dañar el desarrollo normal de las plantas, mientras el hombre es afectado cuando las concentraciones de ozono varía de 0.05 a 0.1 mg kg⁻¹, el cual le causa irritación sobre las fosas nasales y garganta, como también genera sequedad de las mucosas de las vías respiratorias superiores. El ozono troposférico no se emite directamente, sino que se forma a través de reacciones químicas entre contaminantes primarios en presencia de luz solar. Entre las consecuencias para la Salud Humana, se encuentran problemas respiratorios, agravamiento de enfermedades respiratorias, reducción de la función pulmonar, problemas cardiovasculares. Entre las consecuencias para el medio ambiente están: daño a la vegetación, reducción de la biodiversidad, e impacto en cultivos agrícolas (Brönnimann, 2022).

2.2.2. Parámetros meteorológicos

Entre los parámetros meteorológicos se encuentran la temperatura, humedad relativa, la presión atmosférica, radiación solar, y la dirección y velocidad del viento. Estos parámetros son monitoreados y medidos con el objetivo de evaluar la calidad del aire, debido a que estos podrían influenciar en su incremento o reducción de los contaminantes (Senamhi, 2023). Asimismo, estos parámetros medidos ayudan a que se diseñen proyectos de energías renovables. Por ejemplo, la evaluación de la velocidad y dirección del viento a diversas alturas ayuda a determinar cuan viable es realizar un proyecto de parques eólicos, así como, seleccionar de manera eficiente la pala, rotores y altura de instalación. La radiación solar se mide para proyectos de energía solar, mediante el cual se permite saber cuánta cantidad de energía solar se recibe por metro cuadrado.

- a) **Temperatura:** magnitud física escalar que se relaciona a la energía interna dentro un sistema termodinámico, específicamente relacionado a la energía cinética media de sus partículas. Parámetro físico que permite se pueda medir la sensación de frío o calor. Usualmente este parámetro es cuantificado utilizando un termómetro y expresado en °C (Meteo-technology, 2021).
- b) **Humedad relativa:** medida de la cantidad de vapor de agua presente en el aire. Relación que existe entre la presión parcial del vapor de agua y la presión de vapor de equilibrio del agua a una determinada temperatura. Este parámetro es medido usando un higrómetro y es expresado en términos de porcentaje (%) (Meteo-technology, 2021).
- c) **Presión atmosférica:** o barométrica se refiere a la fuerza por unidad de medida que es ejercida por el aire de la atmosfera sobre la superficie terrestre. Esta presión no es igual en todas partes, dado que depende de la altura, y es más alta cuando está más cerca del nivel del mar. El instrumento usado para medir este parámetro es el barómetro y reportada en mmHg (Meteo-technology, 2021).
- d) **Radiación solar:** se refiere a la energía que emite el Sol, el cual tiende a propagarse en todas las direcciones a través del espacio por intermedio de las ondas electromagnéticas. Esta energía es la encargada de determinar la dinámica del clima y de los procesos atmosféricos. Para cuantificar la radiación solar son utilizados los piranómetros (solarímetros) y los pirheliómetros (Figura 1) quienes reportan el valor en irradiancia cuyas unidades están en kWh m² o W.m² (Meteo-technology, 2021).

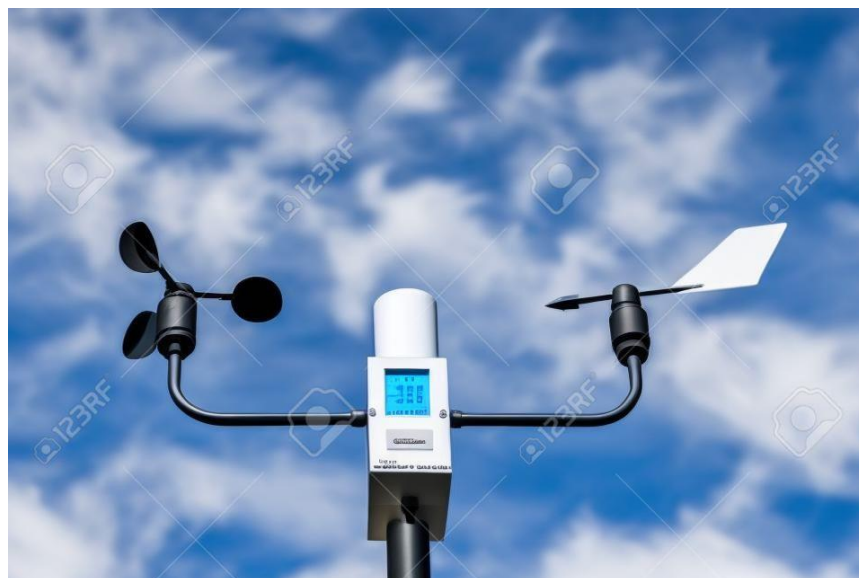
Figura 1
Piranómetro



Fuente: (Senamhi, 2023)

e) **Dirección y Velocidad del viento:** representa el equilibrio logrado por tres fuerzas que actúan sobre este: la fricción superficial, gradiente de presión y la fuerza de Coriolis. Dirección del viento (medido usando una veleta) es referida al punto del horizonte del observador desde el cual sopla (donde proviene). Actualmente, a nivel internacional se usa la rosa de vientos, el cual se divide en 360° . La velocidad del viento si se representa a través de un vector, la longitud indica la velocidad del viento, el cual es medido utilizando un anemómetro y sus valores son expresados en m/s (Figura 2).

Figura 2
Anemómetro y veleta para medir la velocidad y dirección del viento, respectivamente.



Fuente: (Senamhi, 2023)

2.2.3. Imputación de datos

Se refiere a la sustitución de valores que no se informaron en una observación por otros. Es algo necesario que se tiene que hacer para que se traten los datos empleando determinadas técnicas estadísticas. A través de este método se evita que se pierda información, se reduce el sesgo, se logra mejorar la precisión del análisis estadístico.

2.2.4. Datos faltantes

Son aquellos datos que no se guardaron o por algún problema no se logró registrar e refiere a la sustitución de valores que no se informaron en una observación por otros. Es algo necesario que se tiene que hacer para que se traten los datos empleando determinadas técnicas estadísticas como por ejemplo los diferentes métodos de imputación de datos.

CAPÍTULO III

METODOLOGÍA Y TÉCNICAS DE INVESTIGACIÓN

3.1. Ámbito de estudio

El ámbito de estudio está representado por el Área de Lima Metropolitana (ALM-12°00'49"S 76°51'18"O), el cual es un área urbana constituido por 43 distritos que componen la provincia de Lima y la provincia constitucional del Callao (Figura 3). Está localizada en la costa central del Perú, flanqueada por el desierto costero y orillas del océano Pacífico. Este ALM es considerada como una megaciudad, siendo la más extensa y poblada del Perú, mientras a nivel de Latinoamérica es una de las cinco mayores ciudades, y a nivel internacional se encuentra en el puesto 32 (Carranza et al., 2002). Su extensión superficial es de 2819 km² cuya población estimada es de alrededor de 11.3 millones de habitantes aproximadamente (INEI, 2023). Posee un clima subtropical desértico por la influencia de la corriente de Humboldt, siendo que en verano se tiene días calurosos, y bochornosos, mientras que en invierno se tiene aire fresco, seco, ventosos y usualmente despejados. El ALM se caracteriza por su baja presencia de precipitaciones (menos 15 mm por año), mientras que la temperatura varía de 30 grados centígrados a 24 grados centígrados.

El AML alberga la mayor cantidad de vehículos del parque automotor, mucho de estos de antigüedad mayor a quince años, así como también contiene una serie de industrias de diversos sectores, los cuales probablemente sean los que emiten diferentes contaminantes gaseosos hacia la atmósfera.

Carabayllo (11°51'00"S 77°02'00"O), es el distrito más extenso (346.89 km²) y se localiza en ambos márgenes del río Chillón. Cuenta con una población aproximada de 426 985 habitantes (INEI, 2025). Su clima es árido y semicálido, mientras su temperatura promedio anual es de $18 \pm 3^{\circ}\text{C}$, con mayores temperaturas reportadas en febrero (20 a 28 °C). Este distrito alberga industrias de carpintería, metal mecánico, prendas de vestir y de producción de plásticos. La Autoridad de Transporte (ATU) indicó que el distrito de Carabayllo y otros son los que presentan un aire más contaminado (calidad de aire pobre que varía de 38.2 a 38.6 $\mu\text{g}/\text{m}^3$).

Figura 3

Áreas que contienen estaciones de monitoreo de contaminantes gaseosos y parámetros meteorológicos.



3.2. Tipo, Nivel, y diseño de investigación

3.2.1. Tipo de investigación

Es tipo de investigación es longitudinal debido a que los datos se tomarán en diversas temporadas (Landeau, 2007).

3.2.2. Nivel de investigación

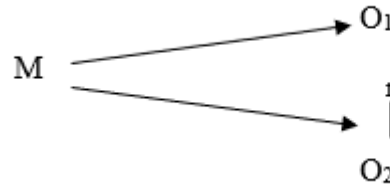
Basado a la profundidad con que es abordado el tema, el nivel de la investigación corresponde al *tipo descriptivo correlacional*. (Arias-Gómez et al., 2016).

Tabla 2

Nivel de investigación

criterio	Nivel de investigación
El grado de asociación o relación entre dos o más variables (no causal)	<i>Descriptivo correlacional</i>

Esto porque las variables no se manipulan de manera intencional para que se determina la relación de la concentración de los gases contaminantes atmosféricos con los parámetros meteorológicos.



Donde:

M = muestra

O₁ = Variable 1 (Concentración de los gases contaminantes atmosféricos)

O₂ = Variable 2 (variables meteorológicas)

r = relación de las variables de estudio

3.2.3. *Diseño de investigación*

El diseño de la investigación es no experimental, longitudinal de tendencia, debido a que se cuantificara la variación de los contaminantes del aire y sus parámetros meteorológicos de un determinado periodo.

3.3. Población y muestra

3.3.1. *Población*

La población está compuesta por el total de estaciones de monitoreo de los gases de contaminantes atmosféricos y meteorológicos presente en los distritos de Lima Metropolitana, del año 2022 y 2023 y que le pertenece a la institución estatal del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI).

3.3.2. *Muestra*

El tipo de muestreo aplicado será el no probabilístico, por conveniencia basado al criterio de los investigadores. En este caso se elegirá la estación de Carabayllo. Esta estación fue elegida porque tiene los datos casi completos de los contaminantes gaseosos (NO₂, O₃, y

CO) correspondientes al año 2022 y 2023, seleccionados en comparación a las otras estaciones de SENAMHI.

3.4. Procedimientos, técnicas e instrumentos de recolección de datos

3.4.1. *Procedimientos para la recolección de los datos*

Los datos de contaminantes gaseosos atmosféricos, así como los parámetros meteorológicos horarios de las estaciones automáticas de monitoreo de la calidad del aire serán solicitadas al Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). La posible información proporcionada por la autoridad competente estará relacionada a:

- Concentración horaria de NO₂
- Concentración horaria de O₃
- Temperatura del aire horaria
- Radiación solar horaria
- Velocidad y dirección del viento

3.4.2. *Técnicas*

Para el trabajo de campo se procederá a seguir con el protocolo Nacional de Monitoreo de Calidad del aire que se aprobó a través el D.S. No 010-2019-NIMAM concerniente a las estaciones automáticas de monitoreo. Este decreto presenta el Protocolo Nacional de Monitoreo de la Calidad Ambiental del Aire que establece las directrices para el monitoreo y control de la calidad del aire en todo el país, con el fin de asegurar que las actividades económicas y los proyectos de inversión cumplan con los estándares ambientales.

Tabla 3.

Diseño y técnicas de acopio de información en campo

DISEÑO	TÉCNICA	INSTRUMENTO
Diseño de colecta de información de la estación automática	Análisis de los datos obtenidos Procesamiento de los datos	Ficha de colecta de datos para gases y variables meteorológicas (*)

Cuadro de registro por cada contaminante gaseoso atmosférico.

*Ficha de colecta de datos (Anexo 1)

Tabla 4.

Estándares nacionales de la calidad ambiental del aire y métodos de análisis

Parámetros	Método de análisis
Benceno	Cromatografía de gases
SO ₂	Fluorescencia UV
NO ₂	Quimioluminiscencia
PM _{2.5}	Separación Inercial/filtración
PM ₁₀	Separación Inercial/filtración
Mercurio gaseoso total (Hg)	Espectrometría de absorción Zeeman
CO	Infrarrojo no dispersivo
O ₃	Fotometría de absorción ultravioleta
Pb en PM ₁₀	Espectrometría de absorción atómica

Fuente: (Senamhi, 2023)

3.4.3. Instrumento de recolección de datos

Instrumentos: los instrumentos para cuantificar los datos meteorológicos y de concentración horaria de los gases contaminantes atmosféricos (NO₂, O₃ y CO) empleados fueron:

- Estación automática de calidad del aire del SENAMHI
- Estación meteorológica automática del SENAMHI

Software: para realizar el análisis descriptivo, temporal y correlacional de las variables meteorológicas y de concentración de los gases contaminantes atmosféricos se emplearán las siguientes herramientas:

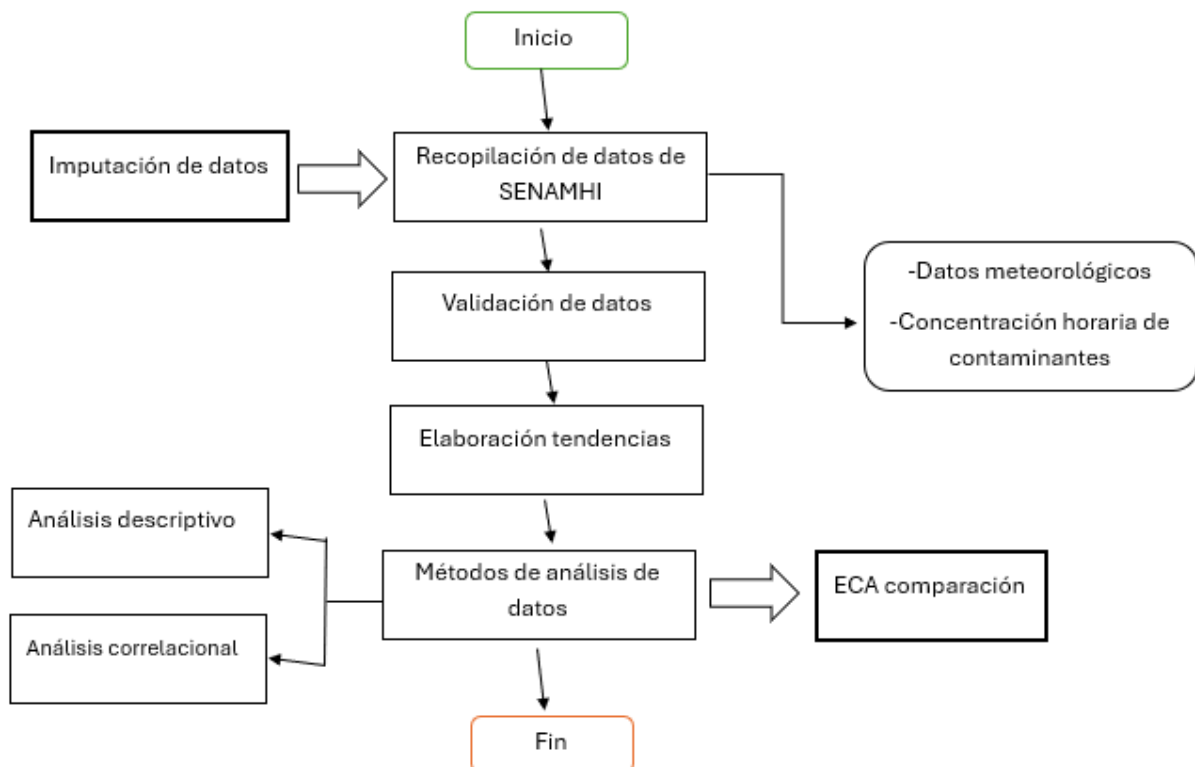
- Microsoft Excel 2019, el cual servirá para ordenar los datos. Aquí los datos se ordenarán en función a su monitoreo horario, y donde los parámetros meteorológicos serán adicionados en sus respectivas líneas y columnas para tener una base de datos del contaminante horario y de su parámetro meteorológico correspondiente
 - El Software libre R, en donde los datos ordenados serán alimentados, y a través del uso de paquetes se generarán tablas de resúmenes de las estadísticas descriptivas y de tendencia central. Asimismo, estos datos serán utilizados para generar gráficos de

comportamiento horario, diario, fines de semana, días de semana, mensual, por estaciones y anual. Además, los promedios de contaminantes gaseosos anuales serán comparados con los Estándares de Calidad del aire (ECA) dados en la normativa peruana. La Figura 4 presenta de manera gráfica el proceso para recolectar y analizar los datos.

- Datos faltantes fueron IMPUTADOS utilizando funciones y paquetes que se instalaron en el software libre R. En este caso, se usó la metodología de la distribución normal, donde se comparó la distribución normal de los datos que se tenía versus la distribución de los datos obtenidos mediante las otras metodologías después de que estos hayan imputado los datos faltantes.

Figura 4.

Flujograma del proceso para la recolecta de los datos y su respectivo análisis



3.5. Análisis de datos

Los datos proporcionados por el SENAMHI fueron analizados previamente. Al analizar estos datos se percató que había demasiados datos faltantes en una primera exploración de los datos, por lo que se aplicó la técnica de imputación de datos. La imputación de datos es un método estadístico que permite se reemplace datos faltantes (NA) por valores que son

estimados, para poder analizarlos con técnicas estadísticas (Castro & Ávila, 2006). Este método es útil si se tiene poco porcentaje de datos faltantes, y no es recomendable si los datos faltantes son muy altos, ya que los resultados podrían no ser tan confiables (Araneda, 2021).

En este estudio se hizo una previa exploración de los datos faltantes, y se aplicó metodologías para reducir el sesgo o se generen resultados no fiables. Por ejemplo, en algunos casos los datos se tomaron por mes y los datos de ese mes fueron completados utilizando tres métodos de imputación: Simple Value Imputation, el paquete Multivariate Imputation by Chained Equations (MICE), y la técnica Miss Forest Imputation que está basada en el algoritmo de Random Forest (Buuren & Groothuis-Oudshoorn, 2011).

La distribución de las variables fue utilizada para evaluar las variables antes y después de la imputación de los datos.

La correlación de los parámetros atmosféricos y gases contaminantes fue evaluada mediante la correlación de Spearman el cual mide la fuerza y dirección que tienen dos variables usando rangos que varían de -1 a +1. El valor positivo indica una correlación directa (si una variable se incrementa la otra también) y el valor negativo una correlación inversa (si una variable se incrementa la otra disminuye). Valores de 0.0 – 0.2 indican una correlación mínima, de 0.2 a 0.4 una correlación baja, mientras de 0.4 a 0.6 correlación moderada, de 0.6-0.8 indica una correlación alta, mientras valores de 0.8 a 1.0 una correlación muy alta (Schober et al., 2018).

3.6. Consideraciones éticas

Plantear los fundamentos básicos para el respeto a la dignidad humana, la investigación científica y los derechos de autor y la ética profesional.

- No haremos falsificación o modificación en los resultados que obtengamos del trabajo de investigación.
- Daremos el consentimiento para brindar la información a la Universidad Nacional Intercultural de la Selva Central Juan Santos Atahualpa, para la divulgación de la investigación, previa consulta de nuestra persona.
- Evitar cualquier daño que pueda afectar a la universidad, ello con fines de mantener la integridad de la institución y de nuestra persona, esto es cuando hay documentos por medio.
- Nos regimos en el cumplimiento normativo de nuestra universidad y cualquier otra entidad que apoye en nuestro proyecto de investigación.

- Mantendremos siempre la integridad personal, profesional y científica, siendo honestas en todas las etapas del proceso de la investigación, desde la recopilación de datos hasta la presentación de los resultados.

CAPÍTULO IV

RESULTADOS Y DISCUSIONES

4.1. Presentación de resultados y discusiones

4.1.1. Concentración anual de los contaminantes gaseosos y parámetros meteorológicos

La concentración anual (2022, 2023 y 2022-2023) de los contaminantes gaseosos (CO, NO₂, y O₃) y valores registrados de los parámetros meteorológicos: Radiación solar (RS), temperatura (Temp) y humedad relativa (HR) de la estación de monitoreo de Carabayllo son presentados en la Tabla 5.

Para el CO, se observa que la máxima concentración registrada fue 4782 µg/m³ (2022), la media más alta de 2086 µg/m³ (2023), el mínimo valor de 70 µg/m³ (2023), mediana de 2437 µg/m³ (2023) y tercer cuartil en 2907 µg/m³ (2023) (Tabla 5). Basado en los resultados descritos arriba, podría concluirse que casi los mayores valores de CO fueron registrados en el año 2023, sin embargo, no lograron exceder las concentraciones de estándar de calidad ambiental (ECA) para el CO en 1 hora (30000 µg/m³) y 8 horas (10000 µg/m³). Asimismo, se observa diferencia significativa ($p > 0.05$) entre el año 2022 y 2023 para el CO. El CO es un gas generado principalmente por la combustión incompleta de los combustibles fósiles. En Lima y Carabayllo su presencia está relacionado probablemente al tráfico vehicular intenso (vehículos antiguos, uso extensivo de diésel y gasolina, mototaxis que no cumplen estándares modernos), la combustión en actividades industriales y domésticas, condiciones geográficas (Lima esta entre el mar y los Andes, por ende, tiene poca ventilación atmosférica) y meteorológicas (en mañanas frías, es formado una capa de inversión térmica que mantiene a los contaminantes cerca del suelo) y la estacionalidad (invierno presenta menor radiación y mayor inversión térmica que impide la dispersión de contaminantes).

EL NO₂ presenta valores que varían de 1.70 µg/m³ a 72.6 µg/m³ entre los periodos de 2022-2023, y una media anual de 28.8 µg/m³ para el 2022, 18.4 µg/m³ en el 2023 y 23.1 µg/m³ para ambos años. La media anual presenta diferencias significativas ($p > 0.05$) entre los años y ambos años. Los valores reportados en este estudio no superaron las concentraciones horarias ni anuales del ECA (1 hora = 200 µg/m³ y anual = 200 µg/m³). El NO₂ es conocido como un gas bastante irritante y contaminante que se genera por la combustión de combustibles fósiles, principalmente en altas temperaturas. Su presencia en la atmosfera de Carabayllo está

relacionado al tráfico vehicular (especialmente a los vehículos antiguos, mal regulado y sin mantenimiento), a las industrias (fábricas, hornos, calderas) y fuentes estacionarias, las condiciones geográficas (baja ventilación atmosférica) y meteorológicas (inversión térmica).

En Carabayllo el O₃ presentó una media anual de 5.07 µg/m³ en el periodo 2022-2023, seguido por el año 2023 con 4.0 µg/m³ y, por último, el valor de 3.24 µg/m³ el año 2022. Estos valores encontrados no superaron los límites de Estándar nacional de Calidad Ambiental para aire (ECA, 100 µg/m³). Silva et al., (2018) en su trabajo de análisis y formación de ozono en Lima metropolitana también reportó que las concentraciones de O₃ no superó el límite ECA para el ozono, siendo que a nivel superficial existe el efecto de la capa límite marina y a la fuerte inhibición de formación basado a la valoración con el óxido nítrico. Estas medias anuales presentaron diferencias significativas (p>0.05). El menor valor de O₃ encontrado fue 0.62 µg/m³ y el máximo valor de 11.6 µg/m³. Su presencia y formación del O₃ está relacionada a los procesos fotoquímicos de contaminantes (NO_x y COVs) emitidos por actividades antrópicas (tráfico vehicular intenso). En este caso, los NO_x y COVs liberados a la atmósfera entran en contacto con la luz solar que descompone al NO₂ en NO y oxígeno atómico, y este último se combina al O₂ para formar el O₃. Lima a pesar de tener casi todo el año nublado, recibe suficiente radiación UV en determinadas épocas que hace que reaccionen contaminantes que no se lavan por falta de lluvia (poca precipitación) o ciertas condiciones meteorológicas (inversión térmica).

Tabla 5.

Estadística descriptiva de las concentraciones anuales de los contaminantes gaseosos y parámetros meteorológicos.

Parámetro	Unidad	Media	Max	Min	Mediana	3er cuartil
2022						
CO	µg/m ³	2015 a	4782 a	203 a	2027 a	2320 a
NO ₂	µg/m ³	28.8 a	72.6 a	1.90 a	28.7 a	37.5 a
O ₃	µg/m ³	3.24 a	11.6 a	0.62 a	2.80 a	4.30 a
RS	Wh/m ²	99.5 a	409 a	0.20 a	77.8 a	134 a
Temper	°C	19.2 a	31 a	11.6 a	18.9 a	22.2 a
HR	%	78.5 a	99.7 a	40.3 a	80.7 a	88.2 a
2023						
CO	µg/m ³	2086 b	4387 b	70 b	2437 b	2907 b
NO ₂	µg/m ³	18.4 b	65.5 b	1.70 b	17.1 b	26.0 b
O ₃	µg/m ³	4.0 b	11.2 a	0.62 a	3.4 b	6.1 b
RS	Wh/m ²	225 b	1121 b	0.0 b	117 a	117 b
Temper	°C	22 b	31 a	11.6 a	18.9 a	18.9 b
HR	%	76.5 b	99.7 a	40.3 a	80.7 a	80.7 b
2022 – 2023						

CO	$\mu\text{g}/\text{m}^3$	2050 c	4782 a	70.1 b	2141 c	2640 c
NO ₂	$\mu\text{g}/\text{m}^3$	23.1 c	72.6 a	1.70 b	23.1 c	32.6 c
O ₃	$\mu\text{g}/\text{m}^3$	5.07 c	11.3 a	0.62 a	3.2 b	5.9 b
RS	Wh/m ²	101 c	359 c	0.2 a	80 a	140 a
Temper	°C	20.6 a	32 a	11.6 a	20.5 b	23.4 c
HR	%	77.5 b	99.7 a	38.1 b	79.3 a	86.7 a

Nota: misma letra en la misma columna y contaminante o parámetro meteorológico no son diferentes significativamente ($p > 0.05$).

4.1.1.1 Normativa usada para comparar las concentraciones

- **Decreto Supremo N.º 003-2017-MINAM** (7 de junio de 2017)

Establece los ECA para aire como referente obligatorio para las actividades productivas, extractivas y de servicios

Deroga normas previas (DS 074-2001-PCM, DS 069-2003-PCM, DS 003-2008-MINAM, DS 006-2013-MINAM)

Contaminantes: CO, NO₂, O₃

Contaminante	Periodo de promedio	Valor ECA ($\mu\text{g}/\text{m}^3$)	Criterio de evaluación	Método de análisis recomendado
Monóxido de carbono (CO)	1 hora	30,000	No debe excederse más de 1 vez por año	Infrarrojo no dispersivo (NDIR)
	8 horas (media móvil)	10,000	Media aritmética móvil	Infrarrojo no dispersivo (NDIR)
Dióxido de nitrógeno (NO₂)	1 hora	200	No debe excederse más de 24 veces por año	Quimioluminiscencia
	Anual	100	Media aritmética anual	Quimioluminiscencia
Ozono (O₃)	8 horas (media móvil)	100	No debe excederse más de 24 veces por año	Fotometría UV

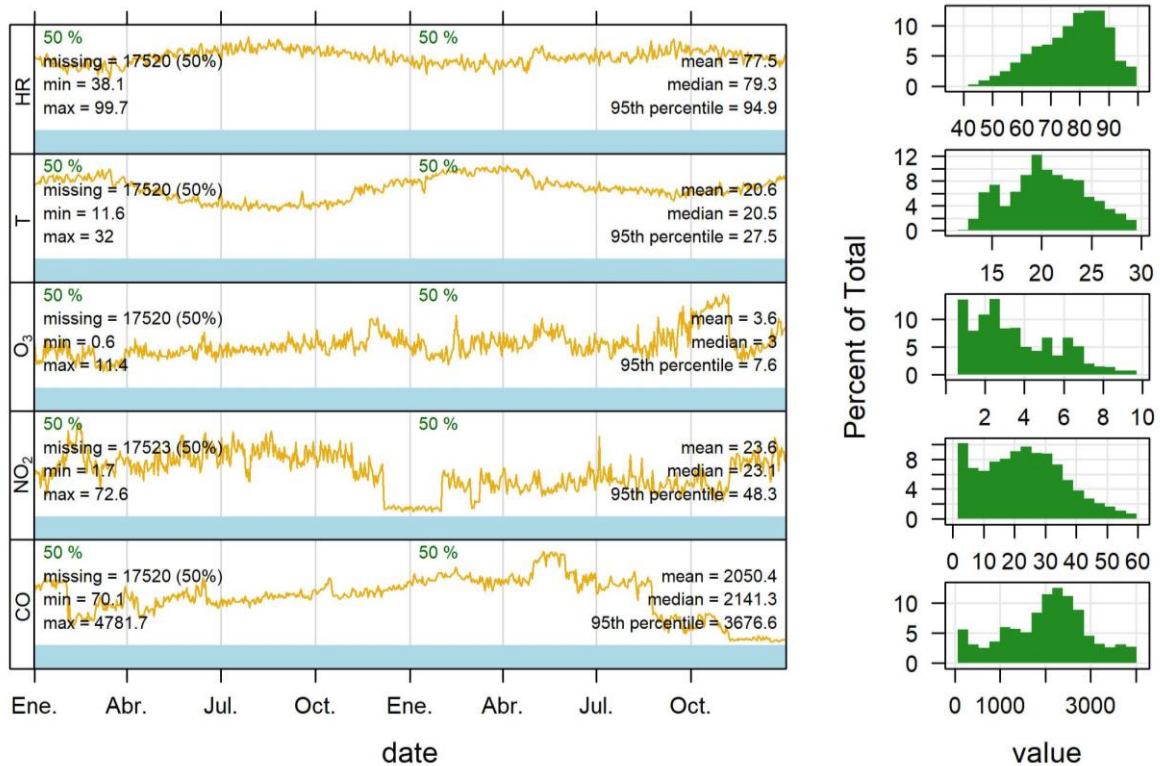
- **Directrices OMS 2021 – Valores recomendados**

Las Directrices Mundiales de Calidad del Aire de la OMS (2021) son un conjunto de recomendaciones internacionales que establecen niveles máximos de concentración de contaminantes en el aire para proteger la salud humana.

Contaminante	Periodo de promedio	Valor guía (AQG)	Unidad	Comentarios
Ozono (O₃)	8 horas (99.º percentil)	100	µg/m ³	Promedio móvil diario; no exceder en más del 3% de los días
	Temporada pico (6 meses)	60	µg/m ³	Promedio de las 6 semanas más contaminadas
Dióxido de nitrógeno (NO₂)	Anual	10	µg/m ³	Media aritmética anual
	24 horas (99.º percentil)	25	µg/m ³	Valor diario a no superar más del 3% de los días
Monóxido de carbono (CO)	24 horas (99.º percentil)	4	mg/m ³	Basado en protección de grupos sensibles

Figura 5.

Distribución de datos en una serie temporal para los contaminantes atmosféricos gaseosos (CO, NO₂, y O₃ en µg / m³) y parámetros meteorológicos (T, HR y RS) durante el periodo de 2022-2023. Son mostrados el valor máximo, valor mínimo, media y mediana para cada parámetro.



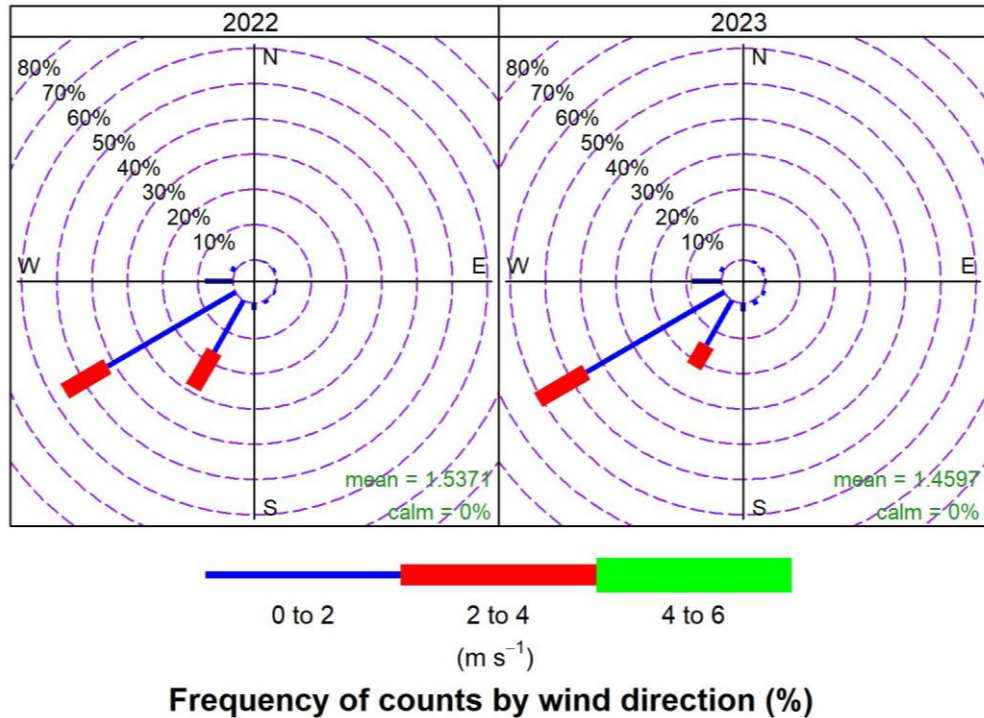
4.1.2. Rosa de vientos de la estación de monitoreo de Carabayllo para el año 2022 y 2023.

La Figura 6 muestra la rosa de vientos para el año 2022 y año 2023 de la estación de monitoreo Carabayllo. De la Figura 6 se puede observar que, durante el periodo de muestreo

del año 2022, el viento sopló desde el suroeste el 48% y 28% del tiempo, mientras que, para el periodo de muestreo del año 2023, el viento sopló también del suroeste el 60% y 21% del tiempo. Los colores de cada radio presentes en la rosa de vientos nos informan sobre la velocidad del viento en nudos ($1 \text{ nudo} = 0.514 \text{ m s}^{-1}$ – metros por segundo), del viento en cada dirección. En este caso, en el 2022 el radio más largo muestra que el viento sopló desde el suroeste a velocidades entre 0 a 2 nudos (azul claro) aproximadamente un 42% y 18% del tiempo, respectivamente, mientras de 2 a 4 nudos (rojo) aproximadamente el 14% y 11% del tiempo, respectivamente. En el año 2023, se puede notar que el radio más largo indica que el viento sopló también desde el suroeste a velocidades entre 0 a 2 nudos (color azul) en alrededor del 45% y el 15% y en velocidades de 2 a 4 nudos (color rojo) en aproximadamente el 15% y 5%, respectivamente. La velocidad promedio del viento para el 2022 fue de 1.537 m.s^{-1} , mientras para el 2023 presentó una velocidad media de 1.4597 m.s^{-1} . Además, se observa un % de 0.0% de calm, lo que significa que no fue detectado viento o que nunca hubo viento totalmente en reposo. Como observado en la Figura 6, la rosa de vientos muestra un comportamiento similar para ambos años indicando que los vientos soplaron del sur oeste para el noreste. Similar comportamiento del aire soplando del suroeste al noreste fueron encontradas en las estaciones de monitoreo de Campo de Marte, Carabayllo, y Naña (Quispe et al., 2020). Encalada-Malca et al., (2021) reportó que los vientos en Carabayllo soplaban del suroeste al noreste.

Figura 6.

Rosa de vientos de la estación de monitoreo de Carabayllo durante el 2022 y 2023.



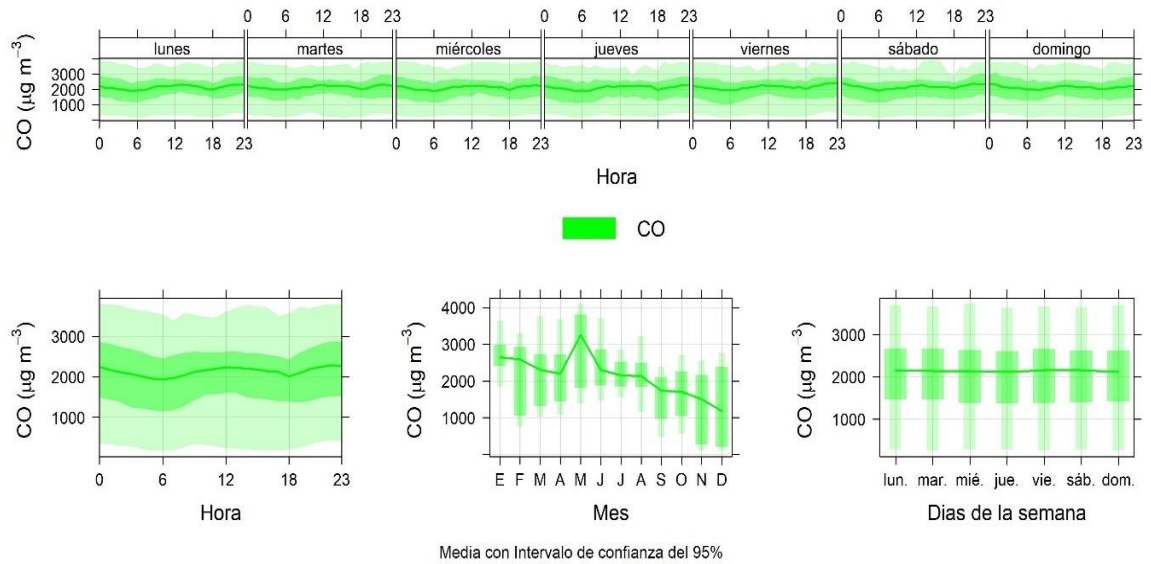
4.1.3. Comportamiento estacional y temporal promedio de CO , NO_2 y O_3 ($\mu\text{g}/\text{m}^3$) para el año 2022-2023.

CO

El comportamiento estacional y temporal del CO horaria, diaria, mensual y anual es mostrado en la Figura 7. De esta Figura 7 es observado un similar comportamiento en el horario diurno y nocturno para todos los días de la semana, pero con pequeños incrementos de concentración en la parte del día de 7:00 horas hasta las 17:00 horas, y después en la parte de la noche de 20:00 a las 04:00 horas. En función a los meses, mayo fue el mes que presento mayores concentraciones de CO que vario aproximadamente de 2000 a 3900 $\mu\text{g}/\text{m}^3$. Entre los días de semana fue observado un comportamiento similar.

Figura 7.

Variación de la concentración promedio de CO ($\mu\text{g}/\text{m}^3$) en función del tiempo (horaria por día, por hora, por mes y días de la semana) para el año 2022 – 2023 de la estación Carabayllo.

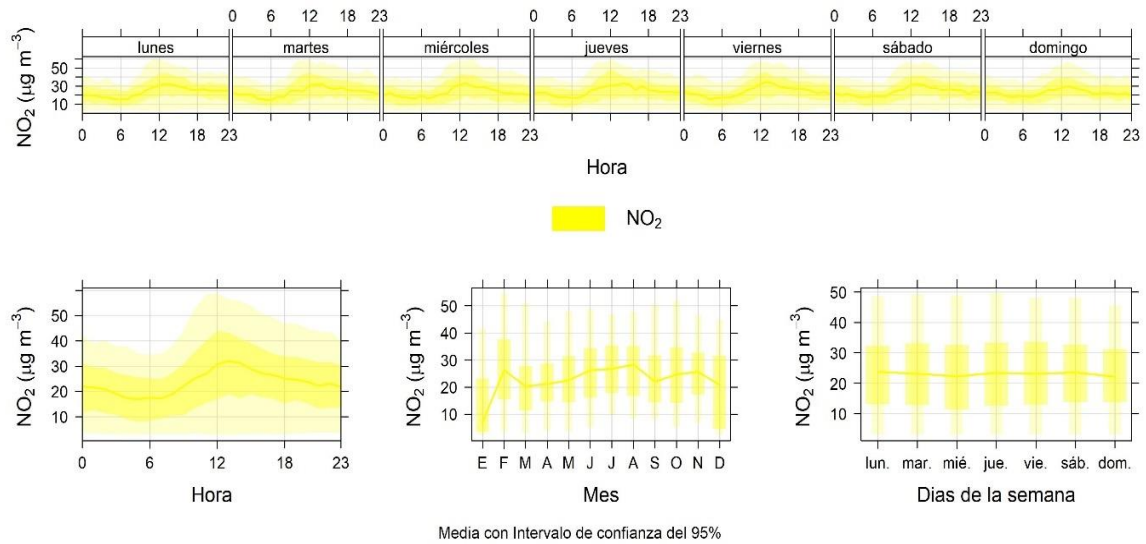


NO₂

El comportamiento horario del NO₂ muestra un incremento de su concentración en el horario que inicia aproximadamente a las 07:00 horas hasta las 12:30 horas, y después se observa una reducción de la concentración hasta alrededor las 18:00 horas. Este mismo comportamiento se puede observar para todos los días de la semana con concentraciones que varían entre los 10 $\mu\text{g}/\text{m}^3$ a 40 $\mu\text{g}/\text{m}^3$ (Figura 8). En función a los meses, febrero presentó las concentraciones promedio más altas de NO₂ (variando de 16 a 38 $\mu\text{g}/\text{m}^3$), seguido por los meses de junio, julio, agosto y setiembre (variando alrededor de 18 a 34 $\mu\text{g}/\text{m}^3$). Basado en los días de la semana la concentración de NO₂ es similar para todos los días con pequeñas variaciones. Esto podría explicarse porque el NO₂ usualmente es liberado por la combustión de combustibles fósiles relacionado al tráfico vehicular intenso que se tiene en Carabayllo y Lima Metropolitana en general que usualmente inician labores desde las 06:00 horas y se mantiene hasta aproximadamente 18:00 horas.

Figura 8.

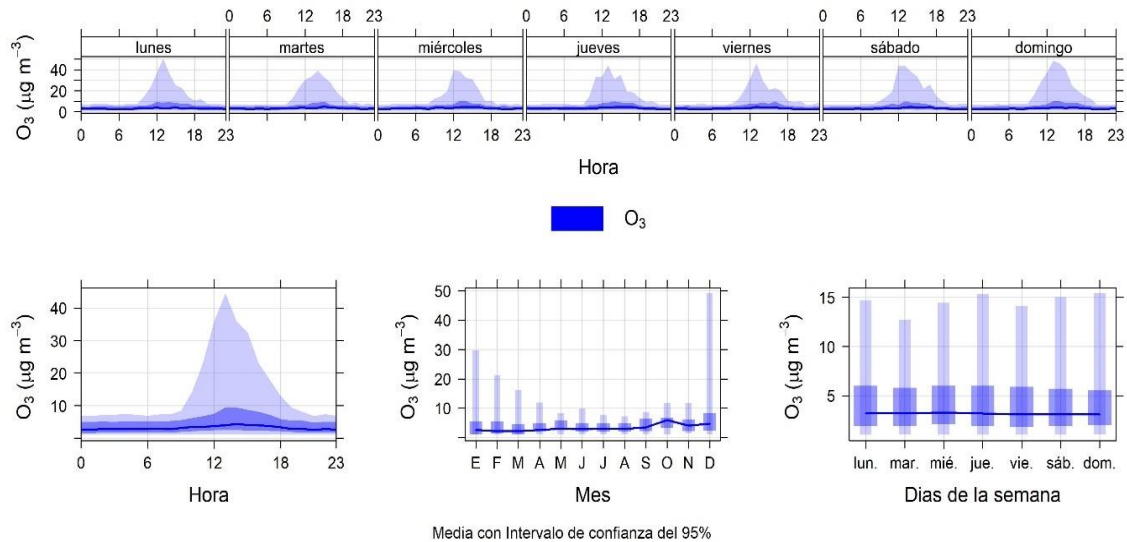
Variación de la concentración promedio de NO_2 ($\mu\text{g}\cdot\text{m}^3$) en diferentes tiempos de medidas (horaria por día, por hora, por mes y días de la semana) para el año 2022 – 2023.

**O₃**

En la Figura 9 se observa el comportamiento horario en función de los días, comportamiento en los meses y días de la generación del O_3 ($\mu\text{g}/\text{m}^3$) en la estación de monitoreo de Carabayllo. En el comportamiento horario es notado que el O_3 se forma principalmente en el horario donde hay presencia de luz solar, iniciando alrededor de las 08:00 hr hasta las 19:00 hr) y con mayor pico aproximadamente a las 14:00 horas en casi todos los días de la semana. Asimismo, es observado que hay mayor generación de ozono en los meses de diciembre a febrero (y los subsecuentes meses va disminuyendo), meses que justamente el área de Lima Metropolitana está en la estación de verano y es donde la presencia de luz solar es mayor. Para los días de la semana se observa una ligera mayor producción de O_3 los fines de semana (sábado y domingo), esto probablemente porque como la mayoría de las personas van a ciudades de viaje o de paseo, empiezan a retornar a sus hogares. Carbo-Bustinza et al., (2022) reveló que la estación de monitoreo de Ate Vitarte comparado el de San Borja, y Campo de Marte era el que mayor prevalencia de ozono presento. Además, reportó que la formación de O_3 estaba entre el horario de 07:00 horas hasta 19:00 horas aproximadamente y que el mayor pico se presentaba en alrededor de la 14:00 horas, algo similar encontrado en este estudio.

Figura 9.

Variación de la concentración promedio de O_3 ($\mu\text{g}\cdot\text{m}_3$) en diferentes tiempos de medidas (horaria por día, por hora, por mes y días de la semana) para el año 2022 – 2023.

**4.1.4. Correlación**

En la **Figura 10**, se presenta la correlación de Spearman de los parámetros meteorológicos vs los contaminantes atmosféricos del periodo 2022-2023 para la estación de monitoreo de Carabayllo. De la **Figura 10** es observado que existe una correlación negativa alta ($r = -0.79$) entre la temperatura y la humedad relativa (HR). Este comportamiento puede explicarse porque cuando se incrementa la temperatura, el porcentaje de saturación tiende a disminuir, lo que hace que la humedad relativa baje. En contraste, cuando la temperatura disminuye, el aire pierde la capacidad de contener vapor, haciendo que la humedad relativa suba (Tejeda et al., 2018). Similares resultados fueron reportados en diversas investigaciones donde reportan correlación negativa entre estas variables (Majumder & Ray, 2021; Zakaria et al., 2020; Zender-Świercz et al., 2024).

Similarmente, fue reportado una correlación negativa moderada ($r = -0.59$) entre la HR y velocidad del viento (wind spd). Este comportamiento no es universal como en el caso de la temperatura y la humedad relativa, porque este depende del contexto climático, condiciones locales y origen del viento. En este caso, esta correlación negativa moderada indica que el viento favorece la evaporación y/o introduce aire más seco, haciendo que la humedad relativa descienda. Por ejemplo, Zakaria et al., (2020) reportó una correlación negativa ($r = -0.26$) entre la humedad y velocidad del viento, mientras una correlación positiva de $r = 0.28$ entre la temperatura y velocidad del viento.

Una correlación negativa baja ($r=-0.26$) es observado entre la HR y radiación solar (RS) (Figura 10). A mayor presencia de radiación solar, la temperatura del aire tiende a subir, por lo que la humedad relativa disminuye, pero la cantidad absoluta de la humedad podría mantenerse constante, lo que explicaría esta correlación negativa baja. Zhang et al., (2021) and Islam et al., (2023) reportaron correlación negativa entre la HR y RS, el primero utilizando modelos bayesianos y el segundo en ciudades de Bangladesh, respectivamente.

Correlación positiva baja ($r=0.27$) fue observado entre la RS y la temperatura. Cuando la radiación solar llega al suelo, esta es absorbida por el suelo, vegetación cuerpos de agua, construcciones etc., que después lo reemiten como calor al medio ambiente haciendo que el aire se caliente. Esta correlación positiva baja podría estar relacionada a que la RS en Carabayllo es poco, por lo tanto, la temperatura no tiende a subir mucho. En Bagdad (Iraq) (Hashim et al., 2023) y Nepal (Kathmandu Valley) (Shrestha et al., 2020) fueron encontrados una correlación positiva entre las variables meteorológicas temperatura y RS.

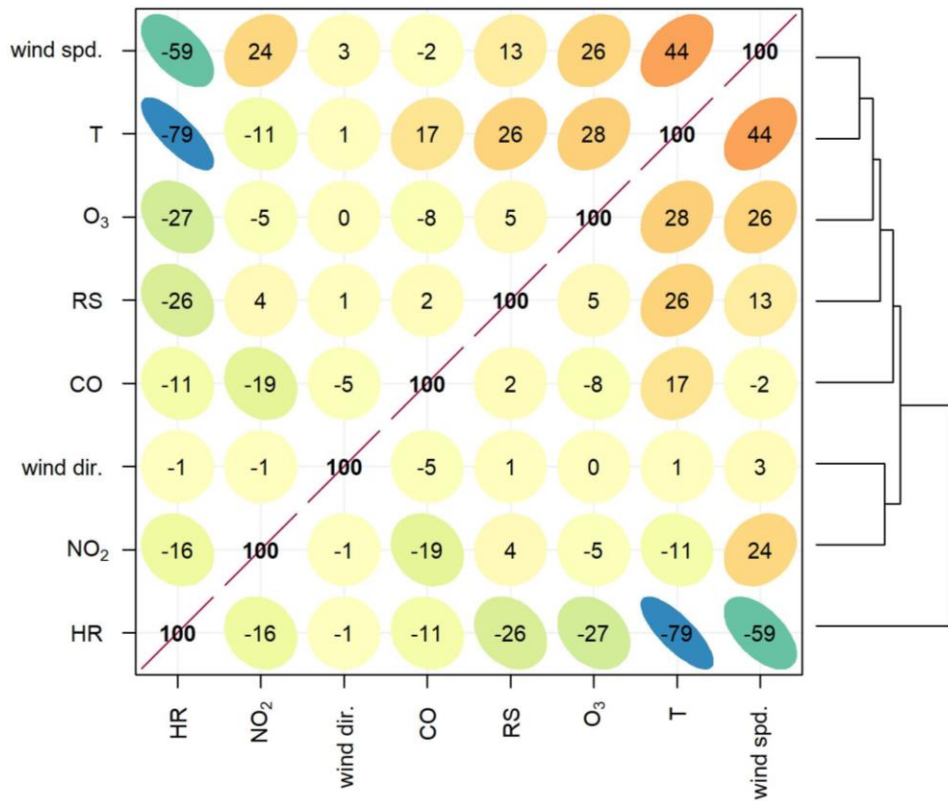
El ozono (O_3) presenta correlación positiva baja con la temperatura ($r=0.28$) y correlación positiva mínima con la radiación solar ($r=0.05$) y velocidad del viento. El ozono se forma cuando los óxidos de nitrógeno (NO_x) y los compuestos orgánicos volátiles (COVs) reaccionan con la radiación solar. Temperaturas más altas aceleran reacciones químicas (fotoquímicas) y favorecen la evaporación de los COVs, por ende, permiten la formación del ozono. Asimismo, mayor RS permite que se activen las reacciones que favorecen la formación del O_3 . En cuanto a la velocidad del viento este podría en situaciones (contexto local) tener una correlación positiva o negativa, para este caso el viento ayuda a que el ozono se distribuya mostrando una correlación positiva. Como Lima en general recibe poca luz solar y radiación solar, esto podría explicar la baja y mínima correlación positiva encontrada en el área de estudio. Un estudio realizado en tres ciudades urbanas (Krakow, Paris y Milán) de Europa reveló una correlación positiva de O_3 con la temperatura y velocidad del viento (Rowland, 2024). En contraste, muestra una correlación negativa baja con la humedad relativa ($r=-0.27$), una correlación negativa mínima con el NO_2 ($r=-0.05$) y con el CO ($r=-0.08$). Una alta presencia de Humedad reduce la presencia de O_3 por diversas razones tales como menor radiación solar disponible (presencia de mayor nubosidad), mayor remoción por radicales OH y deposición húmeda. En el caso del NO_2 , podría explicarse porque este es precursor (O_3 se forma cuando NO_2 se fotodisocia) y a veces destructor (si hay exceso de NO) del O_3 . Paraschiv et al., (2020) reportó una correlación negativa del NO_2 , NO y humedad relativa en relación al ozono troposférico evaluada en una zona industrial de Portugal (Porto). Similarmente, Rowland, (2024) encontró una correlación negativa entre el O_3 y el NO_2 entre ciudades de Europa

(Krakow, Paris y Milan). Para el CO, esta correlación negativa podría estar relacionada a que el CO consume radicales OH necesarios para la formación de O₃, y si el CO se encuentra en altas concentraciones podría ralentizar la formación del ozono. Por ejemplo, Kanchana et al., (2020) después de evaluar la calidad del aire en la ciudad urbana de Telangana (India) encontró una correlación negativa entre en CO y el O₃. El CO en Lima en general se encuentra a altas concentraciones lo que podría explicar esta correlación negativa mínima.

El NO₂ presenta una correlación negativa mínima con la HR ($r=-0.16$), CO ($r=-0.19$) y temperatura ($r=-0.11$), mientras correlación positiva mínima con la RS ($r=0.04$) y velocidad del viento ($r=0.24$). Alta humedad relativa viene acompañada de neblina o lluvia, favoreciendo la remoción del NO₂ (deposición húmeda), además alta humedad es sinónimo de formación de radicales OH, lo que podrían reaccionar con el NO₂ transformándolos. El incremento de temperatura incrementa la intensidad de la RS, y por ende acelera la fotodisociación del NO₂. En NO₂ y CO en ciertos contextos podrían presentar una correlación positiva debido a que usualmente provienen de las mismas fuentes. Sin embargo, en algunos contextos podría darse una correlación negativa en situaciones cuando: a) cuando el CO por ser estable se mantiene más tiempo en la atmosfera que el NO₂ y b) cuando en CO compite con el NO₂ por los radicales OH, produciendo la reducción de NO₂, y mayor presencia de CO favorece la destrucción del NO₂. Por ejemplo Meena et al., (2022) encontró una correlación negativa entre el CO y NO₂ en diversas regiones de la India. Teóricamente la RS y el NO₂ deberían presentar correlación negativa, sin embargo, en este estudio se encontró una correlación positiva lo que indica que durante el día hay mayor presencia de tráfico vehicular y actividad humana, lo que implica mayor emisión de NO₂, haciendo que su destrucción sea superada. En pocas palabras, mayor RS significa mayor actividad antrópica y más reacciones y por ende más NO₂. En el caso de la velocidad del viento podría estar relacionada que este parámetro en vez de dispersar contaminantes transporta contaminantes desde zonas cercanas.

Figura 10.

Análisis de correlación de Spearman de los contaminantes atmosféricos gaseosos y parámetros meteorológicos para el año 2022-2023.



CONCLUSIONES

Se determinó la variabilidad temporal de los contaminantes atmosféricos en función a los datos horarios, diarios y mensuales de la estación de monitoreo de Carabayllo del periodo 2022 – 2023. Aquí se observó que los contaminantes incrementan su concentración principalmente durante el día lo que su presencia posiblemente esté relacionada al intenso tráfico vehicular y actividades industriales desarrolladas alrededor de la estación de monitoreo.

Los contaminantes gaseosos fueron comparados a los valores de estándares de calidad ambiental (ECA) para cada gas, y se encontró que ninguno de los tres contaminantes gaseosos superó o excedieron los valores **ECA** tampoco comparando con las **Directrices OMS 2021**.

Fueron encontrados correlaciones positivas y negativas de los contaminantes en función de los parámetros meteorológicos.

RECOMENDACIONES

- Evaluar el mismo contaminante gaseoso para evaluar su comportamiento estacional y variación temporal en diferentes estaciones de monitoreo.
- Complementar los estudios con los años actuales para verificar si los contaminantes se incrementaron o disminuyeron.
- Evaluar los resultados un poco más profundos, talvez haciendo evaluaciones más exhaustivas por hora, día, estaciones del año o temporada seca y lluviosa.
- Tratar los datos en relación a la imputación de los datos faltantes con otros softwares y metodologías para elegir el más adecuado.
- No abarcar muchos contaminantes que el estudio de hace extenso, pero esto fue realizado por sugerencia de los jurados revisores que al parecer desconocen de la magnitud de datos que se trabaja.

REFERENCIAS BIBLIOGRÁFICAS

- Andreae, M. O., Artaxo, P., Beck, V., Bela, M., Freitas, S., Gerbig, C., Longo, K., Munger, J. W., Wiedemann, K. T., & Wofsy, S. C. (2012). Carbon monoxide and related trace gases and aerosols over the Amazon Basin during the wet and dry seasons. *Atmospheric Chemistry and Physics*, 12(13), 6041–6065. <https://doi.org/10.5194/acp-12-6041-2012>
- Araneda, P. (2021). *Imputación de datos*. Rpubs.
- Arias-Gómez, J., Villasís-Keever, M., & Miranda-Novales, M. (2016). El protocolo de investigación III: la población de estudio. *Rev Alerg Mex*, 63(02), 1–201. <https://doi.org/10.29262/ram.v63i2.181>
- Benites-morales, O. F., & Pacsi-Valdivia, S. A. (2023). Análisis temporal de los contaminantes atmosféricos (NO₂, O₃ troposférico y CO) y su relación con la temperatura del aire y la radiación solar en Lima Metropolitana. *Ecología Aplicada*, 22(1), 1–9. <https://doi.org/10.21704/rea.v22i1.850>
- Brönnimann, S. (2022). Century-long column ozone records show that chemical and dynamical influences counteract each other. *Communications Earth and Environment*, 3(1), 1–11. <https://doi.org/10.1038/s43247-022-00472-z>
- Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice : Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- Carbo-Bustinza, N., Belmonte, M., Jimenez, V., Montalban, P., Rivera, M., Martínez, F. G., Mohamed, M. M. H., De La Cruz, A. R. H., da Costa, K., & López-Gonzales, J. L. (2022). A machine learning approach to analyse ozone concentration in metropolitan area of Lima, Peru. *Scientific Reports*, 12(1), 1–12. <https://doi.org/10.1038/s41598-022-26575-3>
- Carranza, C. C., Maldonado, M. D., Arévalo, W. G., Pacheco, R. A., Giraldo, A. V., & Loayza, S. (2002). Relaciones entre calidad ambiental y calidad de vida en Lima Metropolitana. *Revista Del Instituto De Investigación De La Facultad De Minas, Metalurgia Y Ciencias Geográficas*, 5(9), 47–52.
- Castro, L. M. U., & Ávila, D. M. M. (2006). Una introducción a la imputación de valores perdidos. *Terra*, XXII(31), 127–152.
- Chlebnikovas, A., Paliulis, D., Bradulienė, J., & Januševičius, T. (2023). Short - term field research on air pollution within the boundaries of the large city in the Baltic region. *Environmental Science and Pollution Research*, 30(34), 81950–81965.

<https://doi.org/10.1007/s11356-022-23798-9>

- Cichowicz, R., & Stelęgowski, A. (2019). Average Hourly Concentrations of Air Contaminants in Selected Urban, Town, and Rural Sites. *Archives of Environmental Contamination and Toxicology*, 77(2), 197–213. <https://doi.org/10.1007/s00244-019-00627-8>
- Córdova-Mendoza, P., Barrios-Mendoza, T., & Córdova-Barrios, I. (2021). Primera caracterización de emisiones contaminantes y la calidad del aire en Ica, Perú. *Revista Cubana de Química*, 33(1), 138–152.
- D'Amato, G., Bergmann, K. C., Cecchi, L., Annesi-Maesano, I., Sanduzzi, A., Liccardi, G., Vitale, C., Stanziola, A., & D'Amato, M. (2014). Climate change and air pollution. *Allergo Journal International*, 23(1), 17–23. <https://doi.org/10.1007/s40629-014-0003-7>
- Encalada-Malca, A. A., Cochachi-Bustamante, J. D., Rodrigues, P. C., Salas, R., & López-Gonzales, J. L. (2021). A Spatio-Temporal Visualization Approach of PM10 Concentration Data in Metropolitan Lima. *Atmosphere*, 12(5), 609. <https://doi.org/10.3390/atmos12050609>
- EPA. (2024). *Criteria Air Pollutants*. Environmental Protection Agency.
- Gonzales, A. V., & Alfaro, V. P. D. (2023). *Study of carbon monoxide levels in the most traveled streets of Metropolitan Lima*. Universidad de Lima.
- Hashim, B. M., Al-Naseri, S. K., Hamadi, A. M., Mahmood, T. A., Halder, B., Shahid, S., & Yaseen, Z. M. (2023). Seasonal correlation of meteorological parameters and PM2.5 with the COVID-19 confirmed cases and deaths in Baghdad, Iraq. *International Journal of Disaster Risk Reduction*, 94(January), 103799. <https://doi.org/10.1016/j.ijdr.2023.103799>
- INEI. (2023). *Población peruana alcanzó los 33 millones 726 mil personas en el año 2023*. Instituto Nacional de Estadística e Informática.
- INEI. (2025). *Estadística Poblacional*. Instituto Nacional de Estadística e Informática.
- Islam, N., Toha, T. R., Islam, M. M., & Ahmed, T. (2023). Spatio-temporal Variation of Meteorological Influence on PM2.5 and PM10 over Major Urban Cities of Bangladesh. *Aerosol and Air Quality Research*, 23(1), 220082. <https://doi.org/10.4209/aaqr.220082>
- Kanchana, A. L., Sagar, V. K., Pathakoti, M., Mahalakshmi, D. V., Mallikarjun, K., & Gharai, B. (2020). Ozone variability: Influence by its precursors and meteorological parameters- an investigation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 211(October), 105468. <https://doi.org/10.1016/j.jastp.2020.105468>
- Landeau, R. (2007). *Elaboración de trabajos de investigación*. Editorial Alfa.
- Li, C., Liu, M., Hu, Y., Zhou, R., Huang, N., Wu, W., & Liu, C. (2020). Spatial distribution

- characteristics of gaseous pollutants and particulate matter inside a city in the heating season of Northeast China. *Sustainable Cities and Society*, 61(June), 102302.
<https://doi.org/10.1016/j.scs.2020.102302>
- Majumder, P., & Ray, P. P. (2021). A systematic review and meta-analysis on correlation of weather with COVID-19. *Scientific Reports*, 11(1), 1–10.
<https://doi.org/10.1038/s41598-021-90300-9>
- Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and Health Impacts of Air Pollution: A Review. *Frontiers in Public Health*, 8(February), 1–13. <https://doi.org/10.3389/fpubh.2020.00014>
- Medina, M. E. H. (2023). Análisis temporal y espacial del ozono en tres distritos de Lima y su relación con el parque automotor. *Ciencia Latina*, 7(1), 10168–10185.
https://doi.org/10.37811/cl_rcm.v7i1.5199
- Meena, D., Dashora, H., & Soni, M. (2022). Long-term study of CO , NO2 and meteorological parameters over North-West India. *Strad Research*, 9(2), 1–19.
<https://doi.org/10.37896/sr9.10/008>
- Meteo-technology. (2021). *Meteorological variables*. Meteo-Technology.Com.
- Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidmesser, E., Sommariva, R., Wild, O., & Williams, M. L. (2015). Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer. *Atmospheric Chemistry and Physics*, 15(15), 8889–8973.
<https://doi.org/10.5194/acp-15-8889-2015>
- Moretti-Villegas, L. F., Tafur-Anzualdo, V. I., & Valiente-Saldaña, Y. M. (2023). Contaminación del aire en la ciudad de Lima, Perú. *Revista Arbitrada Interdisciplinaria Koinonía*, 8(2), 822–830. <https://doi.org/10.35381/r.k.v8i2.2975>
- Oujidi, B., Benchrif, A., Tahri, M., Zahry, F., Bounakhla, M., Bazairi, H., Mhammdi, N., & Snoussi, M. (2022). Gaseous Pollutants and Particulate Matter in Ambient Air : First Field Experiment in an Urban Mediterranean Area (Nador , Morocco). *Aerosol and Air Quality Research*, 23(8), 1–19. <https://doi.org/10.4209/aaqr.220451>
- Oyarzún, M. G. (2010). Contaminación aérea y sus efectos en la salud. *Revista Chilena de Enfermedades Respiratorias*, 26(1), 16–25.
- Paraschiv, S., Barbuta-Misu, N., & Paraschiv, S. L. (2020). Influence of NO2, NO and meteorological conditions on the tropospheric O3 concentration at an industrial station. *Energy Reports*, 6(2), 231–236. <https://doi.org/10.1016/j.egyr.2020.11.263>
- Peláez, L. M. G., Santos, J. M., Albuquerque, T. T. D. A., Reis, N. C., Andreao, W. L., &

- Andrade, M. de F. (2020). Air quality status and trends over large cities in South America. *Environmental Science and Policy*, 114(September), 422–435.
<https://doi.org/10.1016/j.envsci.2020.09.009>
- Quispe, J. J. N., Nina, J. H., & Jara, S. T. (2020). Study and application of wind rose for vertical and horizontal axis wind turbine installation in Lima. *Proceedings of the 2020 IEEE Engineering International Research Conference, EIRCON 2020*, 2(October 2020), 1–6. <https://doi.org/10.1109/EIRCON51178.2020.9254076>
- Rebaza, L. T., Paz, D. T., Rebaza, I. T., Choque, F. M., & Picardo, E. M. (2021). Análisis y evaluación de los niveles de CO y CO₂ en la Ciudad de Tacna en relación con el parque automotor y los estándares de calidad ambiental del aire. *ARQUITEK*, 20(1), 1–13.
<https://doi.org/10.47796/ra.2021i20>
- Roberts–Semple, D., Song, F., & Gao, Y. (2012). Seasonal characteristics of ambient nitrogen oxides and ground–level ozone in metropolitan northeastern New Jersey. *Atmospheric Pollution Research*, 3(2), 247–257. <https://doi.org/10.5094/APR.2012.027>
- Romero, Y., Diaz, C., Meldrum, I., Arias Velasquez, R., & Noel, J. (2020). Temporal and spatial analysis of traffic – Related pollutant under the influence of the seasonality and meteorological variables over an urban city in Peru. *Heliyon*, 6, e04029 Contents.
<https://doi.org/10.1016/j.heliyon.2020.e04029>
- Rowland, O. E. (2024). Comparative analysis of meteorological parameters and their relationship with NO₂, PM₁₀, PM_{2.5} and O₃ concentrations at selected urban air quality monitoring stations in Krakow, Paris, and Milan. *Discover Environment*, 2(1), 1–26.
<https://doi.org/10.1007/s44274-024-00060-2>
- Sachs, J. D. (2015). *The Age of Sustainable Development*.
- Saki, H., Goudarzi, G., Jalali, S., Barzegar, G., Farhadi, M., Parseh, I., Geravandi, S., Salmanzadeh, S., Yousefi, F., & Mohammadi, M. J. (2020). Study of relationship between nitrogen dioxide and chronic obstructive pulmonary disease in Bushehr, Iran. *Clinical Epidemiology and Global Health*, 8(2), 446–449.
<https://doi.org/10.1016/j.cegh.2019.10.006>
- Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, 126(5), 1763–1768.
<https://doi.org/10.1213/ANE.0000000000002864>
- Senamhi. (2023). *Datos Hidrometeorológicos a nivel nacional*. Servicio Nacional de Meteorología e Hidrología Del Perú.
- Shrestha, P. M., Poudyal, K. N., Chapagain, N. P., & Karki, I. B. (2020). Study of Affecting Factors of Meteorological Parameters on Solar Radiation in Kathmandu Valley. *The*

- Batuk*, 6(1), 72–80. <https://doi.org/10.3126/batuk.v6i1.32645>
- Silva, J. S., Rojas, J. P., Norabuena, M., & Seguel, R. J. (2018). Ozone and volatile organic compounds in the metropolitan area of Lima-Callao, Peru. *Air Quality, Atmosphere and Health*, 11(8), 993–1008. <https://doi.org/10.1007/s11869-018-0604-2>
- Tejeda, A., Méndez, I., Rodríguez, N. C., & Tejeda, E. (2018). La humedad en la atmósfera: Bases físicas, instrumentos y aplicaciones. In *La humedad en la atmosfera* (1st ed.). Universidad de Colima.
- Ventura, L. M. B., & Pinto, F. de O. (2013). *Inventário de emissões de fontes veiculares: região metropolitana do Rio de Janeiro*. 350.
- Zakaria, N. H., Salleh, S. A., Asmat, A., Chan, A., Isa, N. A., Hazali, N. A., & Islam, M. A. (2020). Analysis of Wind Speed, Humidity and Temperature: Variability and Trend in 2017. *IOP Conference Series: Earth and Environmental Science*, 489(1), 012013. <https://doi.org/10.1088/1755-1315/489/1/012013>
- Zender-Świercz, E., Galiszewska, B., Telejko, M., & Starzomska, M. (2024). The effect of temperature and humidity of air on the concentration of particulate matter - PM2.5 and PM10. *Atmospheric Research*, 312(October), 107733. <https://doi.org/10.1016/j.atmosres.2024.107733>
- Zhang, G., Band, S. S., Jun, C., Bateni, S. M., Chuang, H. M., Turabieh, H., Mafarja, M., Mosavi, A., & Moslehpour, M. (2021). Solar radiation estimation in different climates with meteorological variables using Bayesian model averaging and new soft computing models. *Energy Reports*, 7(2), 8973–8996. <https://doi.org/10.1016/j.egyr.2021.10.117>

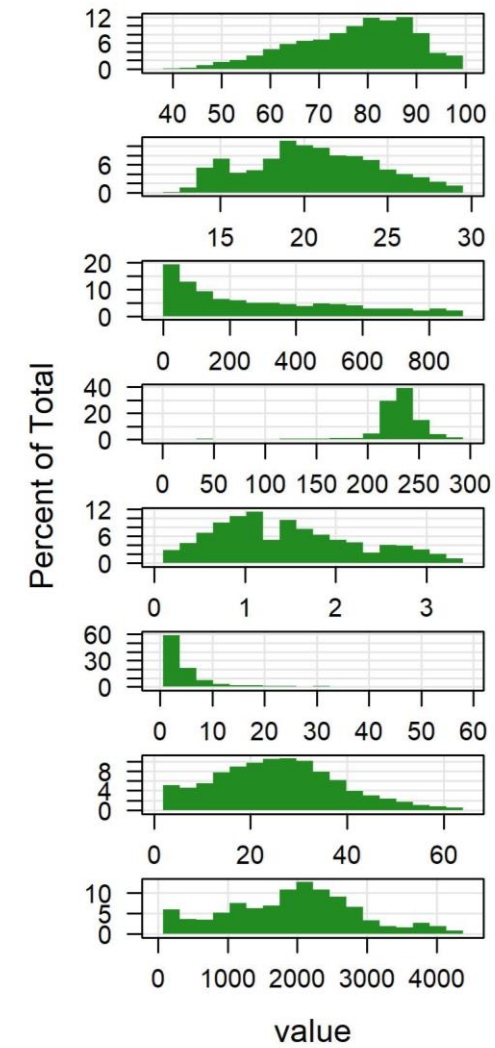
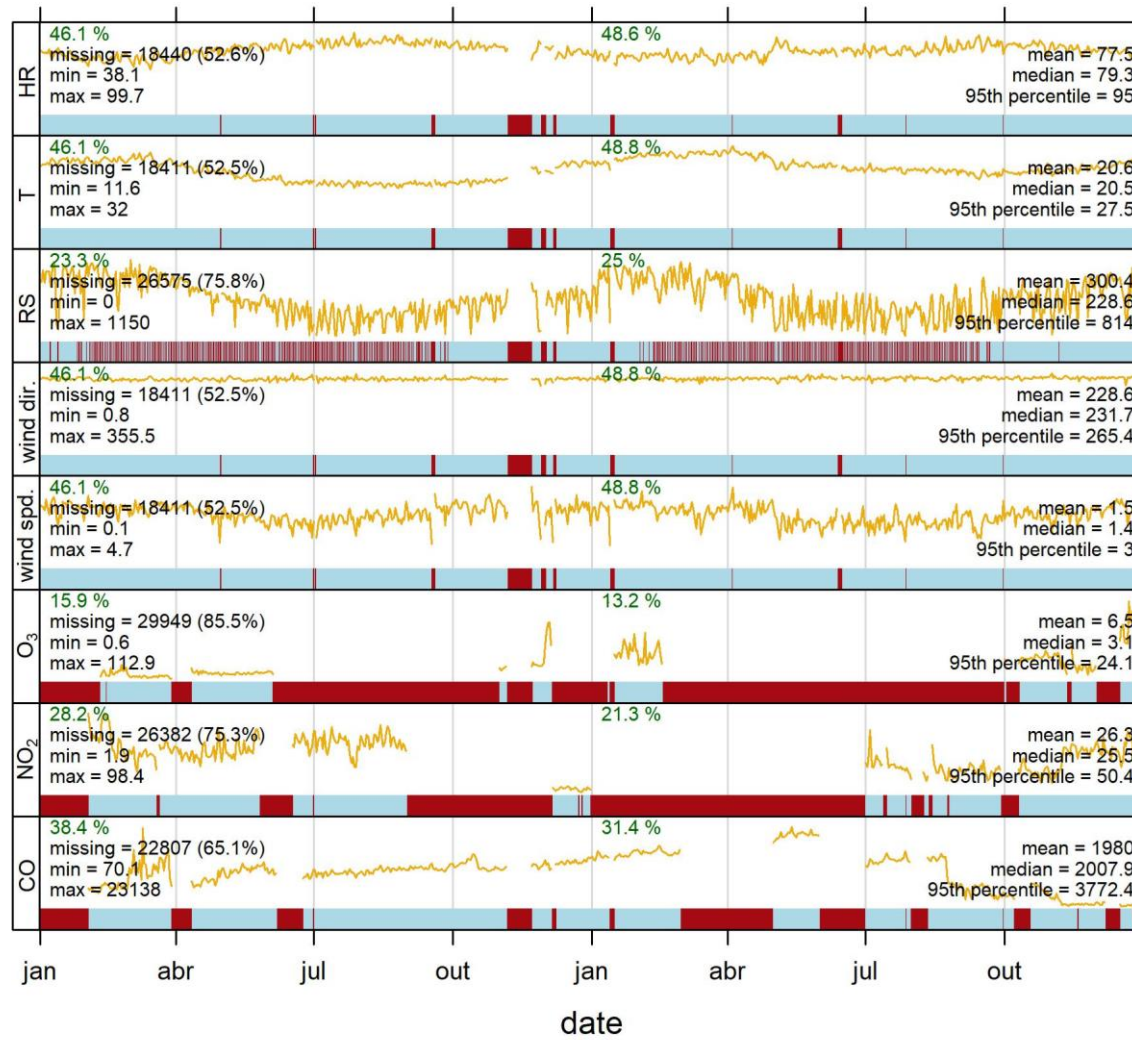
ANEXOS

Anexo 1. Matriz de consistencia

FORMULACIÓN DEL PROBLEMA	OBJETIVOS	HIPÓTESIS	VARIABLES, DIMENSIONES E INDICADORES	METODOLOGÍA DE LA INVESTIGACIÓN
Problema general	Objetivo general	Hipótesis general	Variable independiente:	a) Nivel, tipo y diseño de la investigación:
¿Cuál es la variabilidad temporal de los gases contaminantes atmosféricos (NO ₂ y O ₃) y su correlación con los parámetros meteorológicos en Lima Metropolitana, 2023-2024?	Determinar la variabilidad temporal de los gases contaminantes atmosféricos (NO ₂ y O ₃) y su relación con los parámetros meteorológicos en Lima Metropolitana.	H₀ = Existe relación de los gases contaminantes atmosféricos (NO ₂ y O ₃) en el análisis temporal con los parámetros meteorológicos en Lima Metropolitana H₁ : Existe relación de los gases contaminantes atmosféricos (NO ₂ y O ₃) en el análisis temporal con los parámetros meteorológicos en Lima Metropolitana.	<ul style="list-style-type: none"> Parámetros meteorológicos Dimensiones: <ul style="list-style-type: none"> Estaciones de monitoreo automático de los parámetros meteorológicos. Indicadores: <ul style="list-style-type: none"> Temperatura Radiación solar Velocidad y dirección del viento 	<ul style="list-style-type: none"> Nivel: básico. Tipo: Descriptiva y predictiva. Diseño: no experimental, longitudinal de tendencia. b) Población y muestra <ul style="list-style-type: none"> Población: Distritos de Lima Metropolitana. Muestra: distrito de Carabayllo. c) Técnicas e instrumentos de recolección de datos <ul style="list-style-type: none"> Solicitud de datos de contaminantes y parámetros meteorológicos al SENAMHI Ordenamiento de los datos y unión de los datos de contaminantes con los meteorológicos en el Microsoft Excel Alimentación de los datos ordenados al software libre R Generación de tablas descriptivas y de tendencia central usando el R Generación de gráficos horarios, diarios, días de la semana, fines de semana, mensuales, de estaciones y anuales.
Problemas específicos	Objetivos específicos	Hipótesis específicas	Variable dependiente:	
¿Cuál es la variación en concentración de los gases contaminantes atmosféricos (NO ₂ y O ₃) y características de los parámetros meteorológicos (temperatura, radiación sola, dirección y velocidad del viento) en Lima Metropolitana, 2023-2024? ¿Cuál es el resultado de la comparación de los gases contaminantes atmosféricos con el estándar de calidad ambiental, usando la normativa nacional vigente ¿Cuál es la correlación entre las variables meteorológicas y concentración de los gases	<ul style="list-style-type: none"> Analizar la variabilidad temporal de las variables meteorológicas y concentración de los gases contaminantes atmosféricos (NO₂ y O₃) Comparar la concentración de los gases contaminantes atmosféricos con el estándar de calidad ambiental, usando la normativa nacional vigente Determinar la correlación entre las variables meteorológicas y concentración de los gases 		<ul style="list-style-type: none"> Concentración de gases contaminantes Dimensiones: <ul style="list-style-type: none"> Medida de la concentración de NO₂ y O₃ y CO. Indicadores: <ul style="list-style-type: none"> Concentración NO₂ 	

contaminantes atmosféricos (NO_2 y O_3) mediante la correlación respectiva?	contaminantes atmosféricos (NO_2 y O_3) mediante la correlación respectiva		<ul style="list-style-type: none">• Concentración O_3• Concentración CO	
--	---	--	--	--

Anexo 3. Primer análisis exploratorio de los datos proporcionados por SENAMHI



Anexo 4. Imputación de datos

Figura 11

Evaluación para la imputación de datos del parámetro meteorológico humedad relativa (PPM-imputed distribution)

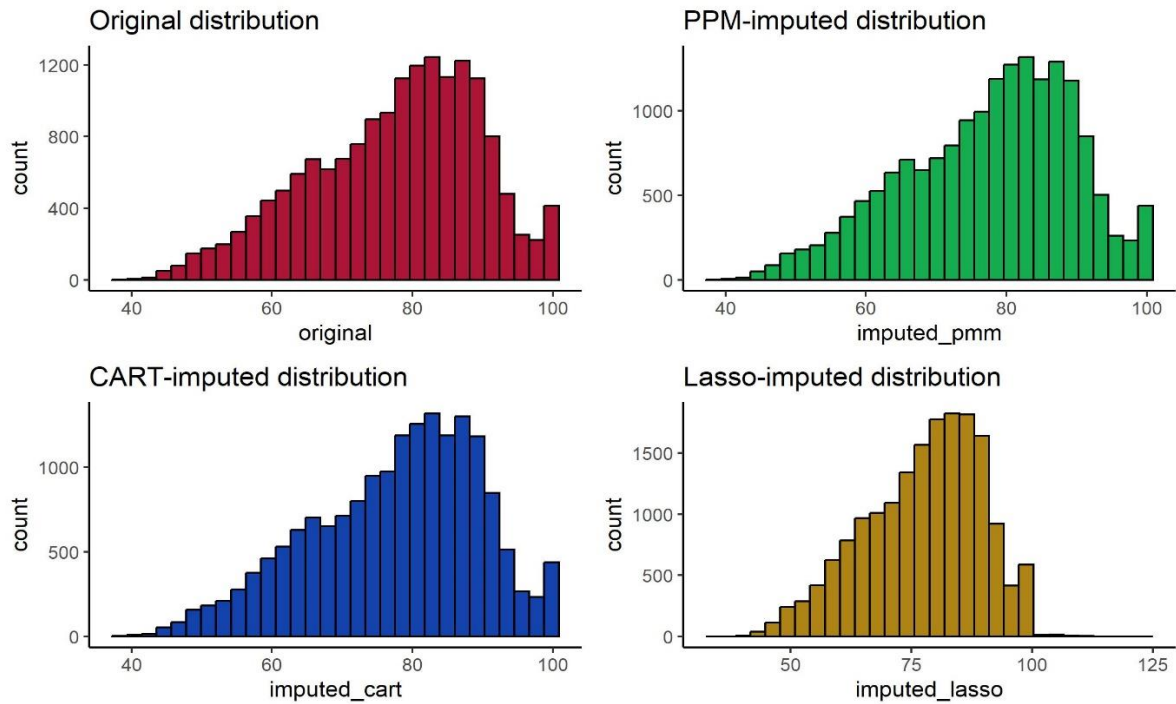


Figura 12

Evaluación para la imputación de datos del parámetro meteorológico temperatura (CART-imputed distribution)

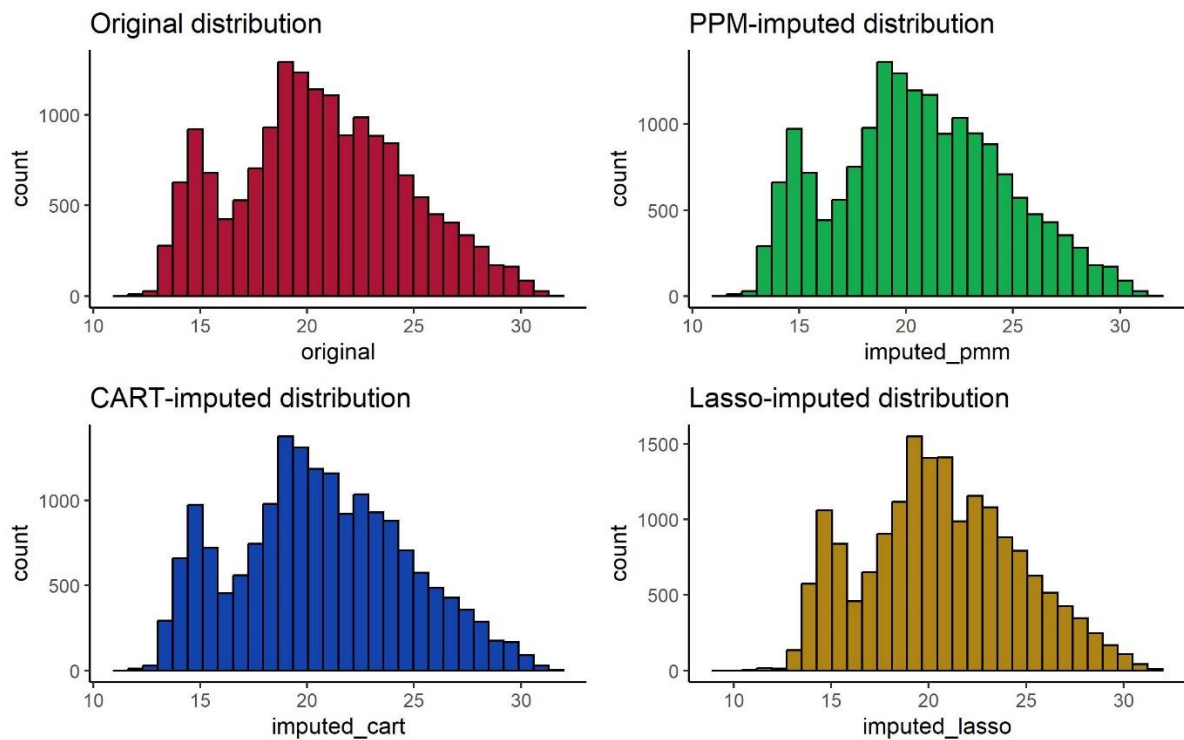
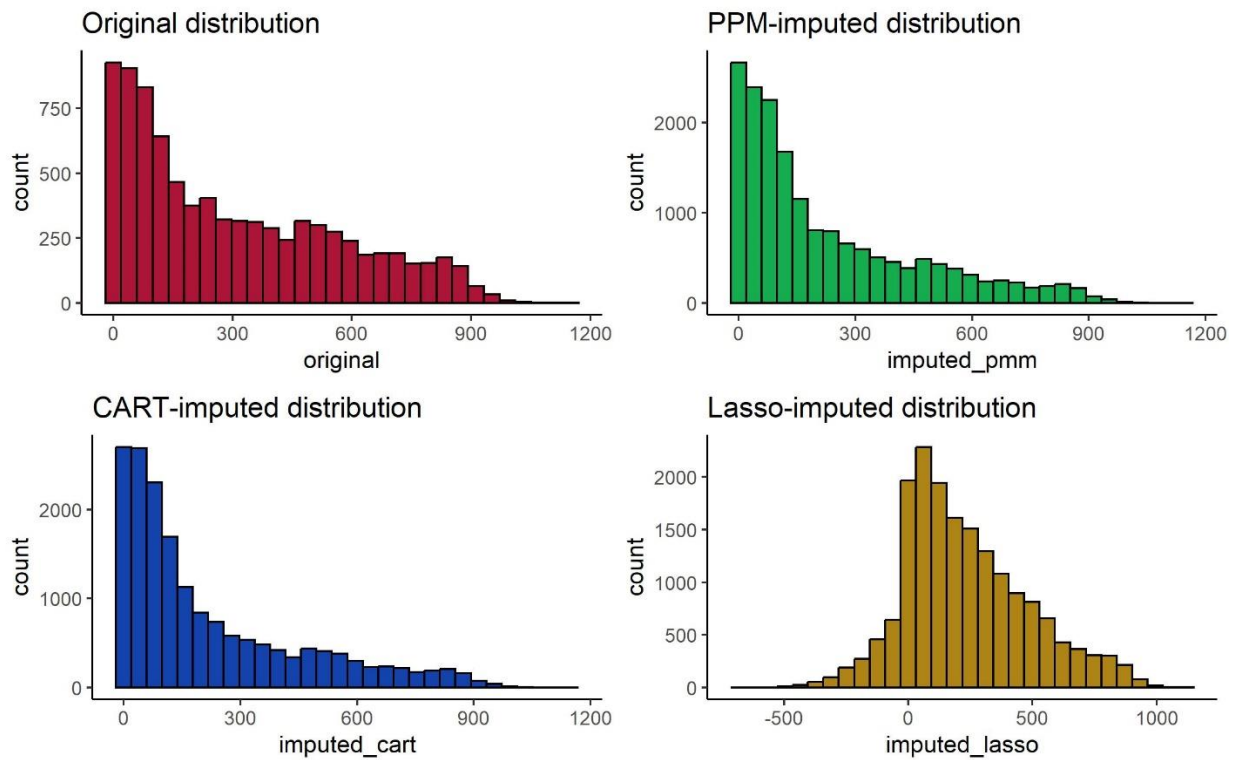


Figura 13

Evaluación para la imputación de datos del parámetro meteorológico radiación solar (CART-imputed distribution)

**Figura 14**

Evaluación para la imputación de datos del parámetro contaminante CO (FOREST)

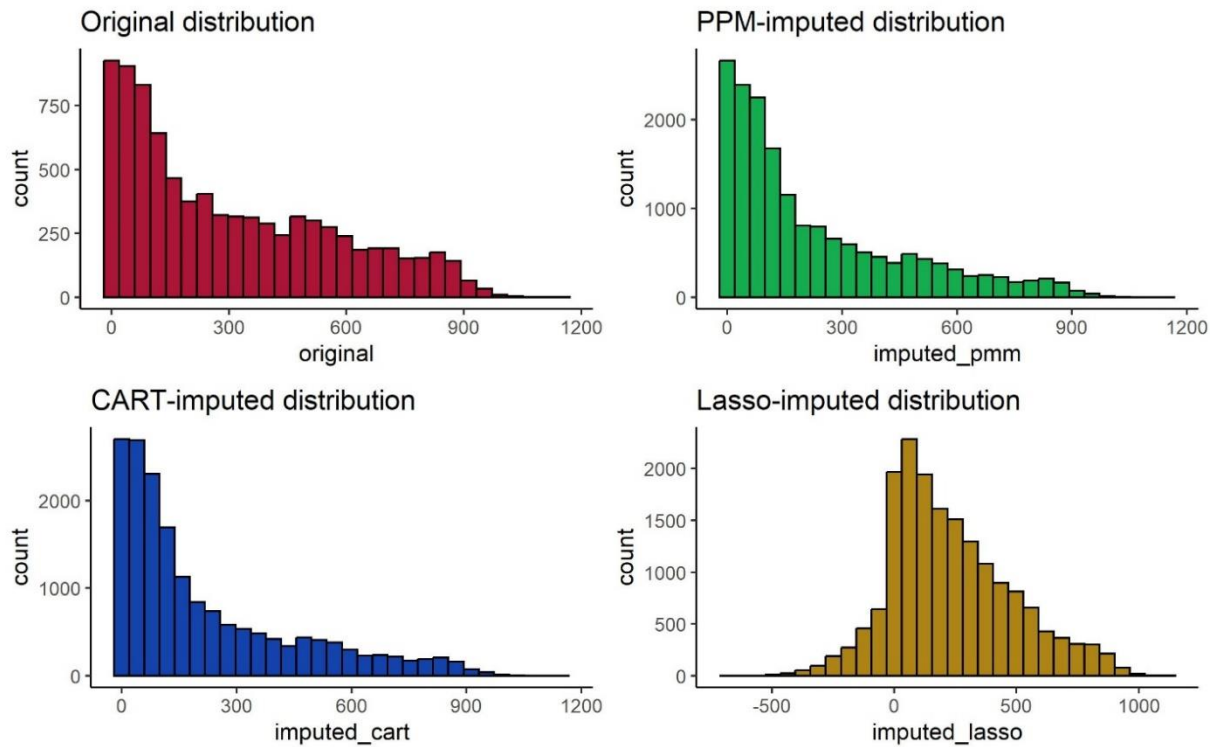
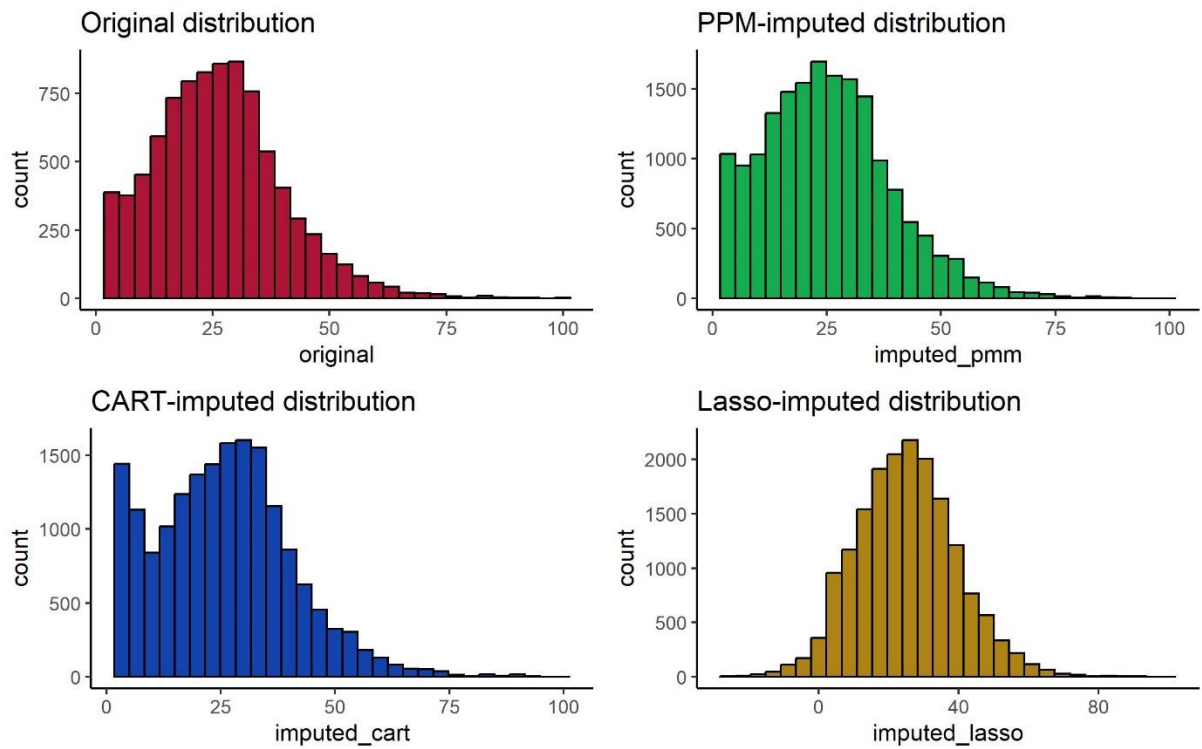


Figura 15

Evaluación para la imputación de datos del parámetro contaminante NO₂ (ppm – imputed distribution)

**Figura 16**

Evaluación para la imputación de datos del parámetro contaminante O₃ (CART – imputed distribution)

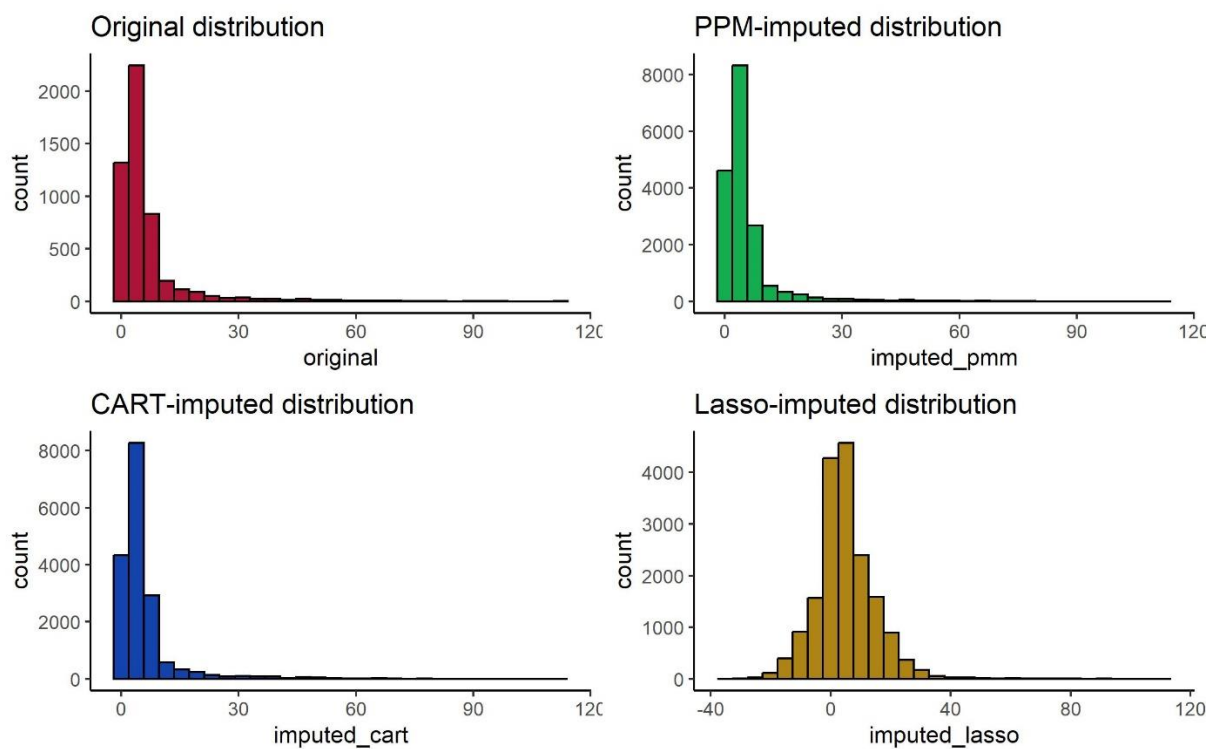
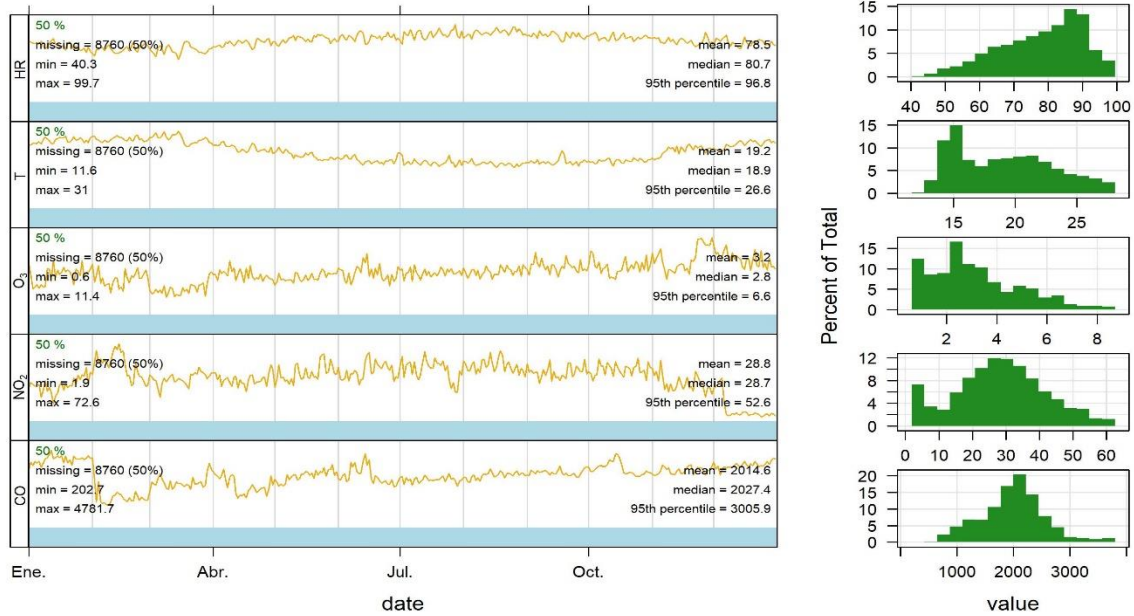


Figura 17

Distribución de datos en una serie temporal para los contaminantes atmosféricos gaseosos (CO , NO_2 , y O_3 en $\mu g / m^3$) y parámetros meteorológicos (T , HR y RS) durante el periodo de 2022. Son mostrados el valor máximo, valor mínimo, media y mediana para cada parámetro.

**Figura 18**

Distribución de datos en una serie temporal para los contaminantes atmosféricos gaseosos (CO , NO_2 , y O_3 en $\mu g / m^3$) y parámetros meteorológicos (T , HR y RS) durante el periodo de 2023. Son mostrados el valor máximo, valor mínimo, media y mediana para cada parámetro.

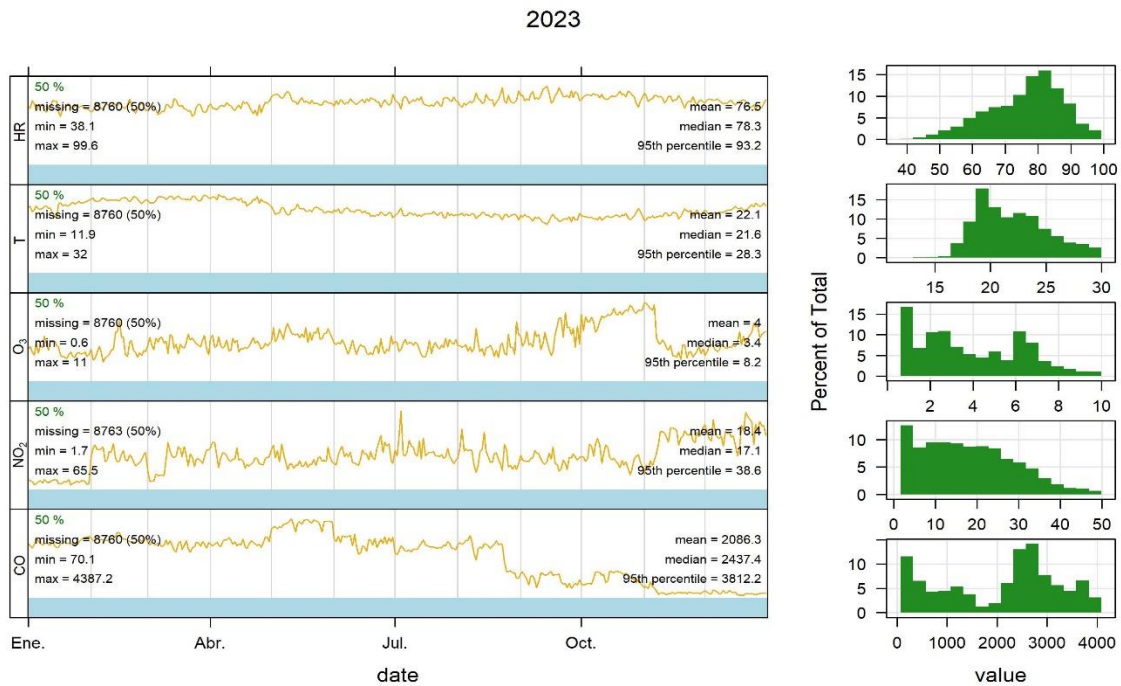


Figura 19

Comportamiento horario, mensual y diario de CO, NO₂, y O₃ para el periodo 2022-2023 de la estación de monitoreo de Carabayllo.

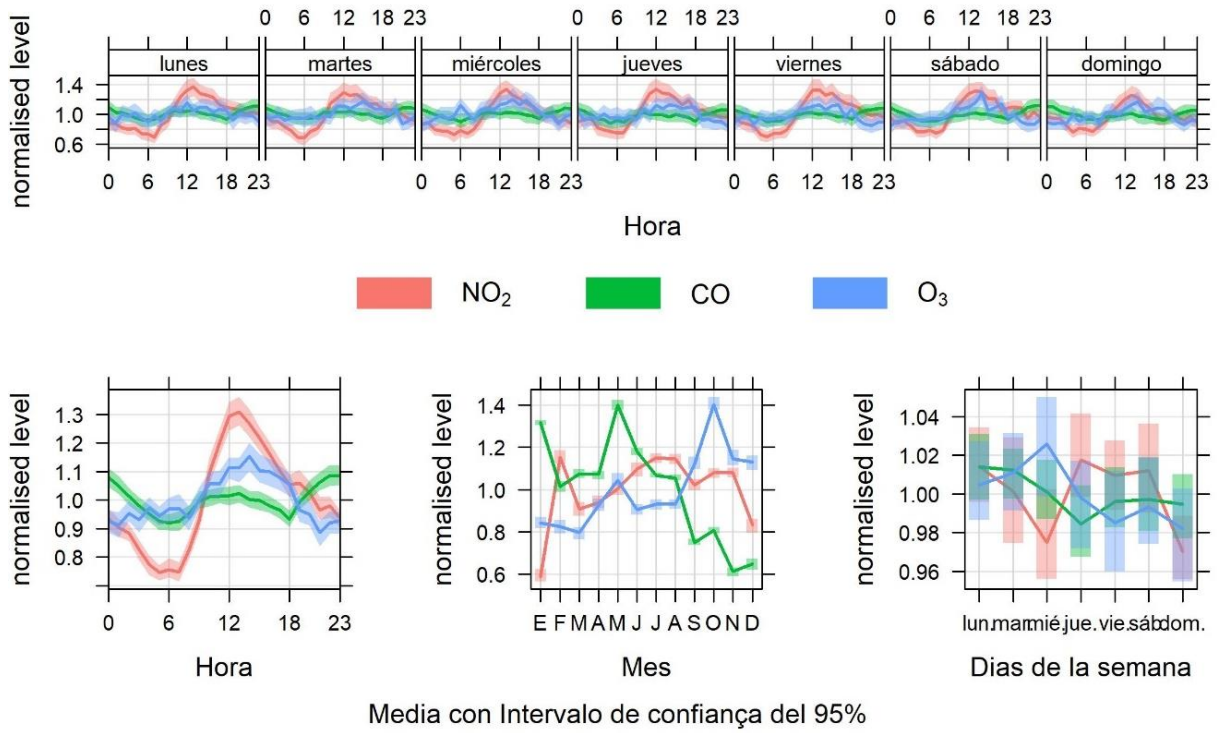


Figura 20

Visita técnica a la Estación de Monitoreo de Carabayllo



Anexo 6. Programación desarrollada en R para la imputación de los datos faltantes

```

library(ggplot2)
library(dplyr)
#install.packages("titanic")
library(titanic)
library(cowplot)
#data
Carabayllo <-read.csv("Carabayllo.csv",
                    header = TRUE, sep = ";", dec = ".")
aqStats(selectByDate(Carabayllo, year = 2022), pollutant = "CO")
#Distribution plot HR
ggplot(Carabayllo, aes(HR)) +
  geom_histogram(color = "#000000", fill = "#0099F8") +
  ggtitle("Variable distribution") +
  theme_classic() +
  theme(plot.title = element_text(size = 18))
# First method value_imputed
value_imputed <- data.frame(
  original = Carabayllo$HR,
  imputed_zero = replace(Carabayllo$HR, is.na(Carabayllo$HR), 0),
  imputed_mean = replace(Carabayllo$HR, is.na(Carabayllo$HR), mean(Carabayllo$HR,
na.rm = TRUE)),
  imputed_median = replace(Carabayllo$HR, is.na(Carabayllo$HR), median(Carabayllo$HR,
na.rm = TRUE))
)
value_imputed
#Plot distribution normal with value_imputed
HR1 <- ggplot(value_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
HR2 <- ggplot(value_imputed, aes(x = imputed_zero)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("Zero-imputed distribution") +
  theme_classic()
HR3 <- ggplot(value_imputed, aes(x = imputed_mean)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("Mean-imputed distribution") +
  theme_classic()
HR4 <- ggplot(value_imputed, aes(x = imputed_median)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Median-imputed distribution") +
  theme_classic()
plot_grid(HR1, HR2, HR3, HR4, nrow = 2, ncol = 2)
# das distribuicoes é observado que nao tem similaridade portantp procuramos outra
metodologia
# impute data Second method - MICE
#install.packages("mice")
library(mice)
Carabayllo_numeric <- Carabayllo %>%
  select(CO, NO2, O3, RS, HR)

```

```

md.pattern(Carabayllo_numeric)
mice_imputed <- data.frame(
  original = Carabayllo$HR,
  imputed_pmm = complete(mice(Carabayllo_numeric, method = "pmm"))$HR,
  imputed_cart = complete(mice(Carabayllo_numeric, method = "cart"))$HR,
  imputed_lasso = complete(mice(Carabayllo_numeric, method = "lasso.norm"))$HR
)
mice_imputed
#Plot distribution normal with value_imputed
HRm1 <- ggplot(mice_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
HRm2 <- ggplot(mice_imputed, aes(x = imputed_pmm)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("PPM-imputed distribution") +
  theme_classic()
HRm3 <- ggplot(mice_imputed, aes(x = imputed_cart)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("CART-imputed distribution") +
  theme_classic()
HRm4 <- ggplot(mice_imputed, aes(x = imputed_lasso)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Lasso-imputed distribution") +
  theme_classic()
jpeg("RH.jpg", width = 8, height = 5, units = "in", res = 300)
plot_grid(HRm1, HRm2, HRm3, HRm4, nrow = 2, ncol = 2)
dev.off()
#Export data in xls
#install.packages("openxlsx")
library(openxlsx)
# for writing a data.frame or list of data.frames to an xlsx file
write.xlsx(mice_imputed, 'ImputedppmHR.xlsx')
# impute data Third method - missForest
#install.packages("missForest")
library(missForest)
missForest_imputed <- data.frame(
  original = Carabayllo_numeric$HR,
  imputed_missForest = missForest(Carabayllo_numeric)$ximp$HR
)
missForest_imputed
#Plot distribution normal with missForest
HRf1 <- ggplot(missForest_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
HRf2 <- ggplot(missForest_imputed, aes(x = imputed_missForest)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("missForest-imputed distribution") +
  theme_classic()
plot_grid(HRf1, HRf2, nrow = 1, ncol = 2)

```

```

#Export data in CSV
write.csv(missForest_imputed, "forest.csv")

#Export data in xls
#install.packages("openxlsx")
library(openxlsx)
# for writing a data.frame or list of data.frames to an xlsx file
write.xlsx(missForest_imputed, 'forest.xlsx')
#Distribution plot Radiacion Solar
ggplot(Carabayllo, aes(RS)) +
  geom_histogram(color = "#000000", fill = "#0099F8") +
  ggtitle("Variable distribution") +
  theme_classic() +
  theme(plot.title = element_text(size = 18))
# First method value_imputed
value_imputed <- data.frame(
  original = Carabayllo$RS,
  imputed_zero = replace(Carabayllo$RS, is.na(Carabayllo$RS), 0),
  imputed_mean = replace(Carabayllo$RS, is.na(Carabayllo$RS), mean(Carabayllo$RS,
na.rm = TRUE)),
  imputed_median = replace(Carabayllo$RS, is.na(Carabayllo$RS), median(Carabayllo$RS,
na.rm = TRUE))
)
value_imputed
#Plot distribution normal with value_imputed
RS1 <- ggplot(value_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
RS2 <- ggplot(value_imputed, aes(x = imputed_zero)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("Zero-imputed distribution") +
  theme_classic()
RS3 <- ggplot(value_imputed, aes(x = imputed_mean)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("Mean-imputed distribution") +
  theme_classic()
RS4 <- ggplot(value_imputed, aes(x = imputed_median)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Median-imputed distribution") +
  theme_classic()
plot_grid(RS1, RS2, RS3, RS4, nrow = 2, ncol = 2)
# das distribuicoes é observado que nao tem similaridade portantp procuramos outra
metodologia
# impute data Second method - MICE
#install.packages("mice")
library(mice)
Carabayllo_numeric <- Carabayllo %>%
  select(CO, NO2, O3, RS, HR, Temp)
md.pattern(Carabayllo_numeric)
mice_imputed <- data.frame(

```

```

original = Carabayllo$RS,
imputed_pmm = complete(mice(Carabayllo_numeric, method = "pmm"))$RS,
imputed_cart = complete(mice(Carabayllo_numeric, method = "cart"))$RS,
imputed_lasso = complete(mice(Carabayllo_numeric, method = "lasso.norm"))$RS
)
mice_imputed
#Plot distribution normal with value_imputed
RSm1 <- ggplot(mice_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
RSm2 <- ggplot(mice_imputed, aes(x = imputed_pmm)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("PPM-imputed distribution") +
  theme_classic()
RSm3 <- ggplot(mice_imputed, aes(x = imputed_cart)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("CART-imputed distribution") +
  theme_classic()
RSm4 <- ggplot(mice_imputed, aes(x = imputed_lasso)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Lasso-imputed distribution") +
  theme_classic()
jpeg("RS.jpg", width = 8, height = 5, units = "in", res = 300)
plot_grid(RSm1, RSm2, RSm3, RSm4, nrow = 2, ncol = 2)
dev.off()
##selected: PPM-imputed distribution
# impute data Third method - missForest
#install.packages("missForest")
library(missForest)
missForest_imputed <- data.frame(
  original = Carabayllo_numeric$RS,
  imputed_missForest = missForest(Carabayllo_numeric)$ximp$RS
)
missForest_imputed
#Plot distribution normal with missForest
RSf1 <- ggplot(missForest_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
RSf2 <- ggplot(missForest_imputed, aes(x = imputed_missForest)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("missForest-imputed distribution") +
  theme_classic()
plot_grid(RSf1, RSf2, nrow = 1, ncol = 2)
#Export data in CSV
write.csv(mice_imputed, "imputedCARTTemp.csv")
#Export data in xls
#install.packages("openxlsx")
library(openxlsx)
# for writing a data.frame or list of data.frames to an excel file

```

```

write.xlsx(mice_imputed, 'imputedCARTRS.xlsx')
#Distribution plot CO
ggplot(Carabayllo, aes(CO)) +
  geom_histogram(color = "#000000", fill = "#0099F8") +
  ggtitle("Variable distribution") +
  theme_classic() +
  theme(plot.title = element_text(size = 18))
# First method value_imputed
value_imputed <- data.frame(
  original = Carabayllo$CO,
  imputed_zero = replace(Carabayllo$CO, is.na(Carabayllo$CO), 0),
  imputed_mean = replace(Carabayllo$CO, is.na(Carabayllo$CO), mean(Carabayllo$CO,
na.rm = TRUE)),
  imputed_median = replace(Carabayllo$CO, is.na(Carabayllo$CO), median(Carabayllo$CO,
na.rm = TRUE))
)
value_imputed
#Plot distribution normal with value_imputed
COh1 <- ggplot(value_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
COh2 <- ggplot(value_imputed, aes(x = imputed_zero)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("Zero-imputed distribution") +
  theme_classic()
COh3 <- ggplot(value_imputed, aes(x = imputed_mean)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("Mean-imputed distribution") +
  theme_classic()
COh4 <- ggplot(value_imputed, aes(x = imputed_median)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Median-imputed distribution") +
  theme_classic()
plot_grid(COh1, COh2, COh3, COh4, nrow = 2, ncol = 2)
# das distribuicoes é observado que nao tem similaridade portantp procuramos outra
metodologia
# impute data Second method - MICE
#install.packages("mice")
library(mice)
Carabayllo_numeric <- Carabayllo %>%
  select(CO, NO2, O3, RS, HR)
md.pattern(Carabayllo_numeric)
mice_imputed <- data.frame(
  original = Carabayllo$CO,
  imputed_pmm = complete(mice(Carabayllo_numeric, method = "pmm"))$CO,
  imputed_cart = complete(mice(Carabayllo_numeric, method = "cart"))$CO,
  imputed_lasso = complete(mice(Carabayllo_numeric, method = "lasso.norm"))$CO
)
mice_imputed
#Plot distribution normal with value_imputed

```

```

COm1 <- ggplot(mice_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
COm2 <- ggplot(mice_imputed, aes(x = imputed_pmm)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("PPM-imputed distribution") +
  theme_classic()
COm3 <- ggplot(mice_imputed, aes(x = imputed_cart)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("CART-imputed distribution") +
  theme_classic()
COm4 <- ggplot(mice_imputed, aes(x = imputed_lasso)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Lasso-imputed distribution") +
  theme_classic()
plot_grid(COm1, COm2, COm3, COm4, nrow = 2, ncol = 2)
# impute data Third method - missForest
#install.packages("missForest")
library(missForest)
missForest_imputed <- data.frame(
  original = Carabayllo_numeric$CO,
  imputed_missForest = missForest(Carabayllo_numeric)$ximp$CO
)
missForest_imputed
#Plot distribution normal with missForest
COf1 <- ggplot(missForest_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
COf2 <- ggplot(missForest_imputed, aes(x = imputed_missForest)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("missForest-imputed distribution") +
  theme_classic()
plot_grid(COf1, COf2, nrow = 1, ncol = 2)
#Export data in CSV
write.csv(missForest_imputed, "forest.csv")
#Export data in xls
install.packages("openxlsx")
library(openxlsx)
# for writing a data.frame or list of data.frames to an xls file
write.xlsx(missForest_imputed, 'forest.xlsx')
#Distribution plot NO2
ggplot(Carabayllo, aes(NO2)) +
  geom_histogram(color = "#000000", fill = "#0099F8") +
  ggtitle("Variable distribution") +
  theme_classic() +
  theme(plot.title = element_text(size = 18))
# First method value_imputed
value_imputed <- data.frame(
  original = Carabayllo$NO2,

```

```

imputed_zero = replace(Carabayllo$NO2, is.na(Carabayllo$NO2), 0),
imputed_mean = replace(Carabayllo$NO2, is.na(Carabayllo$NO2), mean(Carabayllo$NO2,
na.rm = TRUE)),
imputed_median = replace(Carabayllo$NO2, is.na(Carabayllo$NO2),
median(Carabayllo$NO2, na.rm = TRUE))
)
value_imputed
#Plot distribution normal with value_imputed
NO21 <- ggplot(value_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
NO22 <- ggplot(value_imputed, aes(x = imputed_zero)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("Zero-imputed distribution") +
  theme_classic()
NO23 <- ggplot(value_imputed, aes(x = imputed_mean)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("Mean-imputed distribution") +
  theme_classic()
NO24 <- ggplot(value_imputed, aes(x = imputed_median)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Median-imputed distribution") +
  theme_classic()
plot_grid(NO21, NO22, NO23, NO24, nrow = 2, ncol = 2)
# das distribuicoes é observado que nao tem similaridade portantp procuramos outra
metodologia
# impute data Second method - MICE
#install.packages("mice")
library(mice)
Carabayllo_numeric <- Carabayllo %>%
  select(CO, NO2, O3, RS, HR, Temp)
md.pattern(Carabayllo_numeric)
mice_imputed <- data.frame(
  original = Carabayllo$NO2,
  imputed_pmm = complete(mice(Carabayllo_numeric, method = "pmm"))$NO2,
  imputed_cart = complete(mice(Carabayllo_numeric, method = "cart"))$NO2,
  imputed_lasso = complete(mice(Carabayllo_numeric, method = "lasso.norm"))$NO2
)
mice_imputed
#Plot distribution normal with value_imputed
NO2m1 <- ggplot(mice_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
NO2m2 <- ggplot(mice_imputed, aes(x = imputed_pmm)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("PPM-imputed distribution") +
  theme_classic()
NO2m3 <- ggplot(mice_imputed, aes(x = imputed_cart)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +

```

```

ggtitle("CART-imputed distribution") +
theme_classic()
NO2m4 <- ggplot(mice_imputed, aes(x = imputed_lasso)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Lasso-imputed distribution") +
  theme_classic()
jpeg("NO2.jpg", width = 8, height = 5, units = "in", res = 300)
plot_grid(NO2m1, NO2m2, NO2m3, NO2m4, nrow = 2, ncol = 2)
dev.off()
##selected: PPM-imputed distribution
# impute data Third method - missForest
#install.packages("missForest")
library(missForest)
missForest_imputed <- data.frame(
  original = Carabayllo_numeric$NO2,
  imputed_missForest = missForest(Carabayllo_numeric)$ximp$NO2
)
missForest_imputed
#Plot distribution normal with missForest
NO2f1 <- ggplot(missForest_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
NO2f2 <- ggplot(missForest_imputed, aes(x = imputed_missForest)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("missForest-imputed distribution") +
  theme_classic()
plot_grid(NO2f1, NO2f2, nrow = 1, ncol = 2)
#Export data in CSV
write.csv(missForest_imputed, "forest.csv")
#Export data in xls
#install.packages("openxlsx")
library(openxlsx)
# for writing a data.frame or list of data.frames to an xlsx file
write.xlsx(mice_imputed, 'ppmimputedNO2.xlsx')
#Distribution plot O3
ggplot(Carabayllo, aes(O3)) +
  geom_histogram(color = "#000000", fill = "#0099F8") +
  ggtitle("Variable distribution") +
  theme_classic() +
  theme(plot.title = element_text(size = 18))
# First method value_imputed
value_imputed <- data.frame(
  original = Carabayllo$O3,
  imputed_zero = replace(Carabayllo$O3, is.na(Carabayllo$O3), 0),
  imputed_mean = replace(Carabayllo$O3, is.na(Carabayllo$O3), mean(Carabayllo$O3,
na.rm = TRUE)),
  imputed_median = replace(Carabayllo$O3, is.na(Carabayllo$O3), median(Carabayllo$O3,
na.rm = TRUE))
)
value_imputed

```

```

#Plot distribution normal with value_imputed
O31 <- ggplot(value_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
O32 <- ggplot(value_imputed, aes(x = imputed_zero)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("Zero-imputed distribution") +
  theme_classic()
O33 <- ggplot(value_imputed, aes(x = imputed_mean)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("Mean-imputed distribution") +
  theme_classic()
O34 <- ggplot(value_imputed, aes(x = imputed_median)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Median-imputed distribution") +
  theme_classic()
plot_grid(O31, O32, O33, O34, nrow = 2, ncol = 2)
# das distribuicoes é observado que nao tem similaridade portantp procuramos outra
metodologia
# impute data Second method - MICE
#install.packages("mice")
library(mice)
Carabayllo_numeric <- Carabayllo %>%
  select(CO, NO2, O3, RS, HR)
md.pattern(Carabayllo_numeric)
mice_imputed <- data.frame(
  original = Carabayllo$O3,
  imputed_pmm = complete(mice(Carabayllo_numeric, method = "pmm"))$O3,
  imputed_cart = complete(mice(Carabayllo_numeric, method = "cart"))$O3,
  imputed_lasso = complete(mice(Carabayllo_numeric, method = "lasso.norm"))$O3
)
mice_imputed
#Plot distribution normal with value_imputed
O3m1 <- ggplot(mice_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
O3m2 <- ggplot(mice_imputed, aes(x = imputed_pmm)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("PPM-imputed distribution") +
  theme_classic()
O3m3 <- ggplot(mice_imputed, aes(x = imputed_cart)) +
  geom_histogram(fill = "#1543ad", color = "#000000", position = "identity") +
  ggtitle("CART-imputed distribution") +
  theme_classic()
O3m4 <- ggplot(mice_imputed, aes(x = imputed_lasso)) +
  geom_histogram(fill = "#ad8415", color = "#000000", position = "identity") +
  ggtitle("Lasso-imputed distribution") +
  theme_classic()
jpeg("O3.jpg", width = 8, height = 5, units = "in", res = 300)

```

```

plot_grid(O3m1, O3m2, O3m3, O3m4, nrow = 2, ncol = 2)
dev.off()
##selected: PPM-imputed distribution
# impute data Third method - missForest
#install.packages("missForest")
library(missForest)
missForest_imputed <- data.frame(
  original = Carabayllo_numeric$O3,
  imputed_missForest = missForest(Carabayllo_numeric)$ximp$O3
)
missForest_imputed
#Plot distribution normal with missForest
O3f1 <- ggplot(missForest_imputed, aes(x = original)) +
  geom_histogram(fill = "#ad1538", color = "#000000", position = "identity") +
  ggtitle("Original distribution") +
  theme_classic()
O3f2 <- ggplot(missForest_imputed, aes(x = imputed_missForest)) +
  geom_histogram(fill = "#15ad4f", color = "#000000", position = "identity") +
  ggtitle("missForest-imputed distribution") +
  theme_classic()
plot_grid(O3f1, O3f2, nrow = 1, ncol = 2)
#Export data in CSV
write.csv(missForest_imputed, "forest.csv")
#Export data in xls
#install.packages("openxlsx")
library(openxlsx)
# for writing a data.frame or list of data.frames to an xlsx file
write.xlsx(mice_imputed, 'CartO3.xlsx')

```

Anexo 7. Programación desarrollada en R utilizando el paquete openair ()

```

library(openair)
library(dplyr)
# importar los datos
CarabaylloImp <- read.csv("CarabaylloImputed.csv",
  header = TRUE, sep = ";", dec = ".")
# configurar fecha con zona horaria correcta
CarabaylloImp$date <- as.POSIXct(as.character(CarabaylloImp$date),
  format = "%d/%m/%Y %H:%M",
  tz = "America/Lima")
#Create a summary of descriptive statistics;
summary(selectByDate(CarabaylloImp, year = 2022)) # statistics by year 2022
summary(selectByDate(CarabaylloImp, year = 2023))
summary(CarabaylloImp)
#Select a respective period
#mydata2022 <- cutData(CarabaylloImp, type = "year")
# OlsonNames() ofrece zonas horarias de todo mundo
#OlsonNames()
# Function summaryPlot
jpeg("distributionall.jpg", width = 7, height = 5, units = "in", res = 300)
summaryPlot(CarabaylloImp[,1:6]) #plot all years together
dev.off()

```

```

jpeg("distribution2022.jpg", width = 9, height = 6, units = "in", res = 300)
summaryPlot(selectByDate(CarabaylloImp[,1:6], year = 2022), main="2022") #plot only the
designed year
dev.off()
jpeg("distribution2023.jpg", width = 9, height = 6, units = "in", res = 300)
summaryPlot(selectByDate(CarabaylloImp[,1:6], year = 2023), main="2023") #plot only the
designed year
dev.off()
#summaryPlot(CarabaylloImp, period = "months") #plot by months
dev.off()
summaryPlot(Carabayllo[,1:4]) # muestra las tres primeras variables sin el 1ro
#https://github.com/openair-project/openair/blob/master/R/summaryPlot.R
summaryPlot(CarabaylloImp, col.data = "green", col.mis = "blue") #colours data and
missdata
summaryPlot(CarabaylloImp, col.data = "green", col.mis = "blue", col.hist = "red")
summaryPlot(CarabaylloImp, col.data = "green", col.mis = "blue", col.hist = "red",
  col.trend = "yellow")
summaryPlot(CarabaylloImp, col.data = "green", col.mis = "blue", col.hist = "red",
  period = "years")
summaryPlot(CarabaylloImp, col.data = "green", col.mis = "blue", col.hist = "red",
  period = "months")
# Function windrose
windRose(CarabaylloImp)
windRose(CarabaylloImp, type = "year")
windRose(CarabaylloImp, type = "season")
windRose(CarabaylloImp, type = "month")
windRose(CarabaylloImp, type = c("season", "weekday"))
windRose(CarabaylloImp, type = "year", cols = c("blue", "red", "green")) #change colours
jpeg("rosavientos.jpg", width = 6, height = 4, units = "in", res = 300)
windRose(CarabaylloImp, type = "year", cols = c("blue", "red", "green"),
  grid.line = list(value=10, lty=5, col="purple")) # values put ranges, lty=typeline
dev.off()
#windrose with pollutants
windRose(Carabayllo, type = "O3", layout = c(4, 1)) #change colours
windRose(Carabayllo, type = "year", angle = 22.5)
# Function Pollutionrose
pollutionRose(Carabayllo, pollutant = "CO", type="year")
pollutionRose(Carabayllo, pollutant = "NO2", type='O3', layout=c(4,1)) #read more!!!
#PolarPlot
polarPlot(Carabayllo, pollutant = "CO", type="year")
polarPlot(Carabayllo, pollutant = "CO", type="year", key.position = "bottom",
  key.header = "mean CO (ug/m3)", key.footer = NULL)
polarPlot(Carabayllo, pollutant = "CO", type = "hour", min.bin = 2)
polarPlot(Carabayllo, pollutant = "CO", statistic = "nwr")
polarPlot(Carabayllo, pollutant = "CO", type="year", statistic = "nwr")
#Function Polarannulus
polarAnnulus(Carabayllo, pollutant = "CO")
polarAnnulus(Carabayllo, pollutant = "CO", period = "season", Main="Season")
polarAnnulus(Carabayllo, pollutant = "CO", period = "weekday", Main="Weekday")
polarAnnulus(Carabayllo, pollutant = "CO", period = "hour", Main="Hour")
# Function Timeplot

```

```

timePlot(Carabayllo, pollutant = c("NO2", "O3"), y.relation = "free")
## plot monthly means of ozone and no2
timePlot(Carabayllo, pollutant = c("O3", "NO2"), avg.time = "month",
  y.relation = "free")
# Function TimeVariation
Sys.setlocale("LC_ALL", "en_US.UTF-8") #set titleplots in english
Sys.setlocale("LC_ALL", "es") #set titleplots in spaniss, is showed only7days
Sys.setlocale("LC_ALL", "pt_BR") #set titleplots in spaniss, is showed only5days
jpeg("timevariationall.jpg", width = 8, height = 5, units = "in", res = 300)
timeVariation(CarabaylloImp, pollutant = c("NO2", "CO", "O3"),
  normalise = TRUE, ylab="h",
  xlab = c("Hora", "Hora", "Mes", "Dias de la semana"),
  sub = "Media con Intervalo de confianza del 95%")
dev.off()
jpeg("timevariationCO.jpg", width = 13, height = 6, units = "in", res = 300)
timeVariation(CarabaylloImp, pollutant = "CO", statistic = "median",
  ylab = "CO (ug/m3)",
  col = "green", xlab = c("Hora", "Hora", "Mes", "Dias de la semana"),
  sub = "Media con Intervalo de confianza del 95%", fontsize=15)
dev.off()
jpeg("timevariationNO2.jpg", width = 13, height = 6, units = "in", res = 300)
timeVariation(CarabaylloImp, pollutant = "NO2", statistic = "median", ylab = "NO2 (ug/m3)",
  col = "yellow", xlab = c("Hora", "Hora", "Mes", "Dias de la semana"),
  sub = "Media con Intervalo de confianza del 95%", fontsize=15)
dev.off()
jpeg("timevariationO3.jpg", width = 13, height = 6, units = "in", res = 300)
timeVariation(CarabaylloImp, pollutant = "O3", statistic = "median", ylab = "O3 (ug/m3)",
  col = "blue", xlab = c("Hora", "Hora", "Mes", "Dias de la semana"),
  sub = "Media con Intervalo de confianza del 95%", fontsize=15)
dev.off()
#intall hexbin packages
install.packages("hexbin")
library(hexbin)
scatterPlot(Carabayllo, x = "O3", y = "NO2", method = "hexbin", col= "jet")
scatterPlot(Carabayllo, x = "T", y = "O3", method = "hexbin", col= "jet")
scatterPlot(Carabayllo, x = "CO", y = "NO2", method = "hexbin", col= "jet")
scatterPlot(Carabayllo, x = "O3", y = "RS", method = "hexbin", col= "jet")
scatterPlot(Carabayllo, x = "O3", y = "HR", method = "density", col= "jet")
scatterPlot(Carabayllo, x = "O3", y = "HR", smooth = FALSE, linear=TRUE,col= "jet")
##TrendLevel
trendLevel(CarabaylloImp, x = "season", y = "daylight", pollutant = "NO2")
trendLevel(Carabayllo, x = "season", y = "daylight", pollutant = "O3")
trendLevel(Carabayllo, x = "season", y = "daylight", pollutant = "CO")
trendLevel(Carabayllo, x = "season", y = "hour", pollutant = "NO2",
  cols = "increment")
trendLevel(SanBorja, x = "season", y = "hour", pollutant = "PM2.5",
  cols = "increment")
jpeg("trendO3.jpg", width = 13, height = 6, units = "in", res = 300)
trendLevel(CarabaylloImp, pollutant = "O3",
  border = "white", statistic = "max",
  breaks = c(0, 50, 100, 168, 208),

```

```
      labels = c("Good", "Moderate", "high", "unhealthy groups"),
      cols = c("forestgreen", "yellow", "orange", "red"))
dev.off()
jpeg("trendCO.jpg", width = 13, height = 6, units = "in", res = 300)
trendLevel(CarabaylloImp, pollutant = "CO",
  border = "white", statistic = "max",
  breaks = c(0, 1000, 2000, 10000, 17000),
  labels = c("Good", "Moderate", "high", "unhealthy groups"),
  cols = c("forestgreen", "yellow", "orange", "red"))
dev.off()
jpeg("trendNO2.jpg", width = 13, height = 6, units = "in", res = 300)
trendLevel(CarabaylloImp, pollutant = "NO2",
  border = "white", statistic = "max",
  breaks = c(0, 40, 80, 180, 280),
  labels = c("Good", "Moderate", "high", "unhealthy groups"),
  cols = c("forestgreen", "yellow", "orange", "red"))
dev.off()
# Function Corplot
jpeg("corr.jpg", width = 6, height = 6, units = "in", res = 300)
corPlot(CarabaylloImp, dendrogram = TRUE)
dev.off()
```