



Tecnológico de Monterrey

Impacto de áreas verdes en la calidad del aire en Monterrey

Mario Alberto Landa Flores

Rolando Ruiz Martínez

Luis Roberto Garza Sánchez

Septiembre 07, 2024

Aplicación de métodos multivariados en ciencia de datos

Asesores:

Blanca Rosa Ruiz Hernández

Rodolfo Fernandez de Lara Hadad

1. Resumen

Esta investigación tiene como objetivo explorar la relación entre la cantidad de áreas verdes y la calidad del aire en Monterrey durante el año 2022, una ciudad industrializada que enfrenta importantes problemas de contaminación. Partiendo de la premisa de que las zonas con mayor cobertura vegetal podrían tener menores concentraciones de contaminantes atmosféricos, se busca identificar si la vegetación influye en la mitigación de la contaminación del aire y si puede considerarse un factor determinante en la mejora de la calidad del aire en áreas urbanas.

Para la realización de este estudio, se utilizaron datos proporcionados por el Sistema Integral de Monitoreo Ambiental (SIMA), que monitorea continuamente la calidad del aire a través de una red de estaciones distribuidas en diferentes puntos de Monterrey y su área metropolitana. Los datos incluyen mediciones de varios contaminantes clave, tales como el material particulado (PM₁₀ y PM_{2.5}), dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂) y ozono (O₃), que son bien conocidos por sus efectos adversos en la salud pública y el medio ambiente.

Además de los datos sobre la calidad del aire, se recopilaron imágenes satelitales mediante Google Earth Engine para calcular el porcentaje de áreas verdes alrededor de cada estación de monitoreo. Las imágenes fueron procesadas con técnicas automatizadas utilizando Python y la librería rasterio para identificar la vegetación en un radio de 2.5 km desde cada estación. Este análisis permitió clasificar las estaciones en dos grupos: aquellas con alto y bajo nivel de áreas verdes. Se excluyeron las estaciones cuyos datos no se alineaban con ninguna de estas clasificaciones.

Una vez agrupados los datos, se procedió a un análisis exploratorio de las concentraciones de los contaminantes en cada una de las estaciones, con el fin de identificar posibles patrones y relaciones entre la calidad del aire y la cantidad de vegetación. Para modelar esta relación, se utilizó un enfoque de efectos mixtos, donde el porcentaje de áreas verdes se trató como un efecto fijo y el tiempo como un efecto aleatorio, permitiendo controlar las variaciones temporales en la concentración de los contaminantes. Además, para mejorar la robustez de los resultados, se aplicó una transformación de Box-Cox a los datos de concentración, asegurando una distribución más cercana a la normalidad y facilitando la interpretación de los resultados.

Los hallazgos del estudio indican que la vegetación puede tener un impacto positivo en la reducción de algunos contaminantes. Los resultados mostraron que los niveles de dióxido de nitrógeno (NO₂), óxidos de nitrógeno (NO_x) y partículas en suspensión (PM₁₀) tienden a ser más bajos en las estaciones rodeadas de un mayor porcentaje de áreas verdes. Esto sugiere que los árboles y las plantas contribuyen a la filtración del aire y la reducción de la concentración de estos contaminantes. Sin embargo, el comportamiento de otros contaminantes, como el ozono (O₃), fue inverso, mostrando mayores concentraciones en áreas con más vegetación, lo que indica que factores adicionales en el ambiente también pueden influir en la calidad del aire.

En otras palabras, el estudio aporta evidencia significativa sobre el papel de las áreas verdes en la calidad del aire en Monterrey. Los resultados refuerzan la necesidad de

implementar políticas públicas que fomenten la creación y preservación de zonas verdes en áreas urbanas. Finalmente se abren las puertas a nuevas investigaciones relacionadas a los fenómenos observados de concentración y residuales.

2. Introducción y justificación

La calidad del aire es uno de los factores más cruciales para garantizar una vida sana y estable. Tanto para la comunidad científica como para la ciudadanía en general, la calidad del aire representa un tema de suma importancia, ya que afecta directamente la salud pública, el medio ambiente y la calidad de vida en las ciudades. La mala calidad del aire se considera uno de los principales riesgos para la salud integral de una persona. La exposición a la contaminación, día tras día, puede causar enfermedades graves, colocando nuestra salud en una situación crítica.

La contaminación atmosférica no solo afecta a los seres humanos, sino que también afecta gravemente a las plantas y a todos los ecosistemas. Los contaminantes, como el material particulado (PM_{2.5} y PM₁₀), el ozono troposférico y los gases tóxicos como el dióxido de azufre (SO₂) y los óxidos de nitrógeno (NO₂), tienen efectos devastadores sobre la vegetación, ya que afectan en gran medida sus procesos de naturaleza (*Primagas, 2020*).

Los contaminantes más dañinos para la salud incluyen el material particulado (PM_{2.5} y PM₁₀), que puede penetrar profundamente en los pulmones y el sistema cardiovascular, causando enfermedades respiratorias, cardiovasculares e incluso cáncer. El ozono troposférico es otro contaminante peligroso, ya que puede provocar problemas respiratorios, reducir la función pulmonar y agravar enfermedades como el asma. Además, el dióxido de nitrógeno (NO₂) está asociado con enfermedades respiratorias y una disminución en la función pulmonar, mientras que el monóxido de carbono (CO) puede interferir con la capacidad de la sangre para transportar oxígeno, lo que resulta en efectos cardiovasculares y neurológicos. Por último, el dióxido de azufre (SO₂) puede irritar el sistema respiratorio y agravar condiciones preexistentes como el asma.

La calidad del aire se mide con diferentes equipos tecnológicos que permiten registrar la cantidad de contaminantes y clasificarlos de acuerdo a la muestra en diferentes áreas atmosféricas. En el estado de Nuevo León, existen una variedad de tecnologías para realizar mediciones de los contaminantes del aire en distintas áreas geográficas. Entre ellos existen estaciones de monitoreo fijas, las cuales están equipadas con sensores para medir la cantidad de contaminantes criterio (partículas menores a 10 micrómetros y menores a 2.5 micrómetros) conocidas como dióxido de azufre (SO₂), dióxido de nitrógeno (NO₂), ozono (O₃) y monóxido de carbono (CO) (*Mayora, 2019*).

Después de la recopilación de los datos, se implementan procedimientos de calibración y control de calidad en sus estaciones para asegurar precisión y confiabilidad de los datos. En un tablero de control se lleva la supervisión en tiempo real de los datos de calidad del aire recolectados por las estaciones. El tablero facilita la visualización de los niveles de contaminantes y parámetros meteorológicos para poder identificar una anomalía o tendencia preocupante en la contaminación. Esto con el objetivo de que se cumplan los estándares de calidad establecidos por el gobierno (*Jorge Delgado, Centro, y De Metrología, s.f.*).

En el estado de Nuevo León, México, se monitorea la calidad del aire a través del Sistema Integral de Monitoreo Ambiental (SIMA). Su función principal es medir y reportar los niveles de contaminantes en la atmósfera para proteger el medio ambiente y la salud de la ciudadanía. SIMA contiene una red de estaciones de monitoreo, ubicadas en diferentes áreas del estado. Cada una de ellas está equipada con tecnología y sensores avanzados para detectar ciertos contaminantes que afectan la salud. Dentro de estas estaciones, se recopila la información meteorológica como la velocidad y la dirección del viento. Desde 1992, SIMA ha sido la autoridad para tomar las mejores decisiones informadas en el estado. Desempeñando un papel crucial en la mejora del medio ambiente y, por ende, en la calidad de vida de sus habitantes (*"Aire.nl.gob.mx | Home," 2018*).

3. Problemática y Objetivo

Sobre esta misión, se plantea analizar la relación que existe entre la cantidad de áreas verdes y la calidad del aire. ¿En qué medida los árboles determinan la calidad del aire?, ¿Es la falta de árboles un factor importante en la mala calidad del aire en Monterrey? ¿Cuál es el papel de las zonas verdes para la calidad del aire? Así, se concluye en la siguiente pregunta de investigación: *¿Existe una diferencia en la concentración de contaminantes en el aire entre zonas con alto y bajo nivel de áreas verdes en Monterrey durante el año 2022?*

Con esta pregunta de investigación, se persigue el objetivo de crear un reporte detallado, con fundamento estadístico, que clarifique la relación entre porcentaje de zonas verdes y calidad del aire. Con esto se espera como resultado brindar a instituciones e investigadores una herramienta para tomar decisiones, así como concientizar sobre la importancia de las zonas verdes en una comunidad industrializada. Por otro lado, también se espera explorar qué elementos podrían tomar un papel importante a la hora de establecer planes de mitigación contra la contaminación, facilitando a empresas, gobierno y ciudadanía una inversión más informada.

Para llevar a cabo esta tarea, se emplearán datos de gobierno proporcionados por SIMA. Estos datos constan de mediciones y reportes generados por 13 estaciones de monitoreo de calidad del aire distribuidas a lo largo de Monterrey y su área metropolitana. Los datos mencionados abarcan un registro histórico de varios años, datos sobre la presencia en cantidad de diferentes contaminantes y características meteorológicas durante los registros. Cabe mencionar que los datos serán procesados para el posterior análisis.

4. Preparación de la base de datos

4.1. Obtención de porcentaje de áreas verdes

Para realizar un análisis del porcentaje de vegetación o zonas verdes que se encuentra a los alrededores de cada estación de monitoreo se empleó Google Earth Engine. Desde este software, se obtuvieron imágenes de alta definición del área a la redonda de las estaciones de monitoreo, tomando en consideración un radio de 2.5 km, siendo la mitad de la distancia mínima entre pares de estaciones, y con resolución de 10 metros por pixel. Se recopiló una foto por cada mes del 2022 para cada estación, a excepción de septiembre por un error propio del software. Seguidamente, se utilizó Python y la librería *rasterio* para identificar automáticamente

el porcentaje de verde para cada imagen y así obtener un promedio de vegetación del mes. Un ejemplo de mapa obtenido y procesado se ilustra en Figura 1.

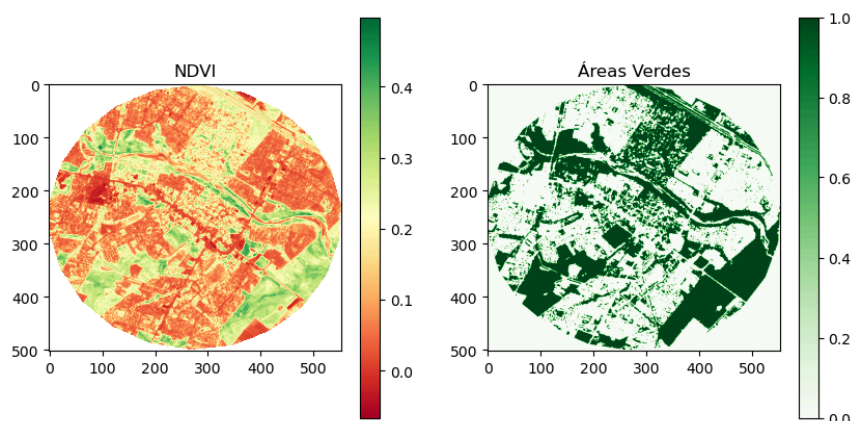


Figura 1. Áreas verdes a 2.5 km a la redonda de la estación de monitoreo Sureste 2 (Juárez).

4.2. Formato y limpieza de datos SIMA

Por otro lado, las bases de datos proporcionadas por el Sistema Integral de Medio Ambiente requerían un procesamiento previo para aplicar las herramientas de análisis y exploración. La base de datos de interés se encontraba en el archivo de Excel “DATOS HISTÓRICOS 2022_2023_TODAS ESTACIONES.xlsx”. Dicho archivo estaba separado en distintas hojas que representaban cada una de las estaciones; y en cada hoja aparecían los datos de todos los contaminantes y factores meteorológicos para todo el periodo. Por lo tanto, para conveniencia de análisis, era necesario agrupar cada una de las hojas en una sola tabla.

Al momento de importar el archivo a un Jupyter Notebook, donde se va a trabajar con la librería Pandas en Python, este aparecía como un arreglo tipo diccionario, donde aparecía el nombre y el contenido de cada una de las hojas de la base de datos en Excel. La función que agrupa cada una de las hojas asegura que en el nombre de las columnas aparezca la estación de donde se toman los datos para así diferenciar el origen.

La tabla generada tiene un tamaño de 226 columnas y 14,255 filas, que comprenden las mediciones realizadas en todas las estaciones para los datos horarios desde el 1 de enero de 2022 hasta el 17 de agosto de 2023, con un total de 164,465 datos nulos. No obstante, se realizaron varios cambios para encaminar el formato a lo necesario para investigar los objetivos. Por ejemplo, se mantuvieron aquellas variables relacionadas con contaminantes y excluimos otros factores meteorológicos. Además, los valores negativos fueron convertidos en nulos porque, para los contaminantes, se trataba de mediciones erróneas.

Asimismo, con la exploración necesaria, encontramos que los datos de las estaciones de Noreste3 y Noroeste3 contenían exceso de datos nulos que no era posible resolver desde la imputación, por lo cual fueron descartadas. Acto seguido, se trabajó con la función que interpola los datos de manera lineal, el cual se apoya de los datos no nulos más recientes para realizar la aproximación. Por último, se hizo un recorte de los registros para mantener los datos del año 2022, sin considerar el primer registro del 1 de enero por la imposibilidad de imputar con interpolación. Cabe recalcar que se agregaron columnas de apoyo para desglosar la fecha en año, mes, día y hora. La tabla resultante tiene 109 columnas y 8,736 registros.

4.3. Creación de bases con nivel de verde y concentración

Para efectuar el modelo en el que se basa esta investigación, explicado de forma detallada en la sección 6, fue necesario generar nuevas bases de datos derivadas de la tabla resultante. La idea detrás de esto es conseguir de forma separada por mes datos sobre el porcentaje de verde en las estaciones, así como estudiar de forma separada los datos capturados de cada contaminante. Por un lado, con los datos de porcentaje de verde, se generó una clasificación mensual de “ALTO” o “BAJO” para aquellas estaciones que tuvieran una mayor o menor proporción, respectivamente, considerando cuatro estaciones para cada grupo. Cabe recalcar que, para cada mes, se quedaban fuera del análisis aquellas estaciones que no entraban en ninguna de las clasificaciones. En la Tabla 1 se muestra el ejemplo de la clasificación de las estaciones según su proporción de áreas verdes para el mes de Enero.

Estaciones	Porcentaje de área verde (%)
SURESTE3 (Cadereyta)	55.80
SUR (Monterrey)	45.00
SURESTE2 (Juárez)	40.63
SUROESTE2 (San Pedro)	37.94
NOROESTE2 (García)	20.45
CENTRO (Monterrey)	19.44
NORESTE (San Nicolás)	17.49
NORTE2 (San Nicolás)	14.98

Tabla 1. Estaciones con menor y mayor porcentaje de áreas verdes en Enero.

A partir de ello, se tomaron los datos de las estaciones que entraron en alguna de estas clasificaciones y fueron separadas en los distintos contaminantes. Por lo tanto, se generó una base de datos para cada par de mes y contaminante en donde se agruparon las ocho estaciones selectas, así como incluir la fecha de cada registro y si el dato proviene de una estación clasificada como “ALTO” o “BAJO”. En la Tabla 2 se muestran algunos registros para la base de datos del mes de junio y contaminante PM10.

PM10	FECHA	PORCENTAJE
------	-------	------------

37	01/06/2022 00:00	ALTO
34	01/06/2022 01:00	ALTO
29	01/06/2022 02:00	ALTO
31	01/06/2022 03:00	ALTO
...		
...		
...		
56	30/06/2022 20:00	BAJO
46	30/06/2022 21:00	BAJO
28	30/06/2022 22:00	BAJO
41	30/06/2022 23:00	BAJO

Tabla 2. Base de datos para PM10 en el mes de junio.

En total se crearon 88 distintas bases de datos, siendo los 8 contaminantes y 11 meses del año (sin considerar septiembre debido a la falta de datos sobre las áreas verdes), cada una de entre 5,760 y 5,952 filas, dependiendo del número de días de cada mes.

5. Análisis exploratorio

Se lleva a cabo el proceso de revisión de las medidas de tendencia central y de dispersión para las variables elegidas. En este caso, los valores mostrados en Tabla 3 representan los cálculos realizados sobre el conjunto de todas las mediciones de cada contaminante sobre todas las estaciones.

	Contaminante							
	CO	NO	NO2	NOX	O3	PM10	PM2.5	SO2
Media	1.50	11.83	15.21	27.03	25.33	61.68	20.77	5.37
Mediana	1.46	4.70	11.80	17.50	21.00	52.00	17.00	4.50
Desviación estándar	0.60	23.98	11.90	31.40	18.26	44.74	16.15	4.54
Mínimo	0.00	0.30	0.00	0.50	1.00	2.00	0.00	0.50
Máximo	7.03	945.10	167.80	971.80	151.00	1001.00	370.00	190.30

Tabla 3. Medidas estadísticas de los contaminantes en todas las estaciones.

La Tabla 3, en general, muestra la existencia de cierto sesgo a la derecha para cada uno de los continentes; esto es particularmente observable cuando se analiza el mínimo y máximo para cada concentración. Se observa que el contaminante con menor concentración promedio (ppm) es la de CO con 1.50 y el de SO₂. Estos, a su vez, tienen un menor y mayor rango respectivamente; lo mismo sucede con sus desviaciones estándar.

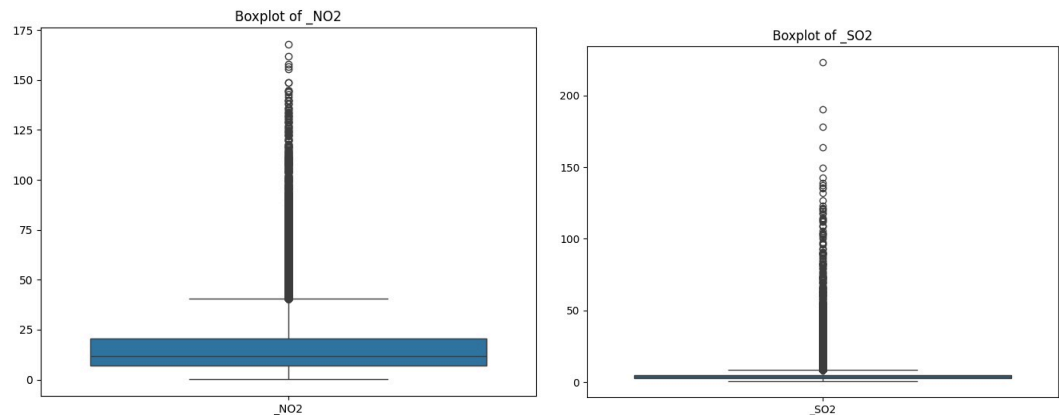


Figura 2. Boxplot de concentración de NO₂ y SO₂ en las diferentes estaciones de monitoreo de calidad del aire.

La figura 2 permite identificar que existen outliers frecuentes en cuanto a la concentración de algunos contaminantes. Estos outliers varían significativamente de las cifras habituales. Tomando en cuenta este dato, es posible estimar que existen factores que promueven la concentración de estos contaminantes de forma espontánea y con poca duración, podrían ser eventos como incendios o sesgos en la percepción de las estaciones de monitoreo.

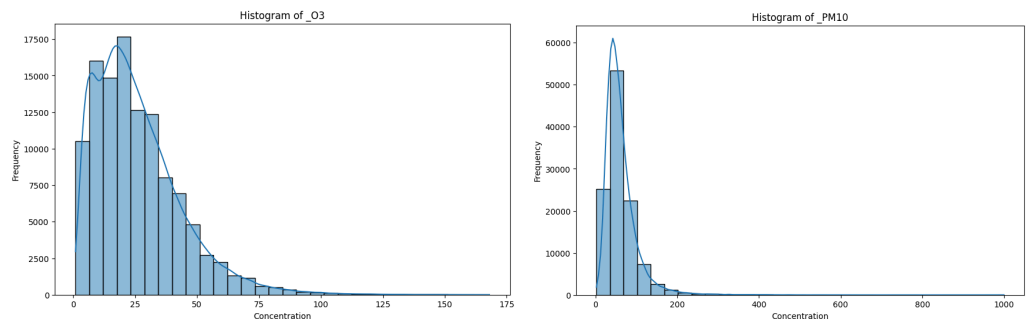


Figura 3. Concentración de O₃ Y PM10 alrededor de las diferentes estaciones de monitoreo de la calidad del aire.

De la figura 3 se observa que la concentración de algunos contaminantes en el aire siguen una distribución en su cantidad de concentración parecida a la normal. Además, se obtiene que las concentraciones estructuran un común o cifra habitual. Esto podría indicar que los factores que promueven la concentración de estos contaminantes se encuentran establecidos de forma constante durante los registros.

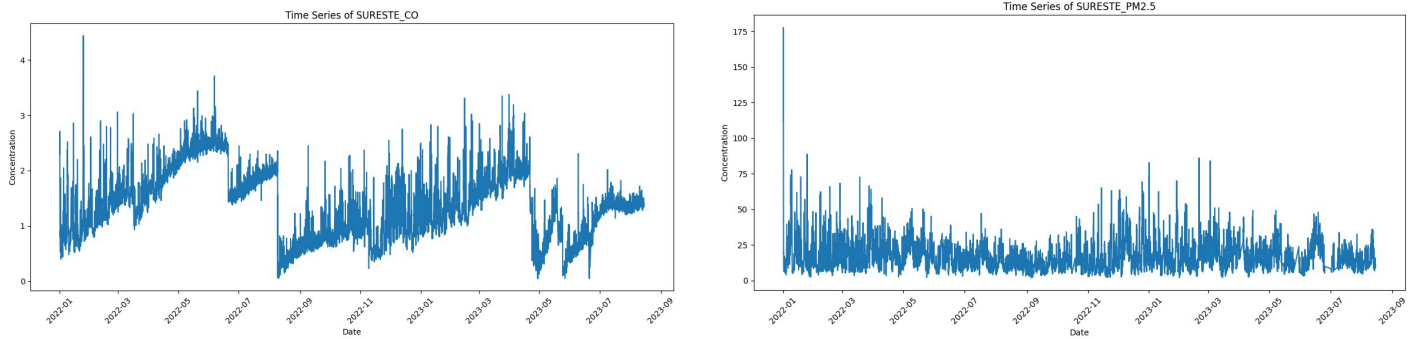


Figura 4. Serie de tiempo sobre concentración de CO y PM2.5 registrada alrededor de la estación de monitoreo sureste (Guadalupe).

Del gráfico 4 se puede desprender que alrededor de la estación de monitoreo sureste en Guadalupe, existe una concentración de PM2.5 variable y espontánea. La concentración presenta picos abruptos, además de niveles de concentración que no siguen un patrón de comportamiento o tendencia claramente definido. Podría interpretarse que la concentración de PM2.5 se comporta de forma variable según sucesos que no están estrechamente relacionados con la periodicidad o el tiempo, sino con sucesos repentinos que pueden incrementar por cortos periodos de tiempo la concentración.

6. Modelación y validación

A partir de las bases de datos derivadas para cada mes y contaminante, se realizó un modelo de efectos mixtos con el fin de explicar el comportamiento de las concentraciones para los grupos de alto y bajo porcentaje de áreas verdes considerando la autocorrelación en los datos. En notación matricial, cada modelo puede escribirse como

$$y = X\beta + Zu + \epsilon,$$

para el cual y es la variable de respuesta que indica la concentración del contaminante; X es la matriz de diseño para los efectos fijos; β es el vector de coeficientes que se desea conocer; Z es la matriz de diseño para los efectos aleatorios; u es el vector desconocido de los efectos aleatorios; y ϵ es un vector desconocido de errores aleatorios (UCLA, s. f.).

Para la modelación, se consideran efectos fijos los niveles de alto y bajo porcentaje de áreas verdes mientras que el tiempo es utilizado como un efecto aleatorio (es decir, se emplea la columna de fecha de cada base de datos) debido a las múltiples observaciones que se

tienen para la concentración del contaminante en un mismo instante. Con ello, utilizando los coeficientes estimados, es posible calcular el comportamiento esperado de la concentración de ambos grupos mediante la obtención de las medias marginales estimadas (MME) y realizar una comparación mediante pruebas t considerando las hipótesis

$$H_0 : \mu_{alto} - \mu_{bajo} = 0 ,$$

$$H_A : \mu_{alto} - \mu_{bajo} \neq 0 ,$$

para las medias de concentración de los niveles alto y bajo con un nivel de significancia de $\alpha = 0.05$. Además, utilizando el signo de las diferencias de las MME es posible inferir si la concentración es mayor o menor para las estaciones con alta presencia de vegetación frente a aquellas con bajo nivel. Sin embargo, debido a la realización de 88 comparaciones, se esperaría un mayor riesgo de cometer un error de tipo 1; con el fin de validar las inferencias frente a esto, se realiza un ajuste de Bonferroni al valor de significancia original y se utiliza en su lugar, un valor de $\alpha_{bonferroni} = \alpha/88 \approx 0.0005$ para rechazar la hipótesis nula.

Por la naturaleza de los datos y la ausencia de normalidad, se transforman los datos de concentración para cada base de datos utilizando una transformación de Box-Cox con un parámetro λ que maximiza la verosimilitud a una distribución normal. Finalmente, para la validez de los modelos, se analizan los supuestos de normalidad, media cero y homocedasticidad en los residuos mediante gráficas de residuales y una prueba de Anderson-Darling para normalidad. Estos procesos se automatizan utilizando el lenguaje R con los paquetes *lmer* para los modelos mixtos y *emmeans* para las MME.

7. Resultados

Tras obtener las conclusiones de las pruebas t , y determinar el signo de las diferencias de las MME para cada mes y contaminante se muestran los resultados por color en la figura 5. Esta imagen conforma los hallazgos principales de esta investigación.

Concentración en zonas de mayor porcentaje de vegetación

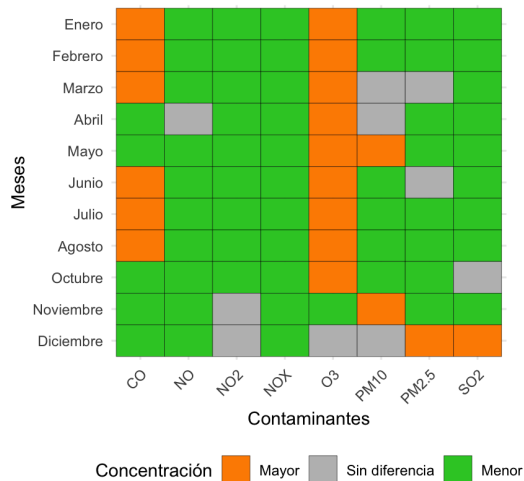


Figura 5. Relación entre concentración de diferentes contaminantes en el aire y porcentaje de vegetación en la zona mensualmente durante 2022.

La figura 5 está caracterizada por una presencia mayoritaria del color verde (59 comparaciones), lo que indica que sí existe una diferencia entre los niveles de alto y bajo y verde ($\alpha = 0.05$) y que, en concreto, la concentración de contaminantes suele ser menor en zonas alrededor de áreas con más vegetación. En concreto, se desprende que, para los contaminantes NO, NO₂ y NO_x, el alto porcentaje de áreas verdes se correlaciona con una menor concentración de los contaminantes en el aire de forma consistente durante el año.

Por otra parte, se observan líneas naranjas persistentes (19 comparaciones) a lo largo de los meses para los contaminantes CO y, especialmente, para O₃. Esto indica que, para las zonas con mayor porcentaje de áreas verdes, se suele registrar mayor concentración de ozono y monóxido de carbono. Finalmente, para diez comparaciones no se registra alguna diferencia entre los niveles de vegetación; su aparición es relativamente esporádica, es decir, no se observa algún patrón evidente, aunque se registran ligeramente más apariciones para el material particulado PM2.5 y PM10.

Respecto a los supuestos estadísticos, se generaron las 88 gráficas correspondientes a los residuales de los modelos contra los valores predichos. En particular, es notorio el cambio de la distribución de los puntos tras aplicar las transformaciones Box-Cox de normalidad a las concentraciones de los contaminantes. Un ejemplo de este cambio se ilustra en la figura 6. Los residuales de la transformación siguen una distribución mucho más uniforme, centrada en cero; sin embargo, es apreciable cierta disminución de concentración de puntos para valores en el contorno de la nube, lo que sugiere cierto grado de heterocedasticidad; además, se identifican ciertos artefactos que muestran una correlación lineal entre estos puntos. Finalmente, para la normalidad, se obtiene un valor $p < 0.05$ para la prueba de Anderson-Darling, con lo se infiere que los residuales no siguen una distribución normal.

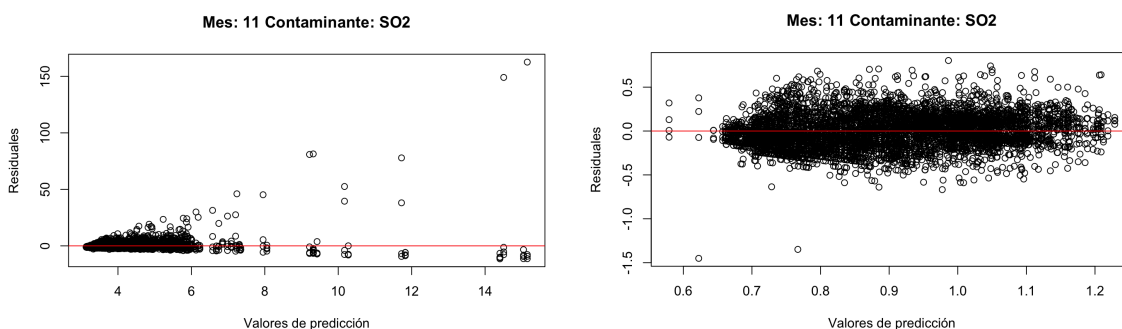


Figura 6. Ejemplo de de residuales antes (izquierda) y después de transformación Box-Cox

Este mismo comportamiento descrito anteriormente es observable en la gran mayoría de los modelos, a excepción de unos cuantos, principalmente aquellos de concentración de CO. En la figura 7, por ejemplo, es apreciable el agrupamiento de los residuales sobre dos

valores de predicción centrados aproximadamente en 0.3 y 0.6. Esta visualización de grupos se presenta casi de forma consistente a lo largo de los meses para este contaminante en particular.

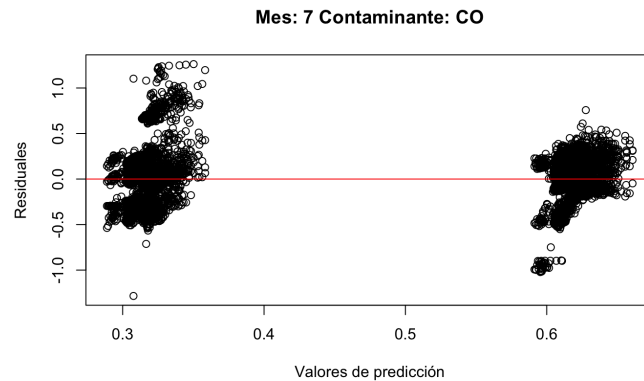


Figura 7. Gráfica de residuos agrupados para CO

Discusión y conclusiones

El tema de la calidad del aire es recurrente en todo el mundo, y más cuando se trata de una ciudad como Monterrey, cuya actividad principal es referente al área industrial. Sin embargo, entre toda esta problemática también están incluidos los temas de la vialidad y, por supuesto, la presencia de áreas verdes. De esta forma, la investigación tuvo un mayor interés en el enfoque de las áreas verdes y la disyuntiva en los efectos que tienen en los contaminantes que abundan en el área metropolitana de la ciudad. Aunado a lo anterior, el manejo de los datos y las variables a considerar se centraron en revisar de forma contrastante las estaciones que tuvieran una mayor y menor proporción de áreas verdes a los alrededores, para así tener resultados más claros en cuanto a la concentración de los contaminantes a lo largo del año 2022.

Los resultados encontrados se inclinan a una respuesta afirmativa de la pregunta de investigación que relaciona las áreas verdes y los contaminantes mediante el efecto que tienen en ellos. Para la mayoría de los contaminantes, se encontró de forma consistente un efecto favorable gracias a la presencia de áreas verdes, donde la mayoría de los meses estudiados reducía la concentración de las variables. Sin embargo, el CO y O₃ representan casos contrarios a lo esperado, donde de forma mezclada o constante existe una mayor concentración de los compuestos a mayor presencia de vegetación. Debido a estas observaciones, nuevas líneas de investigación podrían partir con el objetivo de explicar estos fenómenos como la conexión de compuestos orgánicos volátiles por parte de las plantas y la presencia de ozono en zonas urbanas o el efecto de los incendios forestales para el caso del monóxido de carbono; también se podría buscar explicar la separación de los residuales para CO.

No obstante, a pesar de la consistencia de las conclusiones a lo largo de los meses, las cuestiones referentes con los residuos del modelo y la presencia de correlaciones imprevistas no permiten alcanzar la contundencia y validez de los hallazgos, lo cual abre camino a otras nuevas oportunidades para estudiar la implementación de modelos más complejos, así como buscar la inclusión de variables consideradas para este estudio, como factores meteorológicos, de vialidad o actividad industrial cercana a las estaciones. Asimismo, un obstáculo importante fue la mala calidad de la base de datos, cuya presencia de datos nulos obligaba a aplicar métodos para interpolar espacios importantes en las variables. Para futuras implementaciones, sería adecuado contar con datos menos escasos o métodos más complejos de imputación de mayor complejidad, como las predicciones de Holt Winters de triple suavizado exponencial.

Referencias

- Air Quality and Monitoring*. (2024). Texas Commission on Environmental Quality.
<https://www.tceq.texas.gov/airquality/monops>
- aire.nl.gob.mx | Home*. (2018). Gobierno de Nuevo León.
<http://aire.nl.gob.mx/index.html>
- Calidad del Aire Ambiente. (2018). Paho.org
<https://www.paho.org/es/temas/calidad-aire/calidad-aire-ambiente>
- Contaminación 2. (2022, August 11). National Geographic
<https://www.nationalgeographicla.com/medio-ambiente/2022/08/cuales-son-los-principales-tipos-de-contaminacion-ambiental>
- Jorge, K., Delgado, Centro, N., & De Metrología. (n.d.). La calidad del aire y su relación con la Metrología.
<https://www.cenam.mx/pabiertas/arch/ponencias/Calidad%20del%20aire%20y%20Metrologia-%20Jorge%20Koelliker.pdf>
- Mayora, F. (2019). Contaminación del aire en Monterrey, Nuevo Leon: interpretación del monitoreo ambiental 2005-2018. *Revista de Investigación*, 43(98).
<https://www.redalyc.org/journal/3761/376168604013/html/#:~:text=En%20el%20%C3%81rea%20Metropolitana%20de,%25%20y%20NOx%2C%2034%25>.

Primagas. (2020). *Calidad del aire: la importancia de preservarla*.

<https://www.primagas.es/blog/calidad-del-aire>

Reviewing National Ambient Air Quality Standards (NAAQS): Scientific and Technical Information (2016, September 9). US EPA. <https://www.epa.gov/naaqs>

UNEP (2022). *¿Cómo se mide la calidad del aire?* UNEP

<https://www.unep.org/es/noticias-y-reportajes/reportajes/como-se-mide-la-calidad-del-aire>

UCLA (s. f.). *Introduction to linear mixed models*.

<https://stats.oarc.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/>

WHO Int (2021). *Types of pollutants*. .

<https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants>

WHO Int (2021). *Preguntas y respuestas*.

<https://www.who.int/es/news-room/questions-and-answers/item/who-global-air-quality-guidelines>

World. (2022). *Contaminación del aire ambiente (exterior)*. Who Int.

[https://www.who.int/es/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)

Anexo

Significado de las mediciones

Contaminantes:

PM10 (Material Particulado menor a 10 micrómetros), su unidad es de $\mu\text{g}/\text{m}^3$. Se refiere a partículas en el aire con un diámetro de 10 micrómetros o menos. Las cuales causan problemas respiratorios y otros efectos en la salud.

PM2.5 (Material Particulado menor a 2.5 micrómetros), su unidad es de $\mu\text{g}/\text{m}^3$. Se refiere a partículas finas que pueden llegar profundamente a los pulmones y torrentes sanguíneos, causando problemas de salud.

O3(Ozono), su unidad es ppb (partes por billón). El ozono troposférico es un contaminante que causa reacciones químicas con otros contaminantes en la luz solar. Es un gran irritante pulmonar.

SO2 (Dióxido de Azufre), su unidad ppb. Dióxido de azufre de un gas producido por la quema de combustibles fósiles que contienen azufre.

NO2 (Dióxido de Nitrógeno), su unidad ppb. Es un contaminante derivado de la combustión de vehículos y plantas industriales. Contribuye a la formación de ozono.

CO (Monóxido de Carbono), su unidad ppm. Es un gas tóxico para los humanos en concentraciones altas.

NO (Monóxido de Nitrógeno), su unidad ppb. Gas que contribuye a la contaminación del aire.