

# **Análisis de la Dinámica Temporal y Relación entre las Concentraciones de PM<sub>2.5</sub> y Ozono en Itagüí, Antioquia, mediante Series de Tiempo**

Juan Esteban Bolívar Torres, estudiante de Ingeniería Ambiental, 1003383763

## **Resumen (Abstract)**

La contaminación del aire por material particulado fino (PM<sub>2.5</sub>) y ozono (O<sub>3</sub>) es un desafío significativo en zonas urbano-industriales como el municipio de Itagüí, Colombia. Este estudio tiene como objetivo principal caracterizar y modelar la dinámica de las concentraciones de estos dos contaminantes para identificar sus patrones temporales y evaluar la relación que existe entre ellos. Se utilizaron datos de alta frecuencia temporal, abarcando desde abril de 2023 hasta agosto de 2024, provenientes de la estación de monitoreo del SIATA ubicada en la I.E. Concejo Municipal de Itagüí. La metodología empleó un análisis de series de tiempo que incluyó estadísticas descriptivas, análisis de autocorrelación (ACF/PACF), transformada de Fourier, el ajuste de un modelo ARMA(1,1) y análisis de correlación y regresión. Los resultados revelaron ciclos diarios distintos para cada contaminante: el PM<sub>2.5</sub> presenta dos picos diarios asociados a las horas de mayor tráfico, mientras que el ozono muestra un único pico al mediodía, impulsado por la actividad fotoquímica. El modelo ARMA(1,1) ajustado demostró ser insuficiente para capturar la fuerte componente estacional de las series, lo que se evidenció en la autocorrelación remanente en sus residuales. Adicionalmente, se encontró una correlación lineal negativa, estadísticamente significativa pero extremadamente débil (coeficiente de -0.0656), entre las dos variables. Los modelos de regresión confirmaron esta débil conexión, con un coeficiente de determinación (R<sup>2</sup>) inferior a 0.01, demostrando una capacidad predictiva prácticamente nula. Se concluye que, si bien los patrones temporales de ambos contaminantes son muy marcados, su interrelación es compleja y no puede ser explicada adecuadamente mediante modelos lineales simples o modelos ARMA estándar.

---

## **Introducción:**

La contaminación del aire representa uno de los desafíos ambientales y de salud pública más apremiantes a nivel global, con un impacto particularmente significativo en los centros urbanos densamente poblados e industrializados. El Valle de Aburrá, y en particular el municipio de Itagüí, no es ajeno a esta problemática. La combinación de una alta densidad poblacional, una considerable actividad industrial y las condiciones geográficas del valle, que pueden dificultar la dispersión de contaminantes, genera una compleja dinámica en la calidad del aire. (Pérez-Carrasquilla, 2023)

Entre la variedad de contaminantes atmosféricos, el material particulado fino (PM<sub>2.5</sub>) y el ozono troposférico (O<sub>3</sub>) son dos de los que generan mayor preocupación debido a sus demostrados efectos adversos en la salud humana y el medio ambiente. El PM<sub>2.5</sub>,

compuesto por partículas de un diámetro aerodinámico inferior a 2.5 micrómetros, puede penetrar profundamente en el sistema respiratorio y circulatorio. Por su parte, el ozono, un contaminante secundario formado por reacciones fotoquímicas, es un potente irritante del sistema respiratorio. (Pérez-Carrasquilla, 2023; Wang et al., 2025)

Comprender el comportamiento de estos contaminantes es fundamental para el diseño de estrategias de mitigación efectivas. En este contexto, el presente estudio se enfoca en el análisis de las concentraciones de PM<sub>2.5</sub> y ozono en el municipio de Itagüí, utilizando datos de alta frecuencia temporal. El objetivo principal es caracterizar y modelar la dinámica de estos contaminantes para identificar sus patrones temporales, tendencias y la posible relación entre ellos.

Para alcanzar este objetivo, se empleará una metodología robusta de análisis de series de tiempo. Se realizarán análisis descriptivos para obtener una comprensión inicial del comportamiento de las variables. Posteriormente, se aplicarán técnicas avanzadas como el análisis de autocorrelación, la transformada de Fourier y el análisis de ondículas para descomponer y examinar la estructura temporal de las series, incluyendo sus ciclos diarios y estacionalidad. Finalmente, se ajustarán modelos autorregresivos y de media móvil (ARMA) para intentar capturar la dependencia temporal de los datos y se evaluará la relación lineal y no lineal entre el PM<sub>2.5</sub> y el ozono mediante análisis de correlación y regresión. Los hallazgos de este estudio buscan aportar conocimiento valioso para la gestión de la calidad del aire en una de las zonas más dinámicas del Valle de Aburrá.

## **Metodología:**

Para este estudio, se utilizaron datos obtenidos del Sistema de Alerta Temprana de Medellín (SIATA). Los datos provienen de una estación de monitoreo de calidad del aire ubicada en la Institución Educativa Concejo Municipal de Itagüí. El rango de tiempo de los datos abarca desde abril de 2023 hasta abril de 2025. Estos datos fueron recopilados en 26 archivos CSV, que contenían información sobre PM<sub>2.5</sub> y Ozono, las variables de interés para el estudio. Posteriormente, se realizó un proceso de consolidación para integrar estos datos en un único archivo CSV y asegurar la correcta secuencia temporal de las fechas.

Para cada parámetro medido, los datos incluyen una columna de “calidad” que indica su validez: el valor “1” representa un dato válido, mientras que números iguales o mayores a “2” indican datos de menor calidad. Por lo tanto, se filtraron estos datos para asegurar la alta calidad de la información de cada variable. Se identificó una interrupción de un mes en la recopilación de datos de ozono posterior al 15 de agosto de 2024 a las 10:00:00. Debido a esta brecha de datos, se decidió limitar el análisis a los datos anteriores a dicha fecha. Para abordar fechas con datos faltantes, se aplicó un método de interpolación temporal (por tiempo) a las series de PM<sub>2.5</sub> y Ozono. Finalmente, para mitigar el impacto de valores atípicos (outliers), se eliminaron los datos que excedían tres desviaciones estándar de la media. Aunque esta medida redujo la cantidad total de datos, disminuyó la variabilidad causada por dichos valores atípicos.

Con la intención de hacer un análisis general de ambas variables, se calcularon las principales estadísticas descriptivas. Esto permitió obtener una comprensión de la distribución de las concentraciones de PM2.5 y Ozono. Además, se determinaron los siguientes parámetros estadísticos: el número de observaciones, los valores mínimo y máximo, la media, la mediana, la varianza, la desviación estándar, la asimetría, la curtosis y el coeficiente de variación (CV), con el fin de facilitar su interpretación.

Se crearon visualizaciones iniciales para explorar la naturaleza de los datos:

**Gráficas de series de tiempo:** Se desarrollaron gráficas horarias para las concentraciones de PM2.5 y ozono para tener una primera idea sobre las tendencias, posibles patrones estacionales y los eventos puntuales de alta concentración.

**Histogramas:** Se crearon histogramas para observar la distribución de frecuencias de las concentraciones de ambos contaminantes y analizar visualmente su forma (simetría, curtosis) y su rango.

**Diagramas de Caja y Bigotes:** Se emplearon diagramas de caja y bigotes para obtener una representación visual de estadísticas clave como los cuartiles, la mediana y el rango. Esto permitió confirmar la coherencia con los resultados obtenidos en las estadísticas descriptivas.

Para analizar la estructura temporal de las concentraciones de cada variable, inicialmente se intentó descomponer las series en sus componentes de tendencia, estacionalidad y residuales. Para ello, se utilizó el método `seasonal_decompose` con un período de 24 horas, buscando identificar un patrón diario. Sin embargo, dado que la descomposición del Ozono mostró un patrón estacional diario inconsistente con conocimiento teórico, mejor se optó por calcular y visualizar el ciclo diario promedio directamente, promediando las concentraciones para cada hora del día a lo largo de todo el período de estudio.

Adicionalmente, se calcularon y graficaron las Funciones de Autocorrelación (ACF) y de Autocorrelación Parcial (PACF) para ambas series. Estos gráficos se usaron para analizar una posible autocorrelación y estacionalidad en cada una de las variables.

La Prueba de Mann-Kendall se empleó para evaluar estadísticamente la existencia de tendencias significativas en las series de PM2.5 y Ozono. Dado que las variables ambientales presentan una fuerte dependencia de múltiples factores, se aplicó también la Prueba de Mann-Kendall con pre-whitening. Además, para comprobar la estacionariedad de las series, se realizó la Prueba de Dickey-Fuller Aumentada (ADF).

Para identificar los ciclos importantes en las series, se aplicó la Transformada Rápida de Fourier (FFT) a los datos. Esto permitió obtener periodogramas que facilitaron la determinación de las frecuencias con mayores valores de energía. Adicionalmente, se aplicó la Transformada Continua de Ondícula (CWT), utilizando la ondita 'Morlet'. Los

escalogramas resultantes fueron utilizados para analizar la potencia de diferentes periodos a lo largo del tiempo, y así evidenciar posibles ciclos.

Se ajustaron modelos de series de tiempo con el fin de capturar la estructura de dependencia temporal de los datos. Para asegurar una estacionariedad más robusta, las series se prepararon mediante el cálculo de anomalías (estandarización Z-score) y la aplicación de una diferenciación (utilizando `np.diff`) a estas anomalías. Este proceso permitió obtener series con una media y varianza más estables, y sin tendencia lineal.

También, se ajustaron modelos ARMA(1,1) a las series de anomalías (sin tendencia) de PM2.5 y Ozono. Para ello, se utilizó la función ARIMA de `statsmodels.tsa.arima.model` con un orden (1, 0, 1). La elección inicial de este orden se fundamentó en la observación de los gráficos ACF y PACF de las series diferenciadas.

Para verificar el ajuste del modelo, se llevaron a cabo los siguientes procesos:

**Comparación visual:** Se realizó el trazado de las series de valores reales y los valores predichos por el modelo. Esto permitió una evaluación visual de la calidad del ajuste.

**Estadísticas del Modelo:** Análisis del resumen del modelo (`model_fit.summary()`), incluyendo coeficientes, p-valores, AIC, BIC y pruebas de diagnóstico como Ljung-Box y Jarque-Bera.

**Métricas de Error:** Cálculo del Error Cuadrático Medio (MSE) y la Raíz del Error Cuadrático Medio (RMSE) a los valores del modelo (`model_fit.fittedvalues`).

**Análisis de Residuales:** Se crearon gráficos de diagnóstico que incluyen: histogramas, QQ-plots, gráficos de residuales vs. valores ajustados y ACF de residuales. Estos se utilizaron para diagnosticar y evaluar la independencia, la constancia de la varianza y la normalidad de la distribución de los residuos (`model_fit.plot_diagnostics`). Adicionalmente, la Prueba de Ljung-Box, incluida en el resumen del modelo, fue empleada para evaluar la autocorrelación en los residuales.

**Prueba de Kolmogorov-Smirnov:** Se utilizó la prueba KS para comparar formalmente la distribución de la serie real (anomalías diferenciadas) con la distribución de la serie sintética generada por el modelo ARMA(1,1) ajustado.

Para evaluar la relación lineal y no lineal entre las concentraciones de PM2.5 y Ozono (utilizando las series de anomalías diferenciadas), se calculó el Coeficiente de Correlación de Pearson. Se analizó su valor p para determinar si la correlación era estadísticamente significativa y cuál era la dirección de la relación entre ambas variables. Adicionalmente, se generó un diagrama de dispersión para visualizar la relación entre las anomalías diferenciadas de PM2.5 y Ozono.

Se ajustó un modelo de regresión lineal simple para modelar la concentración de Ozono en función de PM2.5. Este proceso incluyó la evaluación de modelos con y sin intercepto,

utilizando la función OLS de statsmodels.api. Se examinó el resumen de cada modelo para evaluar la significancia de los coeficientes, el R<sup>2</sup> (coeficiente de determinación) y las pruebas de diagnóstico. Adicionalmente, se ajustó un modelo de regresión polinomial de orden 2 (cuadrático) para explorar posibles relaciones no lineales entre las variables, y se calculó el R<sup>2</sup> correspondiente a este ajuste.

De manera similar a la evaluación del modelo ARMA, se analizaron los residuales de los modelos de regresión lineal y cuadrática. Para ello, se utilizaron gráficos de dispersión (residuales vs. valores ajustados), histogramas, QQ-plots y autocorrelogramas. Este análisis permitió evaluar los supuestos de homocedasticidad, distribución normal e independencia de los errores.

Todo el análisis de datos y la generación de visualizaciones se realizaron en el entorno de Google Colab, utilizando el lenguaje de programación Python. Las principales librerías empleadas fueron:

- pandas: Para manejo y manipulación de datos, especialmente series de tiempo.
- numpy: Para operaciones numéricas y matriciales.
- matplotlib.pyplot y seaborn: Para la visualización de datos y resultados.
- scipy.stats: Para funciones estadísticas y pruebas (describe, f.ppf, f.cdf, kstest, pearsonr, levene, ztest).
- statsmodels.tsa: Para análisis y modelado de series de tiempo (seasonal\_decompose, acf, pacf, plot\_acf, plot\_pacf, adfuller, ARIMA).
- statsmodels.api: Para modelos estadísticos, incluyendo regresión OLS.
- pymannkendall: Para la prueba de Mann-Kendall y sus variaciones.
- pywt: Para el análisis de ondículas (wavelets).
- sklearn.metrics: Para métricas de evaluación de modelos como el R<sup>2</sup> y MSE/RMSE.
- matplotlib.dates: Para formatear fechas en gráficos de series de tiempo.

## Resultados:

Inicialmente, se calcularon las principales estadísticas descriptivas con el fin de obtener una comprensión preliminar del comportamiento de las concentraciones.

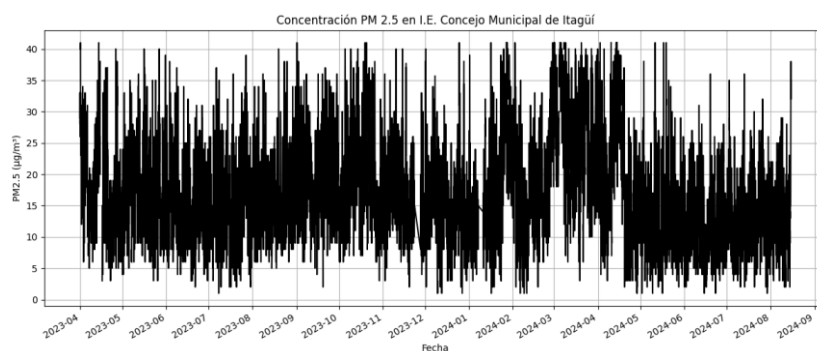
```
Resumen estadístico de PM 2.5:
Los cuartiles son:
Q1: 11.0 [µg/m³] en la posición: 2901.0
Q2: 15.0 [µg/m³] en la posición: 5802.0
Q3: 21.0 [µg/m³] en la posición: 8703.0
-----
Las medidas estadísticas son:
Número de observaciones: 11604
Mínimo: 1.0 [µg/m³]
Máximo: 41.0 [µg/m³]
Media: 16.32829412968857 [µg/m³]
Mediana: 15.0 [µg/m³]
Varianza: 59.71556518708096 [µg/m³]^2
Desviación estándar: 7.72758469297367 [µg/m³]
Asimetría: 0.6968965798109783
Curtosis: 0.197048083576969
Coeficiente de Variación (CV): 47.3263442684019
```

```
Resumen estadístico de Ozono:
Los cuartiles son:
Q1: 5.7 [µg/m³] en la posición: 2926.75
Q2: 14.4 [µg/m³] en la posición: 5853.5
Q3: 29.8 [µg/m³] en la posición: 8780.25
-----
Las medidas estadísticas son:
Número de observaciones: 11707
Mínimo: 0.1 [µg/m³]
Máximo: 72.6 [µg/m³]
Media: 19.58470888093648 [µg/m³]
Mediana: 14.4 [µg/m³]
Varianza: 274.7567217465037 [µg/m³]^2
Desviación estándar: 16.575787213478087 [µg/m³]
Asimetría: 1.0032677167511117
Curtosis: 0.20198265511362967
Coeficiente de Variación (CV): 84.63637276535042
```

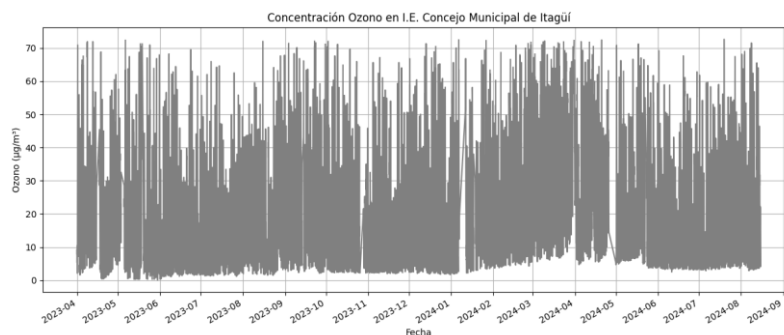
**Imagen 1 :** Estadísticas descriptivas calculadas para ambas series de tiempo (PM2.5 y Ozono) Fuente: *Propia*

Basado en la imagen 1, las concentraciones de PM2.5 y Ozono en la estación de Itagüí presentan patrones distintos; mientras que ambos contaminantes muestran una distribución con asimetría positiva, mostrando la presencia de picos cada cierto tiempo, esta asimetría es notablemente más notoria en el Ozono, reflejada en una diferencia mayor entre su media y mediana. Además, el Ozono exhibe una variabilidad superior a la del PM2.5, que se evidencia por su coeficiente de variación mucho más alto y un rango intercuantil más amplio, lo que resalta su naturaleza altamente fluctuante debido a su ciclo diario.

Se generaron gráficos de las series de tiempo, histogramas y diagramas de caja y bigotes para facilitar la interpretación visual de cada variable.

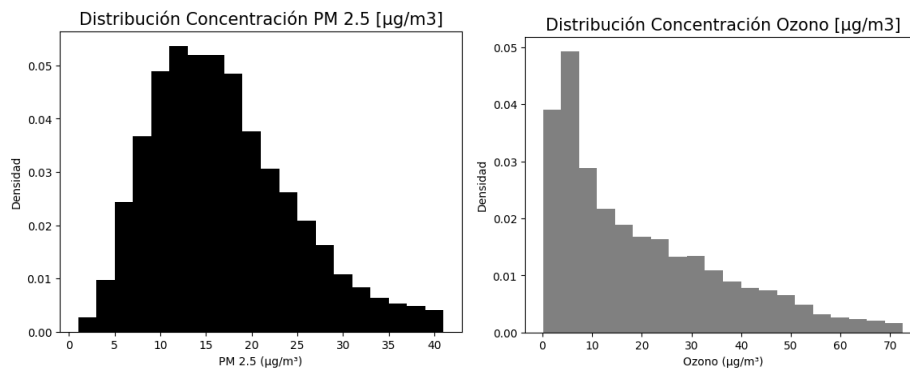


**Gráfica 1 :** Serie de tiempo PM 2.5. Fuente: *Propia*



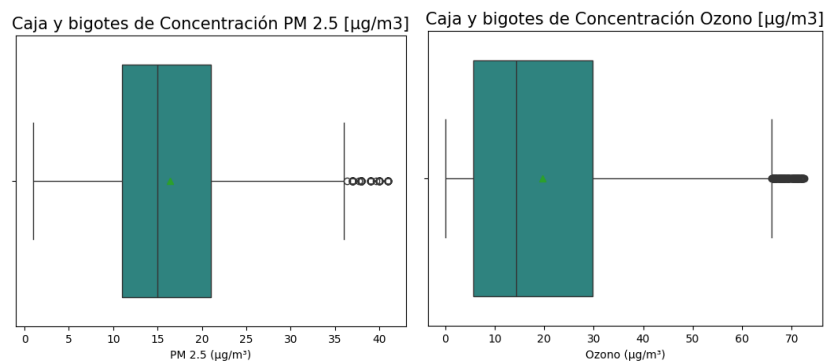
**Gráfica 2 :** Serie de tiempo Ozono. Fuente: *Propia*

Las Gráficas 1 y 2 ilustran las fluctuaciones significativas presentes en las concentraciones a lo largo del tiempo, evidenciando una marcada variabilidad en ambas series. Se observa que la serie de Ozono tiende a ser más estable que la de PM2.5; sin embargo, el Ozono presenta picos de concentración considerablemente más altos. Estas fluctuaciones temporales sugieren la posible existencia de patrones estacionales y una tendencia subyacente en los datos.



**Gráfica 3 :** Histogramas para PM 2.5 y Ozono. Fuente: *Propia*

La Gráfica 3 ilustra la asimetría positiva previamente mencionada, lo cual es evidente a simple vista y sugiere que la distribución de los datos no se ajusta a una distribución normal. En este contexto, una distribución log-normal podría ofrecer un mejor ajuste. Ambos histogramas evidencian una mayor concentración de valores en los rangos bajos, y en el caso del Ozono, se observa la presencia de fechas con concentraciones excepcionalmente altas.

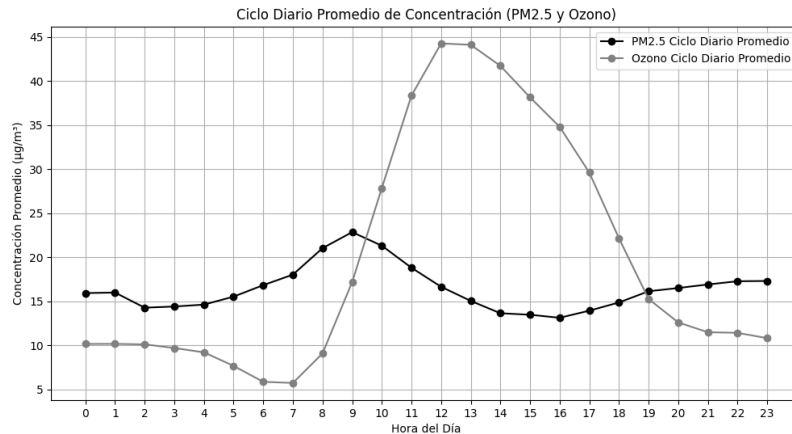


**Gráfica 4 :** Diagramas caja y bigotes para PM 2.5 y Ozono. Fuente: *Propia*

La Gráfica 4 permite confirmar visualmente que la mayoría de los valores se concentran en rangos bajos. Adicionalmente, se observa que la mediana se sitúa a la izquierda de la media, lo cual reafirma la asimetría positiva de la distribución de los datos. También se evidencia que la caja del diagrama correspondiente al Ozono presenta un rango intercuartil mayor, lo que indica una mayor dispersión en sus concentraciones. A pesar de la aplicación de un filtro inicial de outliers, la gráfica sugiere que aún persisten algunos valores atípicos en cada variable.

Luego, se llevaron a cabo los análisis de series de tiempo pertinentes para determinar su estructura temporal.



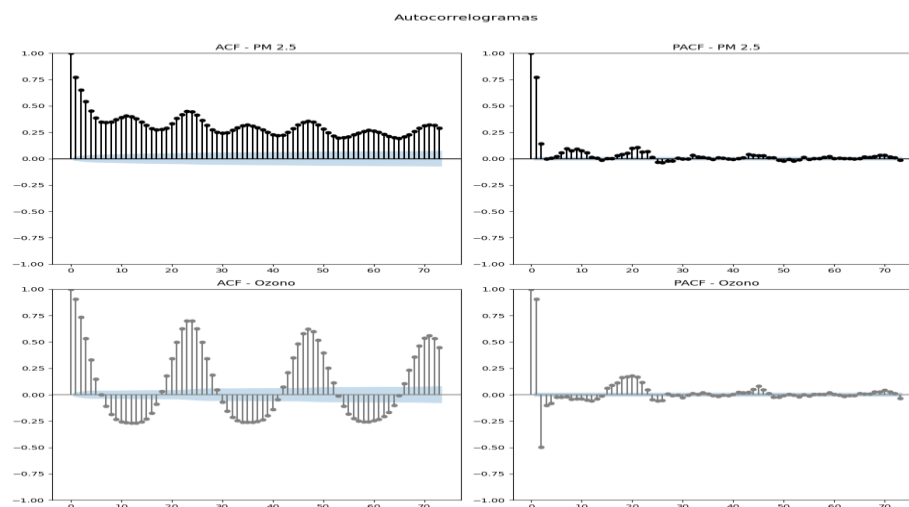


**Gráfica 5:** Ciclo diario promedio para PM 2.5 y Ozono. Fuente: *Propia*

Para analizar la estacionalidad de 24 horas, inicialmente se intentó utilizar la función `decomp.seasonal`. Sin embargo, dado que los resultados obtenidos difirieron de lo esperado según el conocimiento teórico previo, se optó por calcular el promedio de las concentraciones para cada hora del día para PM2.5 y Ozono.

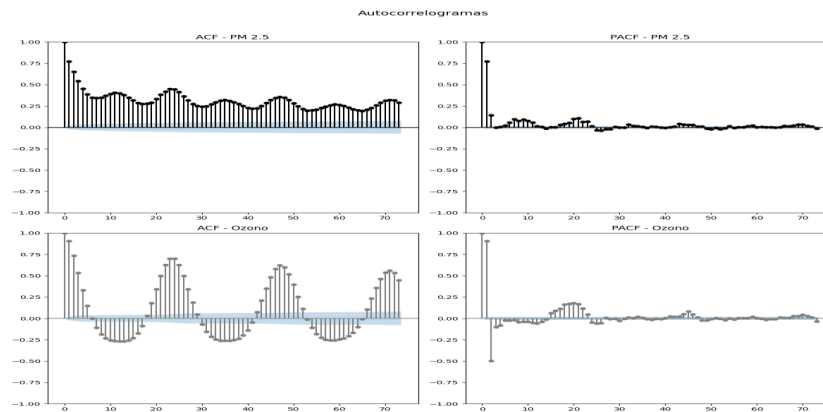
El PM2.5 exhibe dos picos diarios: uno en la mañana (entre las 7:00 y las 9:00 horas) y otro en la noche (entre las 18:00 y las 22:00 horas). Estos picos podrían atribuirse a la confluencia de las horas pico de tráfico y a condiciones meteorológicas específicas que favorecen la acumulación de contaminantes.

En contraste, el Ozono muestra su pico principal alrededor de las 12:00 del mediodía. Este comportamiento sugiere una relación directa con el aumento de las actividades fotoquímicas que generan Ozono, impulsadas por la radiación solar.



**Gráfica 5 :** Análisis de Autocorrelación usando ACF y PACF para PM 2.5. Fuente: *Propia*





**Gráfica 6 :** Análisis de Autocorrelación usando ACF y PACF para Ozono. Fuente: *Propia*

Las Gráficas 5 y 6 demuestran claramente una estacionalidad marcada en ambas variables, junto con una alta persistencia en las series. Para el PM2.5, se observan picos pronunciados aproximadamente cada 12 horas, mientras que en el caso del Ozono, los picos se presentan aproximadamente cada 24 horas. Estos patrones estacionales y la persistencia son indicativos de una falta de estacionariedad en las series de tiempo.

### PM 2.5

### Ozono

Parámetro	Valor	Parámetro	Valor
Tendencia	Decreciente	Tendencia	Creciente
Presencia de Tendencia (h)	Verdadera (True)	Presencia de Tendencia (h)	Verdadera (True)
Valor p	0.0	Valor p	0.0
Estadístico Z	-12.50	Estadístico Z	10.86

**Tabla 1:** Análisis de Tendencias (Mann-Kendall) para PM 2.5 Y Ozono. Fuente: *Propia*

### PM 2.5

### Ozono

Parámetro	Valor	Parámetro	Valor
Tendencia	Decreciente	Tendencia	Sin tendencia (no trend)
Presencia de Tendencia (h)	Verdadera (True)	Presencia de Tendencia (h)	Falsa (False)
Valor p	6.98e-06	Valor p	135
Estadístico Z	-4.49	Estadístico Z	1.49

**Tabla 2:** Análisis de Tendencias (Mann-Kendall con Pre-whitening) para PM 2.5 Y Ozono. Fuente: *Propia*

Como se muestra en la Tabla 1, los resultados de la prueba de Mann-Kendall inicial indicaron que las hipótesis nulas fueron rechazadas para ambas variables. Esto confirmó una tendencia decreciente para el PM2.5 y una tendencia creciente para el Ozono.

Dado el enfoque de este estudio en variables ambientales y la dependencia de estas con múltiples factores, en la Tabla 2 se muestra una prueba de Mann-Kendall con pre-whitening. Este método busca eliminar la influencia de la autocorrelación en la serie, lo

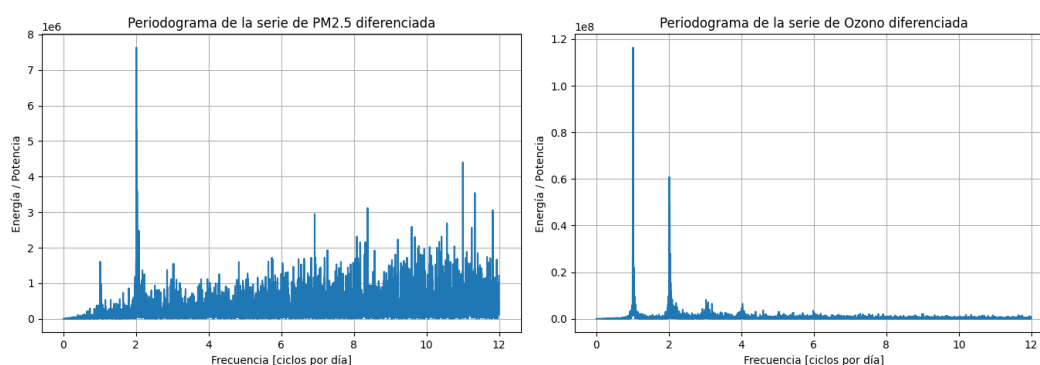
que ayuda a evitar falsos positivos en la detección de tendencias. Tras esta aplicación, la tendencia decreciente del PM2.5 se mantuvo estadísticamente significativa, con la hipótesis nula aun siendo rechazada. Sin embargo, para el Ozono, la hipótesis nula de no tendencia se volvió plausible, lo que sugiere que la tendencia creciente observada inicialmente fue, en efecto, un falso positivo inducido por la autocorrelación de la serie.

PM 2.5		Ozono	
Parámetro	Valor	Parámetro	Valor
Estadístico ADF	-8.669	Estadístico ADF	-10.838
Valor p	4.60e-14	Valor p	1.64e-19

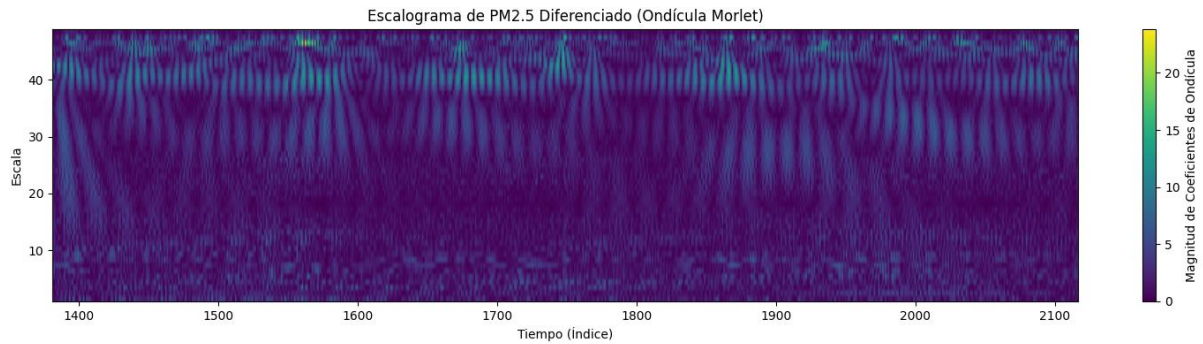
**Tabla 3:** Análisis de Estacionariedad (ADF) para PM 2.5 Y Ozono. Fuente: *Propia*

La Tabla 3 presenta los resultados de la Prueba de Dickey-Fuller Aumentada (ADF), donde en ambos casos (PM2.5 y Ozono) se rechaza la hipótesis nula. Esto evidencia la ausencia de una raíz unitaria, lo que implica que las series se consideran estacionarias.

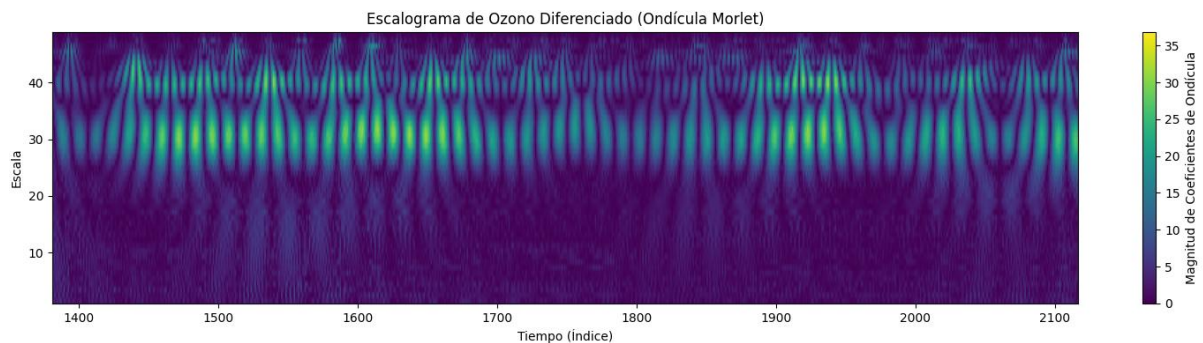
Sin embargo, este resultado podría parecer contradictorio si se compara con la observación visual de los gráficos ACF y PACF, los cuales sugieren la presencia de estacionalidad y persistencia. Aunque la prueba ADF indica estacionariedad, es una buena práctica realizar procesos adicionales para asegurar una estacionariedad más robusta, como la diferenciación y/o el cálculo de anomalías, especialmente si se planea ajustar modelos ARMA.



**Gráfica 7:** Periodograma para ver los niveles de energía por frecuencia para PM 2.5 Y Ozono. Fuente: *Propia*



Gráfica 8: Análisis de Ondículas para PM 2.5 . Fuente: *Propia*



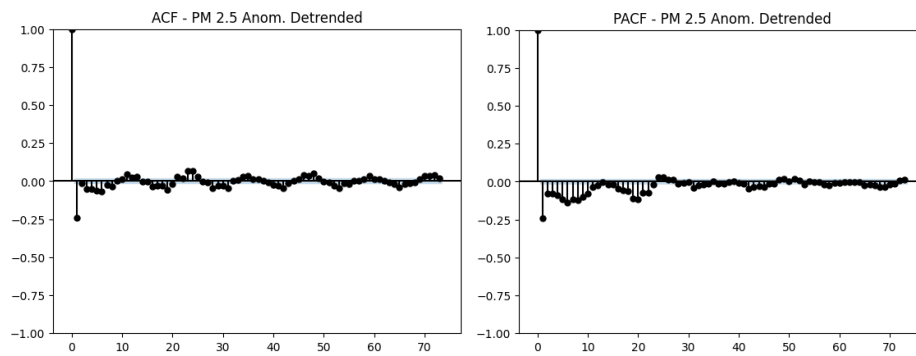
Gráfica 9: Análisis de Ondículas para Ozono. Fuente: *Propia*

La Gráfica 7 ilustra cómo el PM2.5 presenta una mayor concentración de energía en la frecuencia de dos, lo cual se relaciona con un ciclo de aproximadamente medio día. Esto es coherente con la observación previa de dos picos diarios en las concentraciones de PM2.5. En contraste, el Ozono exhibe claramente su energía principal en la frecuencia de uno, correspondiente a un ciclo diario, lo que concuerda con la estacionalidad previamente identificada.

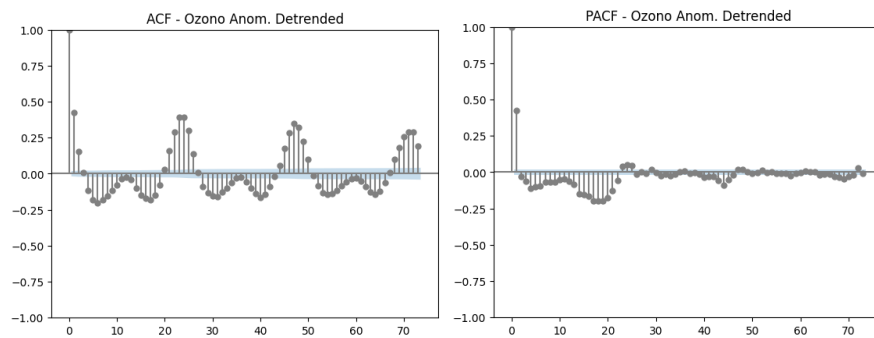
Las Gráficas 8 y 9 (escalogramas) muestran la presencia de pulsos de energía a lo largo del tiempo en periodos de 24 y 12 horas, confirmando los ciclos diarios y semidiarios. Adicionalmente, se observa un nuevo patrón para el Ozono: la existencia de pulsos constantes de energía en periodos de aproximadamente 37 horas. Esto podría deberse a la influencia de factores meteorológicos que operan en escalas de tiempo más prolongadas.

Para asegurar una estacionariedad robusta, las series se prepararon mediante dos pasos: primero, el cálculo de anomalías (estandarización Z-score), y segundo, la aplicación de una diferenciación de primer orden (utilizando la función `np.diff`) a estas anomalías. Este proceso fue fundamental para obtener series con una media y varianza más estables, y sin tendencia lineal aparente.

Posteriormente, a partir del análisis de las Funciones de Autocorrelación (ACF) y Autocorrelación Parcial (PACF) de estas series diferenciadas, se identificó y ajustó un modelo ARMA(1,1) para capturar su estructura de dependencia temporal.



**Gráfica 10:** Análisis de Autocorrelación usando ACF y PACF para las anomalías estandarizadas de PM 2.5. Fuente: *Propia*



**Gráfica 11:** Análisis de Autocorrelación usando ACF y PACF para las anomalías estandarizadas de Ozono. Fuente: *Propia*

Las Gráficas 10 y 11 demuestran cómo el proceso de cálculo de anomalías y la diferenciación lograron reducir significativamente la persistencia y la estacionalidad presentes en ambas variables. En ambas series, se observa un rápido decaimiento de la autocorrelación después del primer rezago, lo que sugiere que un modelo ARMA(1,1) sería una opción adecuada para su ajuste.

No obstante, a pesar de estas transformaciones, aún persiste cierta estacionalidad significativa en rezagos múltiplos de 24, siendo este efecto más evidente en la serie de Ozono. Si bien esta autocorrelación residual podría potencialmente beneficiar el modelo, la presencia de estos patrones estacionales complejos podría, a su vez, complicar su ajuste y capacidad predictiva

Al ajustar el modelo, se obtuvo un resumen que permitió inferir si la calidad del ajuste era adecuada para explicar la serie de tiempo de PM2.5 y Ozono.

**PM 2.5**

**Ozono**

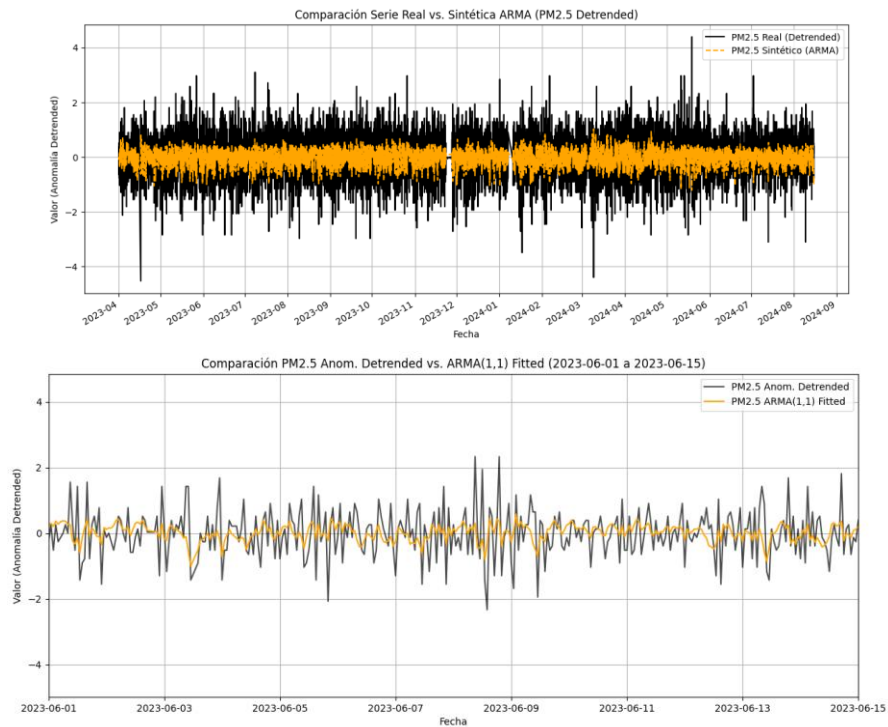
SARIMAX Results							SARIMAX Results						
Dep. Variable:	y	No. Observations:	11603				Dep. Variable:	y	No. Observations:	11603			
Model:	ARIMA(1, 0, 1)	Log Likelihood	-10955.693				Model:	ARIMA(1, 0, 1)	Log Likelihood	-5553.974			
Date:	Tue, 15 Jul 2025	AIC	21919.387				Date:	Tue, 15 Jul 2025	AIC	11115.949			
Time:	01:03:18	BIC	21948.823				Time:	01:03:18	BIC	11145.385			
Sample:	0	HQIC	21929.278				Sample:	0	HQIC	11125.839			
	- 11603							- 11603					
Covariance Type:	opg						Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]		coef	std err	z	P> z	[0.025	0.975]
const	-9.174e-05	0.001	-0.165	0.869	-0.001	0.001	const	0.0001	0.007	0.022	0.982	-0.013	0.013
ar.L1	0.6457	0.008	85.128	0.000	0.631	0.661	ar.L1	0.3774	0.015	24.364	0.000	0.347	0.408
ma.L1	-0.9677	0.003	-352.590	0.000	-0.973	-0.962	ma.L1	0.0605	0.016	3.808	0.000	0.029	0.092
sigma2	0.3869	0.004	96.426	0.000	0.379	0.395	sigma2	0.1525	0.001	146.106	0.000	0.150	0.155
Ljung-Box (L1) (Q):	29.44						Ljung-Box (L1) (Q):	0.02					
Prob(Q):	0.00						Prob(Q):	0.90					
Heteroskedasticity (H):	0.95						Heteroskedasticity (H):	1.13					
Prob(H) (two-sided):	0.12						Prob(H) (two-sided):	0.00					
Jarque-Bera (JB):	1189.01						Jarque-Bera (JB):	19672.25					
Prob(JB):	0.00						Prob(JB):	0.00					
Skew:	0.27						Skew:	-0.44					
Kurtosis:	4.47						Kurtosis:	9.32					

**Tabla 4:** Resumen de los modelos ajustados para PM2.5 y Ozono para PM 2.5 Y Ozono. Fuente: *Propia*

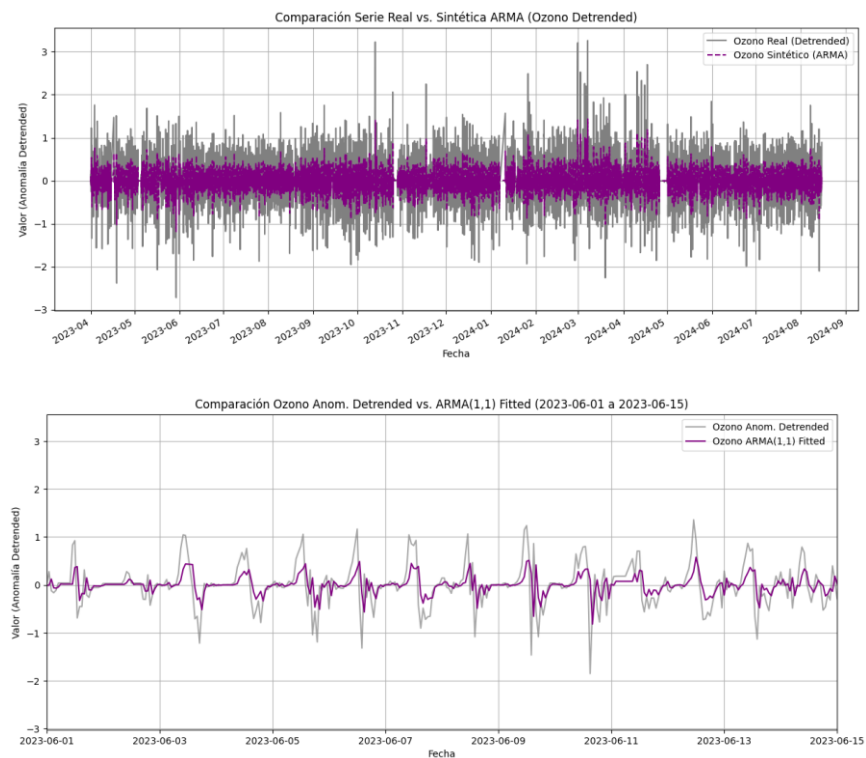
La Tabla 4 nos da información valiosa sobre el ajuste de modelos ARMA(1,1) a las series de anomalías diferenciadas:

- **Relevancia de Componentes:** Los componentes AR(1) y MA(1) mostraron relevancia estadística en ambas series (PM2.5 y Ozono).
- **Ajuste al Ozono:** El modelo ARMA(1,1) se ajustó mejor a la estructura de autocorrelación del Ozono diferenciado. Esto se evidenció en la ausencia de autocorrelación significativa en sus residuales según la Prueba de Ljung-Box.
- **Ajuste al PM2.5:** Para el PM2.5 diferenciado, los residuales presentaron autocorrelación estadísticamente significativa, indicando que el modelo no capturó completamente su estructura de dependencia.
- **Heterocedasticidad:** Los residuales del Ozono diferenciado detectaron heterocedasticidad, a diferencia de los del PM2.5.
- **Normalidad:** Los residuales de ambas series no cumplieron con el supuesto de normalidad.

Para evaluar el ajuste del modelo ARMA (1,1), se realizó una comparación visual entre los valores reales de la serie y los valores predichos por el modelo. Adicionalmente, se llevarán a cabo pruebas de diagnóstico para verificar si los residuos cumplen con los supuestos de independencia, constancia de la varianza y normalidad de la distribución, características esperadas en un proceso de ruido blanco.



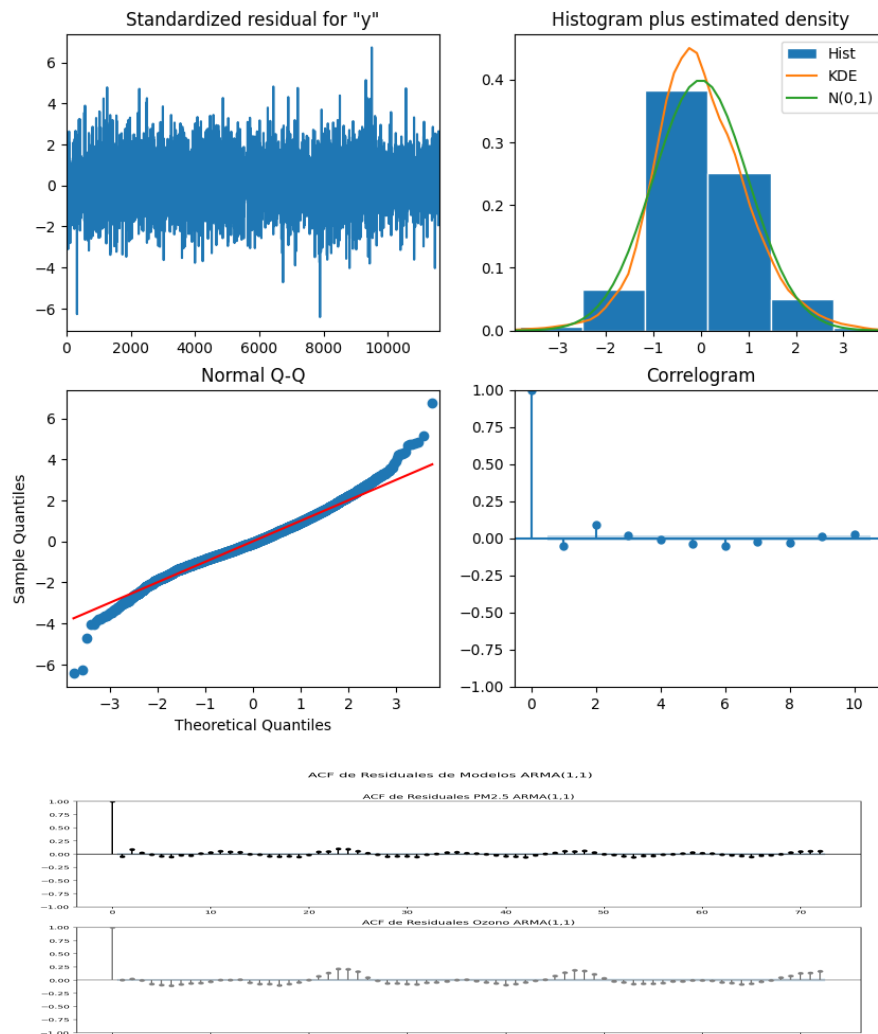
Gráfica 12: Comparación de valores reales con los valores generados por el modelo ARMA con la variable PM 2.5. Fuente: *Propia*



Gráfica 13: Comparación de valores reales con los valores generados por el modelo ARMA con la variable Ozono. Fuente: *Propia*

Gráficas 12 y 13 demuestran que el modelo logra seguir la tendencia general de la serie real, sin embargo, no consigue capturar con precisión la variabilidad ni los picos y valles agudos. Esto resulta en un efecto de suavizado de la serie de tiempo, similar al que se obtendría con una media móvil. La principal limitación observada es que el modelo no logra replicar la fuerte componente estacional presente en nuestras variables.

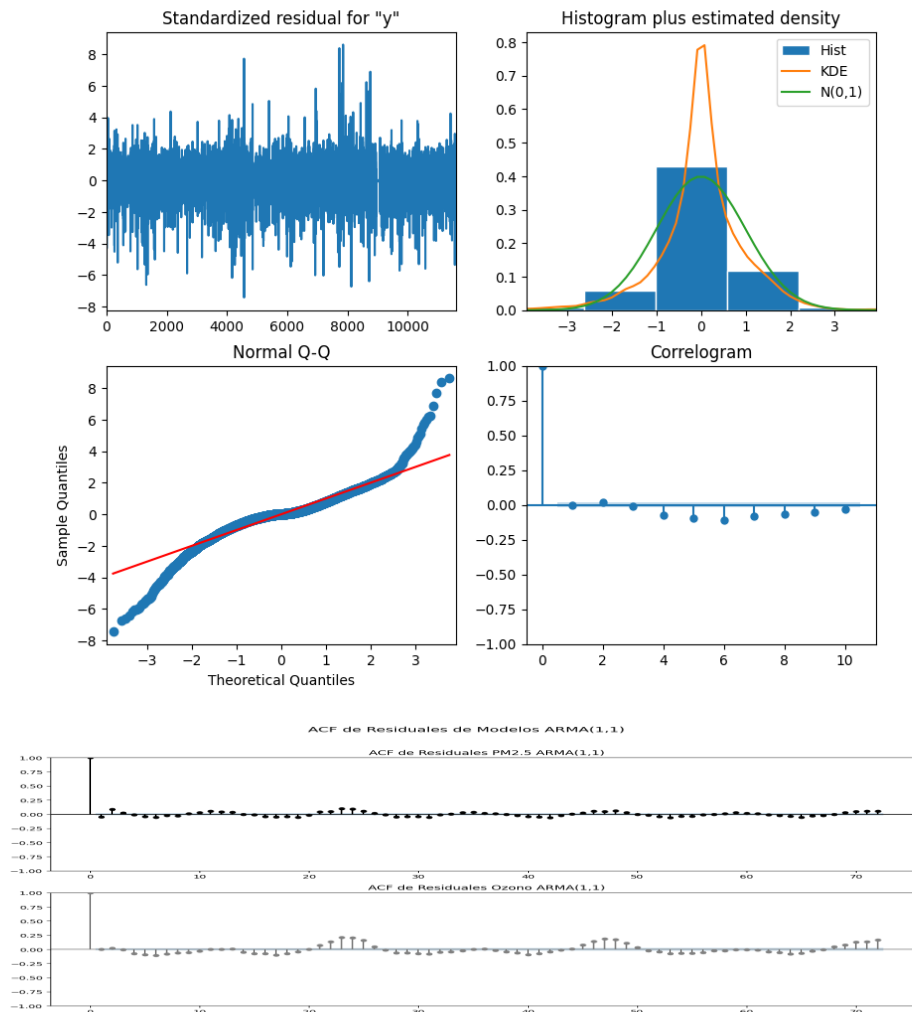
Diagnóstico de Residuales PM2.5 ARMA(1,1)



Gráfica 14: Histograma, QQ-plot, residuales vs. ajustados, ACF de residuales para PM 2.5. Fuente: *Propia*



Diagnóstico de Residuales Ozono ARMA(1,1)



Gráfica 15: Histograma, QQ-plot, residuales vs. ajustados, ACF de residuales para Ozono. Fuente: *Propia*

Las Gráficas 14 y 15 ofrecen un análisis más detallado de los residuales del modelo ARMA, permitiendo evaluar la calidad de su ajuste.

- Residuales estandarizados ("standardized for y"): Se observa que los residuales no presentan un comportamiento completamente aleatorio, conservando cierta información temporal y mostrando ciertos patrones. Esto indica que el modelo no logró ajustar adecuadamente la estacionalidad diaria.
- Histograma: El histograma de los residuales confirma que no exhiben una distribución completamente normal. Esta observación ya había sido sugerida por la Prueba de Jarque-Bera, cuyos resultados se presentaron en el resumen del modelo.
- Normal Q-Q Plot: Los puntos en el gráfico Normal Q-Q se desvían significativamente de la línea recta, lo que confirma visualmente que los residuales no están distribuidos normalmente.

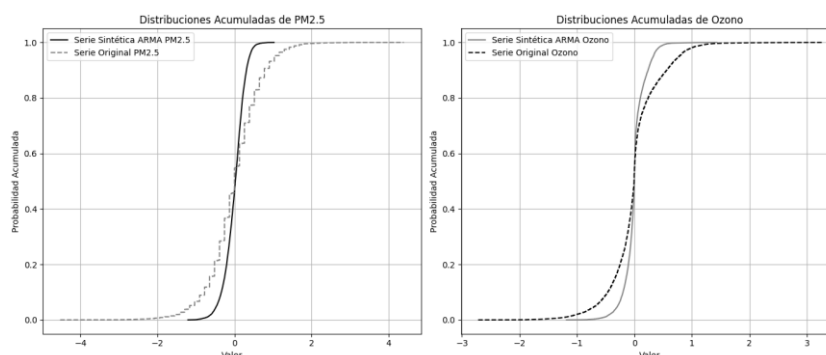
- Correlograma (ACF de residuales): El correlograma muestra picos significativos en la Función de Autocorrelación (ACF) de los residuales, particularmente en los *lags* que son múltiplos de 24. Esta persistencia en la autocorrelación residual es una evidencia clara de que la estacionalidad no fue capturada por el modelo, permaneciendo como un componente remanente en los errores.

### PM 2.5

### Ozono

Parámetro	Valor	Parámetro	Valor
Estadístico KS	0,236	Estadístico KS	0,137
Valor p	9,78E-286	Valor p	1,41E-96

**Tabla 5:** Resultados de la Prueba de Kolmogorov-Smirnov para PM 2.5 Y Ozono. Fuente: *Propia*



**Gráfica 16:** Distribuciones acumuladas para PM 2.5 y Ozono. Fuente: *Propia*

Como se muestra en la Tabla 5, el valor p obtenido es menor que el nivel de significancia (usualmente 0.05). Esto nos lleva a rechazar la hipótesis nula, concluyendo que las distribuciones de las series reales (anomalías detrended) son estadísticamente diferentes de las distribuciones de las series sintéticas generadas por los modelos ARMA(1,1) ajustados.

Adicionalmente, la Gráfica 16 corrobora visualmente este hallazgo: se observa una clara separación entre la distribución acumulada de las series sintéticas y las series reales. Esto confirma aún más que el modelo no logró replicar la distribución de probabilidad completa de los datos originales.

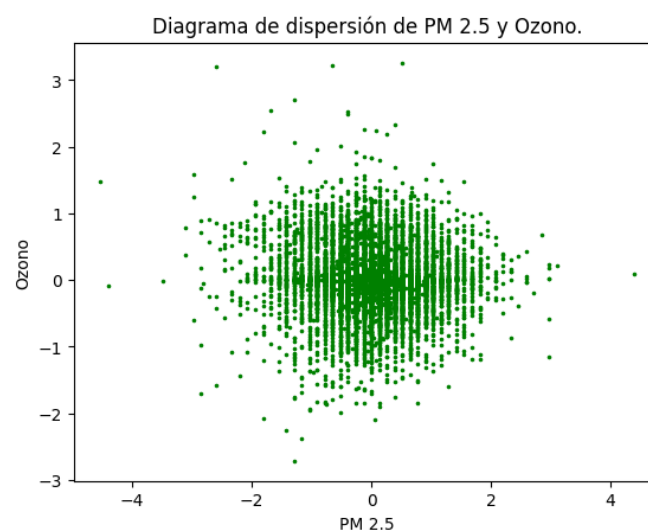
Para finalizar el estudio, se llevó a cabo un análisis de correlación entre las dos variables, con el objetivo de determinar la naturaleza de su relación (lineal o no lineal). Adicionalmente, se intentó realizar un ajuste de regresión utilizando la concentración de PM2.5 como variable independiente y la de Ozono como variable dependiente.

Se calculó el Coeficiente de Correlación de Pearson para evaluar la relación lineal entre las concentraciones de PM2.5 y Ozono. Los resultados obtenidos son los siguientes:

Parámetro	Valor
Coefficiente de Correlación de Pearson	-0,0656
Valor p	1,48E-12

Tabla 5: Coeficiente de Correlación de Pearson entre PM 2.5 Y Ozono. Fuente: *Propia*

El Coeficiente de Correlación de Pearson obtenido fue de -0.0656, lo que indica una correlación lineal débil entre las concentraciones de PM2.5 y Ozono. Sin embargo, el valor p (1.48e-12) es extremadamente bajo y significativamente menor que el nivel de significancia de 0.05. Esto nos permite rechazar la hipótesis nula, confirmando que existe una correlación lineal estadísticamente significativa entre ambas variables, a pesar de la baja magnitud de dicha correlación.



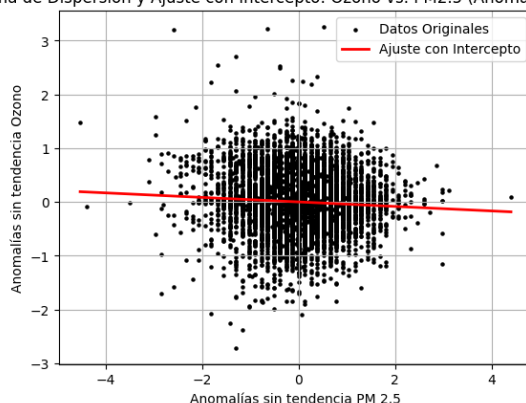
Gráfica 17: Diagrama de dispersión para comparar PM 2.5 y Ozono. Fuente: *Propia*

La gráfica 17 confirma visualmente la debilidad de la relación lineal. Los puntos están muy dispersos y no forman una nube que se acerque claramente a una línea recta, ya sea positiva o negativa.

OLS Regression Results						
Dep. Variable:	y	R-squared (uncentered):	0.004			
Model:	OLS	Adj. R-squared (uncentered):	0.004			
Method:	Least Squares	F-statistic:	50.19			
Date:	Tue, 15 Jul 2025	Prob (F-statistic):	1.47e-12			
Time:	01:03:22	Log-Likelihood:	-6699.3			
No. Observations:	11603	AIC:	1.340e+04			
Df Residuals:	11602	BIC:	1.341e+04			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
x1	-0.0420	0.006	-7.085	0.000	-0.054	-0.030
Omnibus:	1052.189	Durbin-Watson:	1.153			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	6741.134			
Skew:	0.164	Prob(JB):	0.00			
Kurtosis:	6.720	Cond. No.	1.00			

Tabla 6: Resumen del modelo de regresión lineal con intercepto. Fuente: *Propia*

Diagrama de Dispersión y Ajuste con Intercepto: Ozono vs. PM2.5 (Anomalías Detrended)



Gráfica 18: Diagrama de dispersión con ajuste de regresión lineal. Fuente: *Propia*

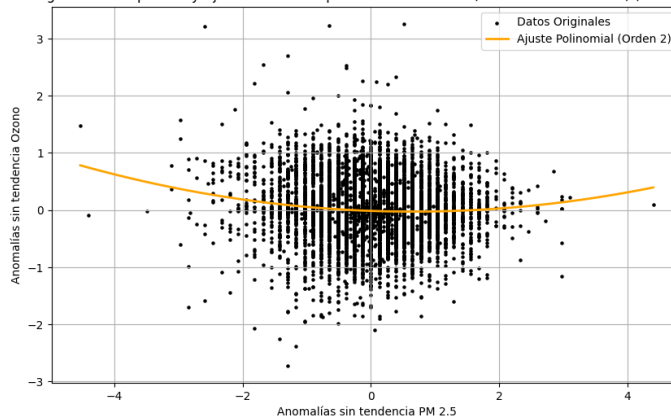
Como se detalla en la Tabla 6, el coeficiente de determinación  $R^2$  es extremadamente bajo (0.004). Esto significa que el modelo de regresión lineal simple (considerando tanto la versión con intercepto como sin él) apenas explica el 0.4% de la varianza en las anomalías *detrended* de Ozono, utilizando las variaciones en las anomalías *detrended* de PM2.5 como predictor.

Este resultado confirma que la relación lineal, si bien es estadísticamente significativa, es prácticamente insignificante en cuanto a su capacidad para predecir las concentraciones de Ozono a partir de las de PM2.5.

Parámetro del Modelo	Valor Estimado
Coeficiente $x^2$	3,011
Coeficiente $x$	-3,937
Término Independiente (Intercepto)	-1,363
$R^2$	0,008

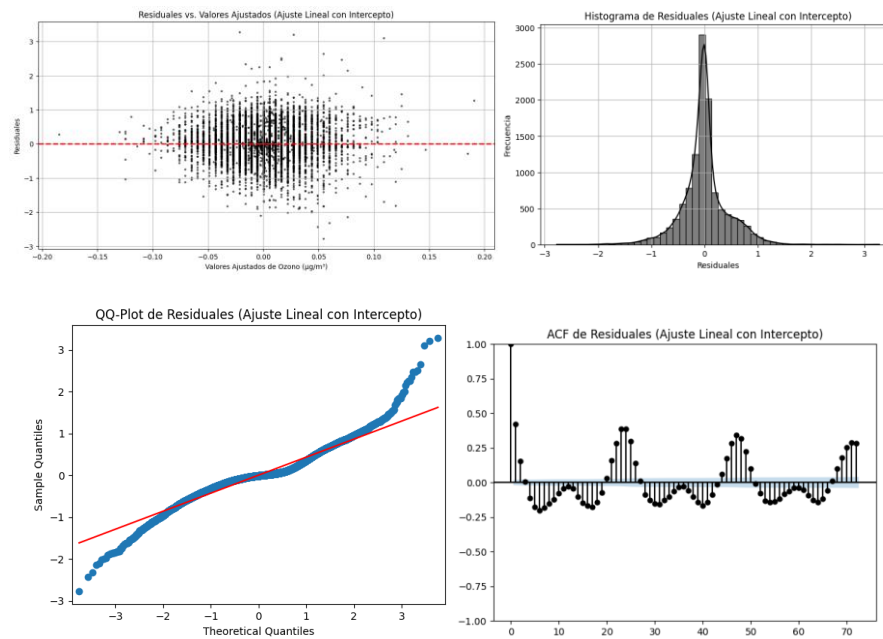
Tabla 7: Resumen del modelo de regresión lineal con intercepto. Fuente: *Propia*

Diagrama de Dispersión y Ajuste con Intercepto: Ozono vs. PM2.5 (Anomalías Detrended) ( $R^2=0.008$ )

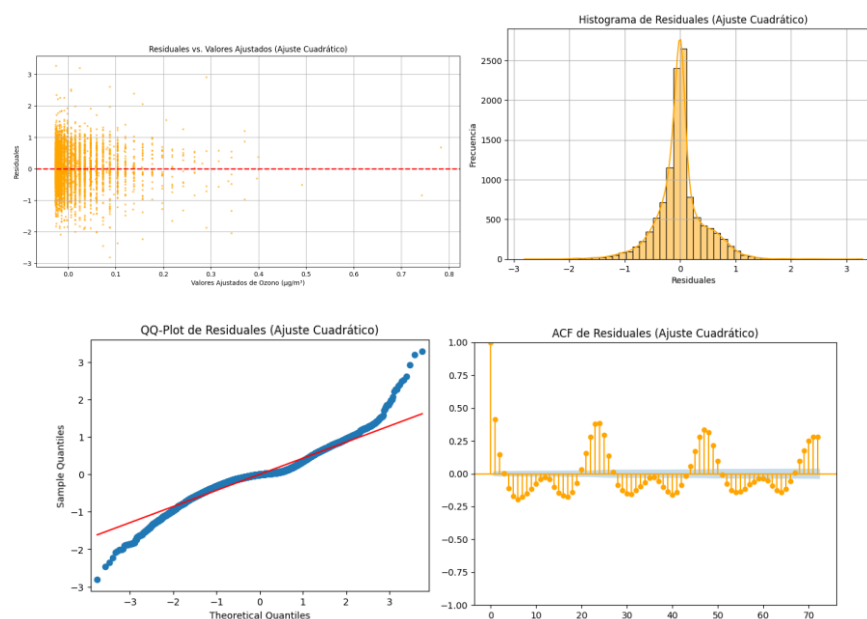


Gráfica 19: Diagrama de dispersión con ajuste de regresión cuadrática. Fuente: *Propia*

Como se presenta en la Tabla 7, el coeficiente de determinación  $R^2$  del ajuste cuadrático (0.008) es ligeramente superior al del ajuste lineal (0.004). Sin embargo, este valor sigue siendo extremadamente bajo (0.8%). Esto sugiere que la inclusión de un término cuadrático mejora el ajuste el doble, pero la relación entre las variables permanece muy débil, y el modelo aún explica una fracción minúscula de la varianza observada en las concentraciones de Ozono.



Gráfica 20: Residuales vs. Ajustados, histograma, QQ-plot, , ACF de residuales para regresión lineal  
Fuente: Propia



Gráfica 21: Residuales vs. Ajustados, histograma, QQ-plot, , ACF de residuales para regresión cuadrática  
Fuente: Propia

Las Gráficas 20 y 21 permiten realizar un análisis de los residuales similar al efectuado para el modelo ARMA.

- Residuales vs valores ajustados: Se observa que los residuales no presentan un comportamiento completamente aleatorio, y además, la variabilidad de los residuales parece cambiar a medida que cambian los valores ajustados.
- Histograma: El histograma de los residuales confirma que no exhiben una distribución completamente normal. Son asimétricos y con colas más pesadas.
- Normal Q-Q Plot: Los puntos en el gráfico Normal Q-Q se desvían significativamente de la línea recta, lo que confirma visualmente que los residuales no están distribuidos normalmente.
- Correlograma (ACF de residuales): Los gráficos de la Función de Autocorrelación (ACF) de los residuales, tanto para el modelo lineal como para el cuadrático, muestran picos significativos, especialmente en los *lags* que son múltiplos de 24. Esto indica claramente que los residuales no son independientes y que aún se observan remanentes de la estacionalidad de las variables originales.

## Discusión:

La Imagen 1 permite observar que, a pesar del predominio de valores bajos, se presentan eventos esporádicos de alta contaminación que resultan en concentraciones significativamente elevadas.

En el caso del PM<sub>2.5</sub>, estos picos pueden estar relacionados con factores como flujos vehiculares excepcionalmente altos o temporadas festivas que implican el uso de explosivos. Estos eventos pueden generar un incremento drástico en la concentración del contaminante durante periodos específicos del día. Adicionalmente, el efecto de inversión térmica, frecuente en el Valle de Aburrá durante las mañanas, puede atrapar contaminantes, contribuyendo a estas altas concentraciones.

La alta variabilidad del Ozono se atribuye a su naturaleza como contaminante secundario. Su formación y destrucción dependen fuertemente de la radiación solar, lo que provoca fluctuaciones drásticas a lo largo del día. En contraste, el PM<sub>2.5</sub> tiende a ser más persistente en la atmósfera, lo que se traduce en concentraciones relativamente más estables.

En la estacionalidad observada en la gráfica 5 podemos analizar cosas importantes.

Con el PM 2.5

**Mañana:** Hay una elevada emisión primaria del tráfico (vehículos diésel/gasolina en hora pico) combinada con una capa límite muy baja y atmósfera estable. Esta mezcla puede producir alta concentración de PM 2.5 cerca del suelo.

**Mediodía:** La intensa radiación solar calienta la superficie, elevando la capa límite y promoviendo turbulencia. Esto diluye el PM 2.5 acumulado por la mañana, de modo que las concentraciones caen. Además, termina la hora pico y el tráfico disminuye. En la gráfica se ve una especie de valle debido a la mayor dispersión atmosférica.

**Tarde:** Al regresar la población, vuelve a aumentar el tráfico vehicular. La capa límite empieza a descender con el atardecer, reduciéndose la dilución vertical. Esto genera un segundo pico vespertino de PM2.5. El efecto de la capa límite no sería tan marcado como por la mañana. Noche 21-5 h: Las fuentes primarias se reducen drásticamente (casi nulo tráfico y emisiones industriales bajas o controladas). Aunque la capa límite es muy baja, la falta de aporte mantiene las concentraciones de PM2.5. en mínimos nocturnos

Con el ozono

**Madrugada/amanecer:** Sin luz solar no hay fotólisis de NO<sub>2</sub> ni generación de radicales. El ozono existente se consume (por reacción con NO y compuestos orgánicos). Por ello, las concentraciones alcanzan su valor mínimo antes de que amanezca

**Día:** Con la luz solar intensa se activa la química fotoquímica. El NO<sub>2</sub> se fotodisocia ( $\text{NO}_2 + h\nu \rightarrow \text{NO} + \text{O}$ ), y el átomo O reacciona con O<sub>2</sub> para formar O<sub>3</sub>. A medida que avanza la mañana y la radiación UV se hace más fuerte, el ozono se acumula rápidamente, alcanzando típicamente su valor máximo alrededor del mediodía o primeras horas de la tarde.

**Atardecer/noche:** Al disminuir la radiación UV cesa la producción fotolítica de ozono. El ozono remanente se va destruyendo gradualmente por reacciones químicas y deposición. Así, tras el pico vespertino las concentraciones de O<sub>3</sub> decaen hacia la medianoche.

(Bougault et al., 2024; Jaeglé, 2019; Wang et al., 2019)

En relación con la tendencia de las series de tiempo, aunque se detecta un ligero decrecimiento en las concentraciones de PM<sub>2.5</sub>, este es de muy baja magnitud. Esto se puede atribuir a que, a pesar del gran volumen de datos, el período de estudio no abarca una extensión temporal suficiente (menos de año y medio después del filtrado) para que sea probable observar una tendencia muy marcada en fenómenos ambientales.

Por conocimiento general del área, es posible que los niveles de contaminantes en Itagüí tiendan a aumentar a largo plazo. Itagüí es una zona industrial que, además de generar sus propias emisiones, recibe contaminantes de otros municipios del Valle de Aburrá. Las condiciones geográficas propias del valle también contribuyen a crear un entorno propicio para la acumulación de contaminantes. A pesar de estos factores, el municipio realiza esfuerzos para mantener los niveles dentro del límite normativo, buscando mitigar un incremento sostenido de la contaminación con el tiempo.

Se observa una discrepancia notoria entre los resultados de la Prueba de Dickey-Fuller Aumentada (ADF) y la inspección visual de los gráficos ACF y PACF. Mientras la prueba ADF indica que las series son estacionarias (rechazando la hipótesis nula de no



estacionariedad), los correlogramas sugieren lo contrario al mostrar patrones claros de dependencia temporal. (CrossValited, s.f.)

Esta discrepancia probablemente se debe a que la prueba ADF detecta la ausencia de una raíz unitaria (es decir, una tendencia estocástica). Sin embargo, la fuerte estacionalidad determinística (el patrón diario repetitivo) y, posiblemente, una tendencia determinística subyacente (a pesar de que Mann-Kendall con *pre-whitening* no encontró una tendencia significativa para Ozono, sí la detectó para PM2.5) aún provocan que la serie se comporte como "no estacionaria" desde una perspectiva visual y para fines de modelado práctico. (CrossValited, s.f.)

El modelo ARMA(1,1) ajustado a las variables presenta numerosas limitaciones, lo que compromete la efectividad de las predicciones. Este tipo de modelos está diseñado para capturar dependencias lineales y autocorrelación de corto plazo (principalmente del valor anterior). Sin embargo, carece de una forma efectiva de modelar patrones que se repiten en intervalos fijos más largos, como la prominente estacionalidad diaria de 24 horas observada en nuestros datos.

Los diagnósticos de residuales revelaron un claro incumplimiento de los supuestos del modelo. Específicamente, se detectó autocorrelación significativa en los residuales, ausencia de normalidad en su distribución y una varianza no constante (heterocedasticidad). Además, la inspección visual de las distribuciones acumuladas mostró una clara diferencia entre las series reales y las generadas por el modelo, lo que confirma que el ARMA(1,1) es altamente limitado en su capacidad para replicar la distribución de probabilidad completa de los datos.

En consecuencia, estas limitaciones del modelo conllevan a predicciones inexactas, interpretaciones engañosas e inferencias estadísticas incorrectas.

Para series de tiempo que presentan una estacionalidad tan marcada, el modelo SARIMA (Seasonal Autoregressive Integrated Moving Average) es una alternativa más adecuada. Un modelo SARIMA es una extensión del modelo ARIMA (el cual, a su vez, amplía el ARMA para manejar la no estacionariedad mediante diferenciación) que incluye componentes estacionales.

La principal ventaja del SARIMA es su capacidad para modelar explícitamente la estacionalidad y manejar la no estacionariedad de la serie. Sin embargo, es crucial que, tras el ajuste, los residuales del modelo sigan exhibiendo las características de un ruido blanco para asegurar su validez. (Majka, 2024 y GeeksforGeeks, 2024)

El análisis de correlación entre las dos variables reveló una relación muy débil. El Coeficiente de Correlación de Pearson fue de -0.0656, lo que indica una correlación lineal negativa, aunque estadísticamente significativa.

Complementariamente, las regresiones lineal y cuadrática reafirmaron esta debilidad, con coeficientes de determinación ( $R^2$ ) de 0.004 y 0.008 respectivamente. Estos valores extremadamente bajos confirman que la relación lineal o cuadrática simple entre las anomalías diferenciadas es extremadamente tenue.

Esto implica que, en promedio y de forma lineal, un incremento en la concentración de una variable se asocia con un descenso muy ligero en la otra. No obstante, la magnitud de esta relación es tan insignificante que la concentración de una variable apenas contribuye a predecir la de la otra en un momento dado.

El PM2.5 se considera un contaminante secundario. Por lo tanto, durante ciertas horas, se puede observar una relación inversa: altas concentraciones de NO y PM2.5 coincidiendo con niveles bajos de Ozono.

Un factor importante para considerar es que una elevada abundancia de PM2.5 puede atenuar la radiación solar que llega al Ozono, lo que podría resultar en una ligera disminución de su concentración. Este fenómeno podría contribuir a explicar, en parte, la débil dependencia lineal inversa observada. (Yadav et al., 2023; Yawei et al., 2023)

Durante el día, la producción fotoquímica de Ozono se vuelve el proceso dominante, y su relación con el PM2.5 se torna más compleja. Aunque ambos pueden compartir precursores, sus mecanismos de formación y dispersión son distintos. Variables meteorológicas como la temperatura, la radiación solar, la velocidad y dirección del viento, y la altura de la capa de mezcla atmosférica influyen de manera diferenciada en la formación, transporte y dispersión tanto del Ozono como del PM2.5, lo que complejiza aún más su interrelación. (Yadav et al., 2023; Yawei et al., 2023)

## **Conclusiones:**

El comportamiento temporal de las concentraciones de PM2.5 y ozono en Itagüí está dominado por ciclos diarios muy definidos y predecibles. El PM2.5 exhibe un patrón bimodal, con picos en la mañana y la noche consistentes con las horas pico de tráfico y las condiciones de la capa límite atmosférica. En contraste, el ozono presenta un patrón con un máximo pronunciado alrededor del mediodía, lo que confirma su fuerte dependencia de la radiación solar para su formación fotoquímica. Sin embargo, la fuerte estacionalidad y persistencia de estas series no fueron capturadas adecuadamente por el modelo ARMA(1,1) implementado. El análisis de los residuales del modelo confirmó esta limitación, mostrando una autocorrelación significativa remanente en lags múltiplos de 24 (también de 12 en el caso de PM 2.5), lo que indica que el modelo es insuficiente para fines predictivos y que se requieren modelos más avanzados como SARIMA que puedan incorporar explícitamente la componente estacional.

La relación entre las concentraciones de PM2.5 y ozono, una vez eliminadas las tendencias y la estacionalidad mediante diferenciación, es estadísticamente significativa pero extremadamente débil y de naturaleza negativa. Los modelos de regresión, tanto

lineal como cuadrático, confirmaron la insignificancia práctica de esta relación, explicando menos del 1% de la varianza en los datos ( $R^2$  de 0.004 y 0.008, respectivamente). Esto demuestra que conocer la concentración de un contaminante no aporta información útil para predecir la del otro a través de un modelo lineal simple. Dicha debilidad sugiere que la interacción entre PM<sub>2.5</sub> y ozono es mucho más compleja, influenciada por una variedad de factores no lineales y variables meteorológicas que afectan sus procesos de formación, dispersión y destrucción de manera diferenciada. Si se desea continuar este estudio, es altamente recomendable incorporar análisis multivariados y modelos que permitan que la relación entre las variables varíe a lo largo del tiempo. Esto facilitará una comprensión más profunda y precisa del comportamiento de las variables.

Para futuros estudios es recomendable incorporar variables meteorológicas clave como temperatura, humedad relativa, velocidad y dirección del viento, radiación solar y altura de la capa de mezcla atmosférica. Estas variables son a menudo las que mejor explican la variabilidad y la interrelación de los contaminantes y mejorarían significativamente la capacidad de los modelos predictivos.

## Bibliografía:

Bougault, V., Valorso, R., Sarda, R., Baisnee, D., Albin, M., Brissebrat, G., Chapuy, D., Cheymol, J., De Marco, G., & Dugué, P. (2024, Julio). *Paris air quality monitoring for the 2024 Olympics and Paralympics: focus on air pollutants and pollen*. British Journal of Sports Medicine. [https://www.researchgate.net/publication/382589184\\_Paris\\_air\\_quality\\_monitoring\\_for\\_the\\_2024\\_Olympics\\_and\\_Paralympics\\_focus\\_on\\_air\\_pollutants\\_and\\_pollen](https://www.researchgate.net/publication/382589184_Paris_air_quality_monitoring_for_the_2024_Olympics_and_Paralympics_focus_on_air_pollutants_and_pollen)

CrossValidated (s.f.). *ACF does not decay though it has passed ADF stationarity test*. Cross Validated. <https://stats.stackexchange.com/questions/412878/acf-doesnot-decay-though-it-has-passed-adf-stationarity-test>

CrossValidated (s.f.). *The difference between the three augmented Dickey-Fuller test (none, drift, trend)*. Cross Validated. <https://stats.stackexchange.com/questions/131148/the-difference-between-the-three-augmented-dickey-fuller-test-none-drift-trend>

CrossValidated (s.f.). *Is my Data stationary? KPSS, ADF Tests and ACF*. Cross Validated. <https://stats.stackexchange.com/questions/418667/is-my-data-stationary-kpss-adf-tests-and-acf>

GeeksforGeeks. (2024, 26 de marzo). *ARIMA vs SARIMA Model*. <https://www.geeksforgeeks.org/machine-learning/arima-vs-sarima-model/>

Jaeglé, L. (2019). *Tropospheric ozone and nitrogen oxides*. [https://faculty.washington.edu/jaegle/558/ozone\\_NOx.pdf](https://faculty.washington.edu/jaegle/558/ozone_NOx.pdf)

Majka, M. (2024, Septiembre). *Seasonal time series analysis: Why SARIMA outshines ARIMA*. LinkedIn. <https://www.linkedin.com/pulse/seasonal-time-series-analysis-why-sarima-outshines-arima-marcin-majka-bg2gf>

Pérez-Carrasquilla, J. S., Montoya, P. A., Sánchez, J. M., Hernández, K. S., & Ramírez, M. (2023). Forecasting 24 h averaged PM<sub>2.5</sub> concentration in the Aburrá Valley using tree-based machine learning models, global forecasts, and satellite information. *Advances in Statistical Climatology, Meteorology and Oceanography*, 9, 121–134.

Wang, C., Hu, X.-M., Feron, S., Leffel, J., & Cordero, R. R. (2025). Compound heat and ozone pollution in the urban environment. *Urban Climate*, 62. <https://doi.org/10.1016/j.uclim.2025.102511>

Wang, C., Jia, M., Xia, H., Wu, Y., Wei, T., Shang, X., Yang, C., Xue, X., & Dou, X. (2019). Relationship analysis of PM<sub>2.5</sub> and boundary layer height using an aerosol and turbulence detection lidar. *Atmospheric Measurement Techniques*, 12(6), 3303–3316. <https://amt.copernicus.org/articles/12/3303/2019/>

Y. Qu, T. Wang, C. Yuan, H. Wu, L. Gao, C. Huang, et al., The underlying mechanisms of PM<sub>2.5</sub> and O<sub>3</sub> synergistic pollution in East China: photochemical and heterogeneous interactions, *Sci. Total Environ.* 873 (2023) 162434, <https://doi.org/10.1016/j.scitotenv.2023.162434>.

Yadav, R. K., Gadhavi, H., Arora, A., Mohbey, K. K., Kumar, S., Lal, S., & Mallik, C. (2023). Relation between PM<sub>2.5</sub> and O<sub>3</sub> over Different Urban Environmental Regimes in India. *Urban Science*, 7(1), 9. <https://doi.org/10.3390/urbansci7010009>