

UNIVERSIDAD PERUANA UNIÓN
FACULTAD DE INGENIERIA Y ARQUITECTURA
Escuela Profesional de ingeniería Ambiental



**Pronóstico a corto plazo de la concentración de ozono en Lima metropolitana
utilizando una combinación híbrida de modelos de series de tiempo**

Tesis para obtener el Título Profesional de Ingeniero Ambiental

Autor:

Flor Vicky Quispe Sanchez

Eddy Albino Salcedo Esquen

Asesor:

PhD. Javier Linkolk López Gonzales

Lima, 26 de marzo del 2024

DECLARACIÓN JURADA DE ORIGINALIDAD DE TESIS

Yo López Gonzales Javier Linkolk docente de la Escuela de Posgrado de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: **“Pronóstico a corto plazo de la concentración de ozono en Lima metropolitana utilizando una combinación híbrida de modelos de series de tiempo”** del (los) autor (autores) Flor Vicky Quispe Sanchez, Eddy Albino Salcedo Esquen, tiene un índice de similitud de 12 % verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de **Lima** a los 16 días del mes de Abril del año 2024.



Javier Linkolk López Gonzales

ACTA DE SUSTENTACIÓN DE TESIS

En Lima, Naña, Villa Unión, a 26 día(s) del mes de marzo del año 2024 siendo las 9:30 horas, se reunieron los miembros del jurado en la Universidad Peruana Unión Campus Lima, bajo la dirección del (de la) presidente(a):

Mg. Lilian del Carmen Gutiérrez Rodríguez, el (la) secretario(a): Mg. Jorl Hugo Fernández Rojas y los demás miembros: Mg. Jackson Edgardo Pérez

Carpio; Ing. Orlando Alan Poma Porras y el (la) asesor(a) Dr. Javier Linkolk López González

con el propósito de administrar el acto académico de sustentación de la tesis titulado: "Pronóstico a corto plazo de la concentración de ozono en Lima metropolitana utilizando una combinación híbrida de modelos de series de tiempo"

del(los) bachiller/es: a) Flor Vicky Quispe Sanchez
b) Eddy Albino Salcedo Esquen
c)

conducente a la obtención del título profesional de:

Ingeniero Ambiental

(Denominación del Título Profesional)

El Presidente inició el acto académico de sustentación invitando al (a la) / a (los) (las) candidato(a)s hacer uso del tiempo determinado para su exposición. Concluida la exposición, el Presidente invitó a los demás miembros del jurado a efectuar las preguntas, y aclaraciones pertinentes, las cuales fueron absueltas por al (a la) / a (los) (las) candidato(a)s. Luego, se produjo un receso para las deliberaciones y la emisión del dictamen del jurado.

Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Bachiller (a): Flor Vicky Quispe Sanchez

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	18	A-	Muy bueno	sobresaliente

Bachiller (b): Eddy Albino Salcedo Esquen

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	18	A-	Muy bueno	sobresaliente

Bachiller (c):

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	

(*) Ver parte posterior

Finalmente, el Presidente del jurado invitó al (a la) / a (los) (las) candidato(a)s a ponerse de pie, para recibir la evaluación final y concluir el acto académico de sustentación procediéndose a registrar las firmas respectivas.

Presidente/a

Secretario/a

Asesor/a

Miembro

Miembro

Bachiller (a)

Bachiller (b)

Bachiller (c)

* Esta sustentación fue realizada de manera virtual u online simultánea conforme al Reglamento General de Grados y Títulos.

Pronóstico a corto plazo de la concentración de ozono en Lima Metropolitana utilizando una combinación híbrida de modelos de series de tiempo.

Short-term forecast of ozone concentration in Metropolitan Lima using a hybrid combination of time series models

Resumen: El aumento de los niveles globales de ozono en las últimas décadas ha perjudicado la salud humana. Lima, Perú, es una de las diez ciudades de América del Sur con los peores niveles de contaminación del aire. Por lo tanto, la elaboración de modelos y pronósticos eficientes y precisos son fundamentales para las concentraciones de ozono en Lima. Para este propósito, utilizamos los datos de concentración de ozono por hora en áreas metropolitanas, para pronósticos de concentración de ozono de varios pasos (uno, dos, tres y siete días de anticipación). Para ello, dividimos la serie horaria del ozono en dos nuevos componentes: determinista y estocástico. Para verificar el desempeño de la metodología propuesta, evaluamos modelos en cuatro horizontes de pronóstico diferentes y calculamos seis errores medios de precisión diferentes, una prueba estadística y evaluaciones gráficas. Los diversos resultados de pronóstico fuera de la muestra de horizonte para los datos considerados sugieren que la técnica de pronóstico basada en componentes propuesta proporciona una ganancia altamente consistente, precisa y eficiente. El modelo NPAR tuvo el error mínimo previsto y el mejor impacto predictivo en comparación con los demás propuestos. Esto puede ampliarse a otros distritos de Lima, otras regiones del Perú e incluso a nivel global para evaluar la eficacia del enfoque propuesto de modelado y pronóstico basado en componentes.

Summary: The increase in global ozone levels in recent decades has harmed human health. Lima, Peru, is one of the ten cities in South America with the worst levels of air pollution. Therefore, efficient and accurate modeling and forecasting are essential for ozone concentrations in Lima. For this purpose, we use hourly ozone concentration data in metropolitan areas, for multi-step ozone concentration forecasts (one, two, three and seven days ahead). To do this, we divide the ozone hourly series into two new components: deterministic and stochastic. To verify the performance of the proposed methodology, we evaluate models at four different forecast horizons and calculate six different mean precision errors, a statistical test, and graphical evaluations. The various horizon out-of-sample forecasting results for the considered data suggest that the proposed component-based forecasting technique provides highly consistent, accurate and efficient gain. The NPAR model had the minimum predicted error and the best predictive impact compared to the other proposed ones. This can be extended to other districts of Lima, other regions of Peru and even globally to evaluate the effectiveness of the proposed modeling and forecasting approach based on components.

Palabras clave: Previsión del ozono con varios pasos adelante; Salud global; Modelo de regresión lineal múltiple; Modelos de series temporales; Técnica de previsión basada en componentes.

Keywords: Ozone forecast several steps ahead; global health; Multiple linear regression model; Time series models; Component-based forecasting technique.

1. Introducción.

Un problema importante en el mundo industrializado moderno es la contaminación del aire, que tiene graves efectos toxicológicos tanto en la salud humana como en el medio ambiente.

El estudio realizado por la Universidad de Monash en Australia encontró que los criterios recomendados por la Organización Mundial de la Salud para la calidad del aire no se cumplían donde residía el 99% de la población mundial.

Aunque diversas actividades físicas liberan contaminantes, la liberación involuntaria de sustancias químicas peligrosas es la principal causa de contaminación. Las elevadas concentraciones de ozono troposférico (O_3) indican un peligro importante para el clima y el medio ambiente. Además, la dispersión de O_3 se ve obstaculizada por el cambio climático como consecuencia de la actividad industrial y urbanización.

Los indicadores importantes de la contaminación del aire incluyen el dióxido de nitrógeno (NO_2), el ozono (O_3), el índice de aerosoles absorbentes (AAI) y el monóxido de carbono (CO). Los principales contribuyentes a la producción de NO_2 atmosférico incluyen las emisiones del suelo, los rayos naturales, los gases de escape de los vehículos de motor, la quema de biomasa y la combustión parcial de combustibles fósiles. Desempeña un papel crucial en la síntesis de ozono troposférico a través de una serie de intrincados procesos químicos que involucran oxígeno y radicales libres producidos por la presencia de la luz solar en compuestos orgánicos volátiles (COV).

Dado que el O_3 se produce como subproducto de la reacción fotoquímica entre CO y VOC y los óxidos de nitrógeno ($NO_x = NO + NO_2$), lo que permite que sus altas concentraciones sean producidas por las emisiones de NO_x procedentes de fuentes de combustión, el O_3 se considera un contaminante secundario.

La Organización Mundial de la Salud afirma que las investigaciones epidemiológicas y toxicológicas han encontrado un apoyo significativo a la relación causal entre el O_3 superficial y las consecuencias respiratorias desfavorables. Especialmente en poblaciones de alto riesgo, estos impactos pueden incluir mortalidad, así como cambios en la función pulmonar y asma. El ozono daña los cultivos agrícolas y el follaje de las plantas. Para informar al público sobre la intervención necesaria y evaluar los efectos inmediatos de las actividades sobre el comportamiento climático, es importante anticipar y comprender la tasa de generación y emisión de ozono.

A nivel mundial, se proyectó que la exposición prolongada al ozono contribuiría a 254.000 muertes adicionales por enfermedad pulmonar obstructiva crónica. China es una de las naciones con mayores emisiones y concentraciones de ozono a escala mundial. Beijing y Shanghai han tenido la peor contaminación del aire en los últimos años, siendo los días críticos de contaminación por O_3 entre un 93 y un 575% mayores que los de otras naciones industrializadas. Por el contrario, algunas regiones de Estados Unidos y el sur de Canadá experimentan una menor exposición al ozono e incluso se las conoce como "lugares limpios". El área metropolitana de la Ciudad de México, la capital del país, tiene 21,8 millones de habitantes y se encuentra en una cuenca de gran altitud (unos 2240 m sobre el nivel del mar). La Ciudad de México está frecuentemente sujeta a episodios de ozono debido a su especial entorno topográfico, en combinación con condiciones meteorológicas y de emisiones. Estos acontecimientos parecen haber vuelto a empeorar últimamente. Sin embargo, debido a que las áreas tropicales y subtropicales tienen condiciones climáticas favorables para la producción y acumulación de ozono, como altas temperaturas, luz solar intensa y convección, la variación en las emisiones de estas regiones se puede ver en la carga global de ozono, lo que demuestra la estrecha conexión entre los factores climáticos y la concentración de O_3 . Se han logrado mejoras en la calidad del aire con ozono, particularmente en

Europa y América del Norte, reduciendo las emisiones antropogénicas de precursores del ozono como los óxidos de nitrógeno (NO_x). El Perú está incluido entre las naciones con mayores niveles de contaminación del aire. Esta condición está ligada al rápido desarrollo económico e industrial del Perú, que resulta en la producción de contaminantes y gases que afectan la calidad del aire. Como capital de Perú, Lima tiene más de un tercio de la población del país, lo que la convierte en la ciudad con mayor contaminación atmosférica de América del Sur.

Para describir las interacciones entre variables, se han aplicado ampliamente enfoques de modelos estadísticos al tema de la contaminación del aire. Una relación entre una serie de factores explicativos (predictores) y una variable de respuesta (objetivo) normalmente se establece mediante técnicas estadísticas basadas en modelos de regresión, como la regresión lineal múltiple (MLR). Pero a pesar de su aparente capacidad para ofrecer resultados razonables en muchas aplicaciones, estos tipos de modelos con frecuencia no describen la complejidad de las relaciones e interacciones no lineales entre variables, y normalmente se prefieren técnicas más avanzadas para lograr un mayor grado de precisión. En las predicciones de los niveles de concentración de contaminantes. La literatura ha proporcionado una serie de métodos estadísticos para pronosticar y evaluar los niveles de contaminación por ozono, incluido el modelo autorregresivo, el modelo autorregresivo integrado, modelo de media móvil y sus variantes.

Actualmente, la investigación científica utiliza cada vez más modelos de aprendizaje automático (ML). Debido a su capacidad para analizar conjuntos de datos enormes y complejos (big data), encontrar patrones y hacer predicciones, se han creado y utilizado cada vez más modelos estadísticos avanzados basados en enfoques de ML en el campo del modelado de la calidad del aire durante las últimas tres décadas.

Durante el brote de COVID-19 en España, el problema del NO₂ se abordó mediante técnicas de ML. realizó un estudio para pronosticar las concentraciones de ozono a nivel del suelo en Jordania utilizando varios métodos. Descubrieron que un algoritmo basado en redes neuronales artificiales funcionaba mejor que todos los demás métodos. El estudio se realizó para pronosticar las concentraciones horarias de ozono para el día siguiente utilizando un enfoque novedoso basado en redes neuronales artificiales con componentes principales como entradas.

Las redes neuronales artificiales de regresión lineal múltiple y feedforward se compararon con el modelo desarrollado, que se basa en los datos originales y utiliza regresión de componentes principales. Los resultados revelaron que, al utilizar componentes principales como entradas, las predicciones de ambos modelos mejoraron al reducir su complejidad y eliminar la colinealidad de los datos. En una investigación empírica, utilizó una máquina de vectores de soporte (SVM) estándar para pronosticar los niveles de ozono basándose únicamente en factores ambientales. Los resultados mostraron que SVM funcionó mejor que las redes neuronales al pronosticar las concentraciones máximas diarias de ozono. Con el fin de modelar las concentraciones de ozono en todo el territorio continental de Estados Unidos, evaluaron trece técnicas de ML con regresión lineal del uso de la tierra (LUR). Las técnicas de ML no lineal superaron a LUR en términos de precisión de predicción, siendo la mejora más significativa para el modelado espaciotemporal. Al ajustar los pesos de las muestras, los modelos espaciotemporales pueden anticipar las concentraciones necesarias para determinar valores de diseño de ozono que son tan buenos o mejores que los modelos espaciales. El objetivo del estudio es predecir la troposfera (O₃) utilizando un conjunto de datos de concentraciones de ozono utilizando una variedad de modelos ML, que incluyen regresión lineal, regresión de árbol, regresión de vector de soporte, regresión de conjunto, regresión de proceso gaussiano y regresión neuronal artificial. modelos de red. Para la predicción de la contaminación por ozono, evalúa la efectividad predictiva de 19 algoritmos ML. Según los hallazgos, los modelos de ML dinámicos que utilizan datos desfasados funcionan mejor que los modelos de ML estáticos y reducidos. Al comparar los modelos de ML con los modelos estáticos y

reducidos, los datos desfasados aumentan la precisión en un 300% y un 200%, respectivamente, según las medidas de RMSE.

Perú es un país sudamericano en la región del Pacífico Sudeste, y su capital, Lima, no es ajena a la contaminación del aire por ozono. Lima se ha convertido en una megaciudad con más de diez millones de habitantes y graves preocupaciones por la contaminación del aire. Romero et al. investigaron el impacto de las variables meteorológicas en las concentraciones de ozono y otros contaminantes en el aire utilizando correlaciones lineales para datos recopilados entre 2015 y 2018 en ocho estaciones de muestreo diferentes en Lima metropolitana y encontraron que este contaminante aumentaba con la irradiación solar entre las 10:00 y 16:00 horas, especialmente en primavera, posiblemente por la interacción de NO_x primarios y emisiones de hidrocarburos de los vehículos. Carbo-Bustinza et al. en cambio, investigó el comportamiento del ozono en invierno en cuatro sitios en Lima utilizando técnicas de ML y descubrió los mayores valores críticos en las regiones ATE. Sin embargo, detectamos una disminución general de los valores durante la estación fría (O₃ 100 µgm³), lo que concuerda con otro estudio. Mientras tanto, es necesario analizar exhaustivamente las series temporales de las zonas más contaminadas para optimizar el pronóstico de la concentración de ozono. En este sentido, el objetivo de este estudio es proporcionar una herramienta mejorada para pronosticar las concentraciones de ozono troposférico en cuatro distritos de la mega ciudad de Lima Metropolitana utilizando un enfoque de estimación basado en componentes. De manera muy precisa, una técnica basada en componentes combina las características de los modelos de regresión clásicos y los modelos de series temporales para crear pronósticos eficientes. Este estudio realizó las siguientes contribuciones: Mejorar la eficiencia y precisión.

Para el pronóstico de la concentración de ozono, se propone una técnica de pronóstico basada en componentes basada en el modelo de regresión lineal múltiple y cuatro modelos estándar de series de tiempo. Aplicación de la técnica de pronóstico por componentes de la base de datos de concentración de ozono en cuatro distritos (ATE, CDM, SB y Santa Anita) con episodios severos de contaminación por ozono entre 2017 y 2019 solo para la temporada de invierno. Para evaluar el rendimiento de la técnica de pronóstico basada en componentes propuesta, se utilizaron seis errores medios de precisión diferentes, incluidos tres errores medios de precisión relativa y tres errores medios de precisión absoluta, una prueba estadística y una evaluación visual. Además, se utilizan cuatro horizontes de pronóstico diferentes para evaluar el desempeño del pronóstico a corto y mediano plazo. Por otro lado, en este trabajo se comparan los resultados del mejor modelo final con los modelos base considerados. Los hallazgos mostraron que el mejor modelo de este estudio es muy preciso y eficiente en comparación con los modelos de referencia. Una propuesta metodológica aplicable al sistema de gestión ambiental para mitigar la contaminación por ozono dirigida a los actores del programa nacional de calidad del aire. Finalmente, el trabajo actual utiliza solo sus cuatro conjuntos de datos de distrito en Lima, Perú. Esto se puede extender a otros distritos de Lima, otras regiones del Perú e incluso a nivel mundial para evaluar el desempeño de la técnica de pronóstico basada en componentes propuesta.

El resto del manuscrito tiene el formato siguiente: El enfoque de pronóstico basado en componentes propuesto se explica en detalle en la Sección 2. La Sección 3 contiene los resultados de los estudios de caso para cada estación de monitoreo analizada y algunas discusiones significativas. La sección 4 presenta los resultados, así como los límites y los desafíos futuros.

2. Metodología.

En esta sección, proporcionaremos una descripción general completa de los diversos modelos y métodos utilizados para construir la técnica de pronóstico y modelado basada en componentes propuesta. Por lo tanto, las siguientes subsecciones son para información detallada sobre cada modelo y método.

2.1. La técnica propuesta de modelado y pronóstico basada en componentes.

El objetivo principal de este estudio fue predecir la concentración de ozono con uno, dos, tres y siete días de anticipación en cuatro estaciones de monitoreo: ATE, CDM, SB y STA en Lima, Perú. Sea O_h la concentración de ozono para la hora h . Para tener en cuenta con precisión los cambios en la concentración de ozono a lo largo del tiempo, sugerimos modelar O_h de la siguiente manera:

$$O_h = d_h + s_h$$

La serie de concentración de ozono (μgm^3) se divide en dos partes: un componente determinista (denotado como d_h) y un componente estocástico (denotado como s_h). d_h incluye la tendencia (patrón a largo plazo) y ciclos horarios, mientras que s_h representa fluctuaciones aleatorias. Matemáticamente, d_h se define de la siguiente manera:

$$d_h = t_h + n_h$$

El símbolo t_h representa la tendencia a largo plazo, mientras que n_h representa el componente de periodicidad horaria. Por otro lado, s_h es un componente estocástico, también conocido como residual, que define la dinámica aleatoria. Se utiliza un modelo de regresión lineal múltiple para estimar el componente determinista d_h . Para estimar los componentes estocásticos, este estudio examina cuatro modelos distintos para univariados Análisis de series de tiempo: autorregresivo, autorregresivo no paramétrico, promedio móvil autorregresivo y autorregresivo de red neuronal no lineal. Como resultado, existen cuatro combinaciones posibles a efectos de comparación cuando se combinan modelos deterministas y estocásticos.

2.1.1. Procedimiento para modelar el componente determinista

Esta sección discutirá cómo estimar el componente determinista utilizando un modelo de regresión lineal múltiple. Para lograr esto, modelaremos linealmente la variable de respuesta O_h estimando el componente de tendencia (largo plazo) t_h mediante regresión lineal para el tiempo h . Además, describiremos la periodicidad horaria mediante dummies:

$$n_h = \sum_{i=1}^{24} \zeta_{li,h}$$

A la variable $l_{i,h}$ se le asigna el valor 1 cuando h se refiere a la i -ésima hora del día, y 0 en caso contrario. Los coeficientes de regresión asociados con estos componentes se determinan utilizando el método de mínimos cuadrados ordinarios. Luego de obtener todos los coeficientes de regresión se presenta la ecuación estimada.

$$\hat{d}_h = \hat{\zeta}_0 + \sum_{i=1}^{24} \hat{\zeta}_{li,h}$$

Una vez obtenido el componente determinista estimado, el componente residual o estocástico se puede derivar como:

$$sh = Oh - (\hat{th} + d^{\wedge} h)$$

2.2. Modelando el componente estocástico.

Para estimar el componente estocástico en este trabajo, la serie residual se obtuvo de ambos modelos mediante un modelo de regresión lineal múltiple. Sin embargo, para modelar y pronosticar el componente estocástico, se consideran cuatro modelos diferentes de series de tiempo univariados: modelos autorregresivos, autorregresivos no paramétricos, promedio móvil autorregresivo y modelos autorregresivos de redes neuronales no lineales. Los detalles de estos modelos son los siguientes:

2.2.1. Modelo autorregresivo.

Un proceso autorregresivo (AR) lineal y paramétrico describe la dinámica a corto plazo de sh y considera una combinación lineal del tiempo anterior d observaciones de sh , se denota como:

$$sh = \alpha + \beta_1 sh_{-1} + \beta_2 sh_{-2} + \dots + \beta_n sh_{-n} + \epsilon_d$$

En la fórmula anterior, α es un término de intersección, y $\beta_j (j = 1, 2, \dots, n)$ es el parámetro de pendiente del proceso AR subyacente, y ϵ_d es el término de perturbación. Después de trazar el correlograma (trazar las series ACF y PACF), se encuentra que el mejor ajuste es una serie AR(5) sh .

2.2.2. Modelo de media móvil autorregresiva.

El modelo de media móvil autorregresiva (ARMA) no sólo incorpora los valores pasados de la variable objetivo, sino que también utiliza información importante en forma de media móvil. En nuestro caso, la variable de estudio sd se explica en los n términos anteriores, así como en los valores rezagados de los residuos.

$$sd = \alpha + \beta_1 sd_{-1} + \beta_2 sd_{-2} + \dots + \beta_n sd_{-n} + \epsilon_d + \zeta_1 \epsilon_{d-1} + \zeta_2 \epsilon_{d-2} + \dots + \zeta_m \epsilon_{d-m}$$

En la última ecuación, α denota intercepto $\beta_j (j = 1, 2, \dots, n)$ y $\zeta_k (k = 1, 2, \dots, m)$ son los parámetros del proceso AR y MA respectivamente, y ϵ_d es un Serie de ruido blanco gaussiano con media cero y varianza σ . La selección del orden del modelo ARMA se establece inspeccionando los correlogramas (es decir, ACF y PACF). Ajustamos un modelo ARMA (5, 2) a series temporales de precios de cierre (sd).

2.2.3. Modelo autorregresivo no paramétrico.

La contraparte aditiva no paramétrica del proceso AR conduce al modelo aditivo (NPAR), donde la asociación entre sh y sus términos anteriores no tiene ninguna forma paramétrica específica, permitiendo, probablemente, cualquier tipo de no linealidad que se expresa como:

$$sh = q_1 (sd_{-1}) + q_2 (sd_{-2}) + \dots + q_n (sd_{-n}) + \epsilon_d$$

donde q_j ($j = 1, 2, \dots, n$) muestran funciones de suavizado y describen la asociación entre s_h y sus valores anteriores. En este trabajo, las funciones q_i se denotan mediante splines de regresión cúbica. Al igual que en el caso de la forma paramétrica, se consideran rezagos 5 al estimar el NPAR.

2.2.4. Red neuronal autorregresiva.

Una red neuronal autorregresiva (ANN) es un modelo de aprendizaje automático que predice los valores de las variables de entrada en el futuro. El modelo ANN predice valores futuros de una serie temporal s_d en función de sus observaciones pasadas,

como la función matemática dada por $s_{d-1}, s_{d-2}, \dots, s_{d-n}$. En esta expresión, n es el parámetro de retardo de tiempo. Para reducir el error al cuadrado entre los valores reales y predichos, el modelo ANN se entrena utilizando el método de retropropagación y el enfoque de descenso más pronunciado. En este trabajo se utiliza un diseño ANN (5,2), lo que significa que la capa oculta tiene dos entradas retardadas y cinco nodos.

Además de los modelos mencionados anteriormente, también incluimos dos modelos de referencia, como el modelo Naive y el modelo Seasonal Naive, para evaluar el rendimiento de los modelos de pronóstico basados en componentes propuestos.

Por lo tanto, una vez que se pronostican los componentes determinista y estocástico utilizando los modelos respectivos, los pronósticos finales con uno a siete días de anticipación se derivan como

$$\hat{O}_{h+1} = (\hat{t}_{h+1} + \hat{n}_{h+1} + \hat{s}_{h+1})$$

2.3. Medidas de precisión

Se calcularon seis medidas de precisión estándar diferentes para validar el rendimiento de la técnica de pronóstico y modelado basada en componentes propuesta, incluido el error absoluto medio (MAE), Error porcentual absoluto medio (MAPE), error porcentual absoluto medio simétrico (SMAPE), raíz error cuadrático medio (RMSE), error logarítmico cuadrático medio (RMSLE) y error cuadrático relativo (RRSE). Las fórmulas MAE, MAPE, SMAPE, RMSE, RMSLE y RRSE se muestran a continuación:

$$\begin{aligned}
MAE &= \frac{1}{h} \sum_{h=1}^H |O_h - \hat{O}_h|, \\
MAPA &= \frac{1}{h} \sum_{h=1}^H \frac{O_h - \hat{O}_h}{O_h}, \\
SMAPE &= \frac{1}{h} \sum_{h=1}^H \frac{|O_h - \hat{O}_h|}{(|O_h| + |\hat{O}_h|)/2}, \\
RMSE &= \sqrt{\frac{1}{h} \sum_{h=1}^H (O_h - \hat{O}_h)^2}, \\
RMSLE &= \sqrt{\frac{1}{h} \sum_{h=1}^H (\log(O_h + 1) - \log(\hat{O}_h + 1))^2}, \\
RRSE &= \frac{\frac{1}{h} \sum_{h=1}^H (O_h - \hat{O}_h)^2}{\frac{1}{h} \sum_{h=1}^H (O_h - \bar{O})^2}
\end{aligned}$$

En las ecuaciones anteriores, se observa O_h y $\hat{O}_h$. Se pronostica que los valores de ozono cencertaron para h th observación ($h=1, 2, \dots, 2256=H$).

Además de las medidas de rendimiento de precisión, para evaluar la importancia de las diferencias en el rendimiento de predicción de los modelos propuestos, se realizó la prueba de Diebold-Mariano [34]. La prueba DM es una prueba estadística ampliamente utilizada para comparar predicciones obtenidas de diferentes modelos. La estadística DM se puede determinar mediante la ecuación:

$$DM = \frac{\bar{X}}{\sqrt{\text{Var}(X)}}$$

Donde:

$$\begin{aligned}
\bar{X} &= \frac{1}{h} \sum_{h=1}^H X_h, X_h = (O_h - \hat{O}_h)^2 - (O_h - \hat{O}_{2h})^2, \\
\text{Var}(X) &= \frac{1}{h} \left(2 \sum_{j=1}^{h-1} r_j + r_0 \right), \text{ y } r_j = \text{cov}(X_h - X_{h-j}).
\end{aligned}$$

$\hat{O}_{1h}$ es el valor predicho del primer modelo predictivo y $\hat{O}_{2h}$ es el valor predicho del segundo modelo predictivo en el momento h . Si la estadística DM es negativa, el primer modelo predictivo es estadísticamente mejor que el segundo modelo predictivo.

3. Evaluación y discusión

En este trabajo, utilizamos conjuntos de datos de concentración de ozono horaria de cuatro estaciones de monitoreo en el área metropolitana de Lima: la ATE, la CDM, la SB y la STA, durante tres años consecutivos: 2017, 2018 y 2019, respectivamente. Sólo se consideran los días de invierno de cada año. Como resultado, para una estación hay 6768 puntos de datos. Una sección de entrenamiento (para ajuste del modelo) y una sección de prueba (para pronóstico fuera de muestra). La sección de capacitación comprende datos de los dos primeros años de 2017 a 2018 (4512 horas), mientras que la del año completo de 2019 (2256 horas) se utiliza como muestra fuera de la muestra (pruebas).

Es una práctica común preparar los datos antes de comenzar el proceso de modelado. El propósito del preprocesamiento generalmente es facilitar el modelado de datos. Para hacer esto, la base de datos se ordena, categoriza y evalúa para cada estación de monitoreo teniendo en cuenta la temporada invernal de la ciudad, que se extiende del 21 de junio al 22 de septiembre, para el ozono. Del 2017 al 2019 se propusieron cuatro estaciones de monitoreo en lugares importantes de Lima, capital de Perú. Cabe mencionar que la capital, Lima, cuenta con diez estaciones de monitoreo; sin embargo, cuatro fueron elegidos por falta de datos en el registro. Se utilizó un analizador Teledyne para probar las concentraciones de ozono cada hora. Las pruebas de cero y span, la calibración y la detección de fugas son ejemplos de actividades del analizador. Luego de corregir ceros, duplicados y/o anomalías, los datos se transmiten vía telemetría al SENAMHI (Servicio Nacional de Meteorología e Hidrología del Perú) para su certificación. De igual forma, el SENAMHI cuenta con una red sistemática de estaciones que monitorean y reportan las variables investigadas a un centro de procesamiento de manera periódica y automatizada. Cada hora, estas estaciones emplean instrumentación y sensores de alta calidad para detectar la temperatura, la humedad relativa, la velocidad y la dirección del viento. Además, se utilizó un enfoque inductivo conocido como Imputación Múltiple por Ecuaciones Encadenadas. Este enfoque se basa en una especificación completamente condicional, con cada variable incompleta dada por su modelo [33].

Para obtener los pronósticos de concentración de O₃ con un día de anticipación, dos días de anticipación, tres días de anticipación y siete días de anticipación utilizando la metodología basada en componentes propuesta para el pronóstico de series de tiempo presentada en la Sección 2 para todas las estaciones de monitoreo consideradas, se debe realizar lo siguiente. Se deben seguir varios pasos: Primero, para estabilizar la varianza de la serie temporal de concentración de O₃, se aplicó la transformación logarítmica natural. En segundo lugar, dividimos la serie temporal de concentración horaria de O₃ en dos nuevos componentes: determinista y estocástico. El componente determinista contiene un componente lineal de tendencia larga y un componente estacional por hora, mientras que el estocástico es el resto. Modelar la parte determinista utilizando un modelo de regresión lineal múltiple y el componente estocástico con varios modelos de series de tiempo discutidos en la última sección. Finalmente, combine los pronósticos de ambos componentes y obtenga los resultados finales del pronóstico para cada modelo de combinación posible. Por lo tanto, los pronósticos de un día de anticipación, dos días de anticipación, tres días de anticipación y siete días de anticipación se obtuvieron utilizando la técnica de ventana en expansión durante 94 días (2256 horas), y los modelos se estimaron en consecuencia. Asimismo, los pronósticos de concentración de ozono se lograron a través de la ecuación 2.9. Las medidas de desempeño, incluidas MAE, MAPE, SMAPE, RRAE, RMSLE y RMSE, se utilizan luego para la evaluación y el desempeño comparativo de los modelos. Por lo tanto, el detalle de los resultados de cuatro estaciones de monitoreo: la Ate (en el Cuadro 1), el Campo de Marte (en el Cuadro 2), la San Borja (en el Cuadro 3) y la Santa Anita (en el Cuadro 4), todas ubicadas en Lima Metropolitana, Perú.

Tabla 1. Concentración de ozono (μgm^3) en la estación ATE: error de pronóstico medio para todos los modelos para un pronóstico fuera de muestra con veinticuatro horas de antelación.

Con un día de antelación (24 horas de antelación)								
MODELO MAE SMAPE MAPE RRSE RMSLE RMSE								
ARIMA	2.997	0,212	0,218	0,192	1.113	0,247	4.010	
NPAR 2.692 ARMA		0,194	0,199	0,200	1.022	0,229	3.685	
2.790 NNA 3.276		0,224	0,228	0,219	1.043	0,233	3.760	
NATIVE 3.181		0,231	0,277	0,289	Con	1.464	0,286	5.276
SNAIVE 3.829		dos días de antelación			1.219	0,263	4.394	
		(48 horas de antelación)			1.391	0,326	5.012	
Con tres días de antelación (72 horas de antelación)								
MODELO MAE MAPE SMAPE RRSE RMSLE RMSE								
ARIMA	3.132	0,231	0,224	1,174	0,262	4.192		
NPAR 2.699 ARMA		0,194	0,190	1,073	0,237	3.832		
2.823 NNA 3.452		0,202	0,203	1,078	0,239	3.849		
NATIVE 3.467		0,245	0,230	1,578	0,308	5.635		
SNAIVE 3.792		0,255	0,237	1,360	0,288	4.855		
		0,287	0,275	1,392	0,324	4.971		
Con tres días de antelación (72 horas de antelación)								
MODELO MAE MAPE SMAPE RRSE RMSLE RMSE								
ARIMA	3.112	0,241	0,231	1,219	0,326	4.361		
NPAR 2.944 ARMA		0,218	0,217	1,185	0,316	4.237		
2.879 NNA 3.940		0,216	0,218	1,116	0,309	3.990		
NATIVE 3.769		0,294	0,263	1,812	0,391	6.479		
SNAIVE 3.853		0,297	0,265	1,523	0,371	5.446		
		0,292	0,288	1,415	0,372	5.060		
Siete días por delante (168 horas por delante)								
MODELO MAE MAPE SMAPE RRSE RMSLE RMSE								
ARIMA	3.416	0,258	0,268	1,207	0,342	4.321		
NPAR 2.900 ARMA		0,205	0,223	1,117	0,309	3.999		
2.971 NNA 3.147		0,217	0,228	1,122	0,310	4.018		
NATIVE 4.043		0,212	0,243	1,213	0,327	4.343		
SNAIVE 3.826		0,309	0,320	1,415	0,402	5.065		
		0,290	0,287	1,404	0,370	5.027		

Los resultados del ATE se enumeran en la Tabla 1. Esta tabla muestra los errores medios de precisión de la cuatro horizontes, como un día, dos días, tres días y siete días de anticipación, de los seis modelos: cuatro modelos combinados desde dentro de la técnica de pronóstico basada en componentes propuesta: el AR, el Modelos NPAR, ARMA y NNA; y dos modelos de referencia: el ingenuo y el ingenuo estacional modelos.

Las siguientes dos conclusiones se extrajeron de la Tabla 1. Los errores de precisión medios de la El modelo NPAR fue mínimo. Como se muestra en la Tabla 1, el modelo NPAR tuvo el mejor efecto de pronóstico, con exactitud los errores medios (MAE, MAPE, SMAPE, RRSE, RMALSE y RMSE) del día (2,692, 0,192, 0,194, 1,022, 0,229 y 3,685), dos días (2,879, 0,216, 0,217, 1,116, 0,309 y 3,990), y pronósticos anticipados de siete días (2.900, 0.205, 0.223, 1.117, 0.309 y 3.999), respectivamente, menos que los modelos AR, ARMA y NNA dentro de la metodología de pronóstico propuesta, y también significativamente mínimo con respecto a los modelos de referencia (los modelos ingenuo y ingenuo estacional). El predictivo el efecto del modelo NNA fue el peor y fue mucho mayor que los errores medios del AR y modelos ARMA. Sin embargo, sólo en el caso de una precisión de pronóstico de tres días los errores medios son los mejores resultados mostrados por los modelos ARMA con las siguientes métricas: MAE = 2.879, MAPE = 0.216, SMAPE = 0,217, RRAE = 1,116, RMSLE = 0,309 y RMSE = 3,990. Aunque el modelo NPAR también muestra los segundos mejores resultados. Por otro lado, al comparar el mejor modelo dentro de la propuesta enfoque de pronóstico con los modelos de referencia (modelos ingenuos y modelos ingenuos estacionales), se confirma a partir de Tabla 1 que los modelos de referencia son superados por el modelo NPAR. Esto indicó que el O3 La concentración se ve afectada en gran medida por los valores a corto plazo más que por los valores a largo plazo. Además de arriba, el error de predicción de un día fue mínimo en comparación con el de los otros tres horizontes.

Tomando el modelo NPAR como ejemplo, el MAE, el MAPE, el SMAPE, el RRAE, el RMSLE y el RMSE de la predicción de un día fue 2,692, 0,192, 0,194, 1,022, 0,229 y 3,685, menos que 2,879, 0,216, 0,217, 1,116, 0,309 y 3,990 para la predicción de dos días, 2,944, 0,218, 0,217, 1,185, 0,316 y 3,987 para la predicción de tres días, y 2,900, 0,205, 0,223, 1,117, 0,309 y 3,999 para la predicción de tres días. la predicción de siete días. Los efectos predictivos de horizonte de los modelos AR, ARMA y NNA fueron los mismos que el modelo NPAR, lo que indica que cuanto más corto sea el horizonte predictivo del modelo, mayor mejor el efecto predictivo. Cuanto más largo sea el horizonte predictivo, peor será el efecto predictivo.

De este modo, se puede ver a partir de estos resultados que el error predictivo del modelo NPAR fue el más pequeño, y el efecto predictivo fue el mejor en comparación con el resto dentro de los modelos de pronóstico propuestos y los modelos de base. Además, esto también indicó que la información reciente fue más efectiva en pronosticar la concentración de ozono que la información anterior después de comparar los errores de predicción de los cuatro horizontes de los tres modelos.

Tabla 2. Concentración de ozono (μgm^3) en la Estación Campo de Marte: El error medio de pronóstico para todos los modelos para un pronóstico fuera de muestra con veinticuatro horas de antelación.

Con un día de antelación (24 horas de antelación)						
MODELO	MAE	SMAPE	MAPE	RRSE	RMSLE	RMSE
Arkansas	4.6696		0,1171	0,1303	0,5597	0,1343
NPAR 4.752			0,1182	0,6667	0,1463	0,6295
ARMA			0,1274	0,1382	0,645	0,1322
5.2546 NNA 5.4503			0,131	0,6035	0,4033	1,3213
NATIVE			0,1245			
5.1361 SNAIVE				0,2066		7.8429
10.4678				0,1985		7.3385
			0,2587	0,5135		16.0659
Con dos días de antelación (48 horas de antelación)						
MODELO	MAE	SMAPE	MAPE	RRSE	RMSLE	RMSE
Arkansas	5.909		0,198	0,158	0,732	0,301
NPAR 6.009			0,201	0,166	0,714	0,306
ARMA			0,213	0,169	0,787	0,315
6.485 NNA 6.688			0,205	0,175	0,788	0,307
NATIVE			0,203	0,176	0,777	0,303
6.737 SNAIVE			0,397	0,256	1,275	0,509
10.313						15.907
Con tres días de antelación (72 horas de antelación)						
MODELO	MAE	SMAPE	MAPE	RRSE	RMSLE	RMSE
AR 6.489			0,220	0,178	0,795	0,352
NPAR 6.695			0,231	0,184	0,808	0,358
ARMA			0,220	0,178	0,793	0,352
6.485 NNA 7.794			0,244	0,210	0,877	0,372
NATIVE			0,228	0,191	0,833	0,362
7.064 SNAIVE			0,367	0,264	1,178	0,510
10.011						15.309
Siete días por delante (168 horas por delante)						
MODELO	MAE	SMAPE	MAPE	RRSE	RMSLE	RMSE
Arkansas	7.171		0,294	0,201	11,016	0,801
NPAR 7.288			0,295	0,204	11,021	0,801
ARMA			0,285	0,201	10,943	0,796
7.169 NNA 9.913			0,368	0,269	15,171	1,103
NATIVE			0,356	0,294	14,588	1,061
10.963 SNAIVE			0,329	0,274	13,384	0,973
10.156						0,434

De la misma manera, la Tabla 2 contiene los hallazgos para la Estación Campo de Marte. La exactitud significa fueron los mismos que el modelo NPAR, lo que indica que cuanto más corto sea el horizonte predictivo del modelo, mayor mejor el efecto predictivo. Cuanto más largo sea el horizonte predictivo, peor será el efecto predictivo. De este modo, 0,130, 0,560, 0,107 y 6,927), dos días (5,909, 0,198, 0,158, 0,732, 0,301 y 8,905). Sin embargo, Los errores medios son los mejores resultados mostrados por los modelos ARMA en términos de precisión de pronóstico de tres y siete

días, con las siguientes métricas: tres días (6,485, 0,220, 0,178, 0,793, 0,352 y 10,304) y siete. pronósticos anticipados de varios días (7,169, 0,285, 0,201, 0,796, 0,395 y 10,943), respectivamente, que son menores que los modelos NPAR y NNA dentro de la metodología de pronóstico propuesta y también significativamente menores que los modelos de referencia. El modelo NNA tuvo el efecto predictivo más débil y tuvo errores medios sustancialmente mayores que el modelo NPAR. A pesar de esto, el modelo NPAR produce los terceros mejores resultados. Al comparar el mejor modelo dentro de la técnica de pronóstico propuesta con los modelos de referencia (modelos ingenuos y sin estacionalidad), la Tabla 2 muestra que el modelo AR supera a los modelos de referencia en pronósticos de uno y dos días, mientras que el modelo ARMA supera al modelo de referencia. modelos en pronósticos de tres y siete días. Por lo tanto, una vez más, la investigación demostró que los valores a corto plazo tienen un mayor impacto en la concentración de O₃ que los valores a largo plazo. Junto con lo mencionado anteriormente, la inexactitud del pronóstico de un día fue modesta en comparación con los otros tres períodos de tiempo. Tomando el modelo AR como ejemplo, el MAE, el MAPE, el SMAPE, el RRAE, el RMSLE y el RMSE de la predicción de un día fueron 4,670, 0,117, 0,130, 0,560, 0,107 y 6,927, menos que 5,909. 0,198, 0,158, 0,732, 0,301 y 8,905 para la predicción de dos días, 6,489, 0,220, 0,178, 0,795, 0,352 y 10,333 para la predicción de tres días, y 7,171, 0,294, 0,201, 0,801, 0. 398 y 11.016 para la predicción de siete días. Los modelos AR, ARMA y NNA tuvieron todos los mismos resultados predictivos de horizonte que el modelo NPAR, lo que sugiere que cuanto más corto sea el horizonte de predicción del modelo, mejor será el impacto predictivo. Cuanto más largo sea el horizonte previsto, menos precisa será la predicción. Por lo tanto, el error de predicción de los modelos AR y ARMA fue el menor y el impacto predicho fue el mejor en comparación con el resto de los modelos de pronóstico y modelos de referencia sugeridos. Además, la comparación de los errores predichos de los cuatro horizontes de los tres modelos reveló que la información reciente fue más útil para estimar los niveles de ozono que la información antigua.

Tabla 3. Concentración de ozono (μgm^3) en la Estación San Borja: El error de pronóstico medio para todos modelos para un pronóstico fuera de muestra con veinticuatro horas de antelación.

Con un día de antelación (24 horas de antelación)						
MODELO	MAE	MAPE	SMAPE	RRSE	RMSLE	RMSE
Arkansas	3.3365	0.2689	0.3442	0.6845	NPAR 3.1904	0.2505
					0,3543	4.1535
0.2997	0.6502	ARMA 3.2937	0.2658	0.3273	0.6808	NNA 3.2018
					0,3517	3.9458
0.2601	0.3643	0.653	NATIVE 3. 3159	0,2635	0,303	0,6809
					0,3553	4.1312
5,8544	0,4466	0,78	1,1846	Con dos días de antelación (48 horas de antelación)		
					0,3732	3.9624
					0,3526	4.1318
					0,6638	7.1885
Con tres días de antelación (72 horas de antelación)						
MODELO	MAE	MAPE	SMAPE	RRSE	RMSLE	RMSE
Arkansas	3.987	0,432	0,313	0,827	0,412	4.969
NPAR 3.758	ARMA	0,440	0,294	0,765	0,404	4.596
3.879	NNA 3.708	0,417	0,303	0,800	0,395	4.807
	NATIVE 4.007	SNAIVE	0,381	0,288	0,760	0,358
						4.569
5.781		0,389	0,311	0,829	0,410	4.980
		0,767	0,440	1,185	0,657	7.122
Con tres días de antelación (72 horas de antelación)						
MODELO	MAE	MAPE	SMAPE	RRSE	RMSLE	RMSE
Arkansas	4.000	0,486	0,316	0,861	0,466	5.167
NPAR 3.920	ARMA	0,397	0,299	0,824	0,442	4.950
3.923	NNA 3.961	0,399	0,305	0,836	0,453	5.022
	NATIVE SNAIVE	0,521	0,308	0,831	0,473	4.988
		4.043	0,432	0,316	0,874	0,458
		5.875	0,767	0,453	1,213	0,692
						7.284
Siete días por delante (168 horas por delante)						
MODELO	MAE	MAPE	SMAPE	RRSE	RMSLE	RMSE
Arkansas	4.975	0,468	0,404	1,054	0,552	6.276
NPAR 4.578	ARMA	0,625	0,350	0,940	0,517	5.599
5.020	NNA 4.491	0,594	0,393	1,043	0,514	6.207
	NATIVE 5.151	SNAIVE	0,442	0,341	0,928	0,505
						5.527
5.773		0,598	0,400	1,068	0,545	6.356
		0,753	0,444	1,211	0,685	7.211

Asimismo, los resultados preliminares para la Estación San Borja se presentan en la Tabla 3. La precisión media errores de los seis modelos en cuatro horizontes, como un día, dos días, tres días y siete días de anticipación se presentan en la Tabla 3: cuatro dentro de la técnica de pronóstico basada en componentes propuesta: el Modelos AR, NPAR, ARMA y NNA; y dos modelos de referencia: el modelo ingenuo y el modelo ingenuo estacional.

La Tabla 3 arrojó los siguientes dos hallazgos. Los errores medios de precisión de los modelos NPAR y NNA. son mínimos. Como se indica en la Tabla 3, el modelo NPAR tuvo el mejor efecto de pronóstico para un día y pronóstico con tres días de anticipación, con errores medios de precisión (MAE, MAPE, SMAPE, RRSE, RMALSE, y RMSE) del pronóstico de un día (3.190, 0.251, 0.300, 0.650, 0.352 y 3.946) y de los pronósticos de tres días (3.920, 0.397, 0.299, 0.824, 0.442 y 4.950), mientras que el modelo NNA tenía el mejor efecto de pronóstico para pronósticos de dos y siete días de antelación, con los siguientes errores medios de dos días (3,708, 0,381, 0,288, 0,760, 0,358 y 4,569) y de siete días (4,491, 0,442, 0,341, 0,928, 0,505 y 5,527) pronósticos futuros, respectivamente, menores que los modelos AR y ARMA dentro del metodología de pronóstico propuesta, y también significativamente mínima para los modelos de referencia. La RA El modelo tuvo el efecto predictivo más pobre, con errores medios sustancialmente mayores que el modelo ARMA.

Aunque el modelo NPAR produce el segundo mejor rendimiento, comparar el mejor modelo dentro a técnica de pronóstico propuesta a los modelos de referencia (modelos ingenuos y modelos ingenuos estacionales), Tabla 5 muestra que el modelo NPAR supera a los modelos de referencia. Esto sugirió que los valores a corto plazo tuvo un mayor impacto en la concentración de O₃ que los valores a largo plazo. Sumado al mismo, el día El error de pronóstico fue pequeño en comparación con los otros tres períodos de pronóstico. Tomando el NPAR modelo como ejemplo, el MAE, el

MAPE, el SMAPE, el RRAE, el RMSLE y el RMSE de la predicción de un día fueron 3,190, 0,251, 0,300, 0,650, 0,352 y 3,946, menos que 3,758, 0,440, 0,294, 0,765, 0,404 y 4,596 para la predicción de dos días, 3,920, 0,397, 0,299, 0,824, 0,442 y 4,950 para la a predicción de tres días y 4,578, 0,625, 0,350, 0,940, 0,517 y 5,599 para la predicción de siete días.

Los modelos AR, ARMA y NNA tuvieron los mismos resultados predictivos de horizonte que el modelo NPAR, lo que sugiere que cuanto más corto sea el horizonte de predicción del modelo, mejor será el impacto predictivo. Mientras más largo sea el horizonte del pronóstico, menos preciso parece ser el pronóstico. Así, se puede observar desde estos datos que el modelo NPAR tuvo el error mínimo previsto y el mejor impacto predictivo cuando en comparación con los otros modelos de pronóstico y modelos de referencia propuestos. Además, comparar Los errores pronosticados de los cuatro horizontes de los tres modelos revelaron que la información reciente era más más útil para estimar los niveles de ozono que la información histórica.

Tabla 4. Concentración de ozono (μgm^3) en la Estación Santa Anita: El error de pronóstico medio para todos modelos para un pronóstico fuera de muestra con veinticuatro horas de antelación.

Con un día de antelación (24 horas de antelación)							
MODELO MAE MAPE SMAPE RRSE RMSLE RMSE							
Arkansas	3.7604	0,361	0,4296	1,0933	0,4317	4.8461	
NPAR 3.1036		0,296	0,3415	0,955	0,3666	4.2332	
ARMA 3,2337	0,3096	NNA 3,1828	0,378	0,9786	0,3856	4.3378	
0,3064	NATIVE 3,925	0,3692	SNAIVE	0,3529	0,9679	0,3765	4.2901
5,2327	0,4827		0,4319	1,1339	0,4408	5.0262	
			0,6442	1,4868	0,6032	6.5903	
Con dos días de antelación (48 horas de antelación)							
MODELO MAE MAPE SMAPE RRSE RMSLE RMSE							
AR 3.890	NPAR	0,446	0,369	1,131	0,447	5.054	
3.235	ARMA 3.313	0,344	0,305	0,956	0,363	4.269	
NNA 3.252	NATIVE	0,384	0,314	0,977	0,387	4.366	
4.234	SNAIVE 5.199	0,349	0,308	0,964	0,371	4.306	
		0,474	0,389	1,237	0,471	5.526	
		0,635	0,478	1,466	0,598	6.549	
Con tres días de antelación (72 horas de antelación)							
MODELO MAE MAPE SMAPE RRSE RMSLE RMSE							
Arkansas	4.162	0,488	0,405	1,166	0,492	5.243	
NPAR 3.584	ARMA	0,385	0,349	1,022	0,415	4.595	
3.597	NNA 3.522	0,427	0,348	1,009	0,425	4.538	
NATIVE 4.484	SNAIVE	0,384	0,344	1,003	0,411	4.511	
5.195		0,510	0,418	1,260	0,505	5.666	
		0,636	0,486	1,454	0,610	6.537	
Siete días por delante (168 horas por delante)							
MODELO MAE MAPE SMAPE RRSE RMSLE RMSE							
Arkansas	4.694	0,533	0,443	1,326	0,534	5.905	
NPAR 3.565	ARMA	0,373	0,346	1,042	0,413	4.640	
3.705	NNA 3.577	0,445	0,354	1,071	0,441	4.768	
NATIVE 5.472	SNAIVE	0,381	0,348	1,052	0,420	4.683	
5.129		0,625	0,478	1,532	0,574	6.819	
		0,626	0,480	1,456	0,604	6.479	

Por el contrario, la Tabla 4 tabuló los resultados para la Estación Santa Anita. Esta tabla ilustra la exactitud errores medios de los cuatro horizontes, como un día, dos días, tres días y siete días adelante, de los seis modelos: cuatro dentro de la técnica de pronóstico basada en componentes propuesta: el Modelos AR, NPAR, ARMA y NNA; y dos modelos de referencia: el ingenuo y el estacional modelo ingenuo. Las dos conclusiones siguientes se extrajeron de la Tabla 4. Los errores medios de precisión del modelo NPAR fueron mínimos. Como se muestra en la Tabla 4, el modelo NPAR tuvo el mejor efecto de pronóstico, con exactitud los errores medios (MAE, MAPE, SMAPE, RRSE, RMALSE y RMSE) del día (3,104, 0,296, 0,342, 0,955, 0,367 y 4,233), dos días (3,235, 0,344, 0,305, 0,956, 0,363 y 4,269), y pronósticos anticipados de siete días (3,565, 0,373, 0,346, 1,042, 0,413 y 4,640), respectivamente, menos que los modelos AR, ARMA y NNA dentro de la

metodología de pronóstico propuesta, y también significativamente mínimo con respecto a los modelos de referencia (los modelos ingenuo y ingenuo estacional).

El efecto del modelo NNA fue el peor y fue mucho mayor que los errores medios del AR y Modelos ARMA. Sin embargo, sólo en el caso de una precisión de pronóstico de tres días los errores medios son los mejores resultados mostrados por los modelos NNA con las siguientes métricas: MAE = 3.522, MAPE = 0.384, SMAPE = 0,344, RRSE = 1,003, RMSLE = 0,411 y RMSE = 4,511. Aunque el modelo NPAR también muestra los segundos mejores resultados. Por otro lado, comparar el mejor modelo dentro del pronóstico propuesto enfoque con los modelos de referencia (modelos ingenuos y modelos ingenuos por temporada), se confirma en la Tabla 3 que Los modelos de referencia son superados por el modelo NPAR. Esto indicó que la concentración de O₃ se ve más afectado por los valores a corto plazo que por los valores a largo plazo. Junto con lo anterior, el error de pronóstico de un día fue pequeño en comparación con los otros tres tramos. El MAE, MAPE, SMAPE, RRAE, RMSLE y RMSE de la predicción de un día para el modelo AR fueron 3,7604, 0,361, 0,4296, 1,0933, 0,4317 y 4,8461, que eran inferiores a 3,890, 0,446, 0,369, 1,131, 0,447 y 5,054 para la predicción de dos días, 4.162, 0.488, 0.405, 1.166, 0.492 y 5. NPAR, ARMA y NNA.

Todos los modelos tuvieron los mismos efectos predictivos de horizonte que el modelo AR, lo que demuestra Cuanto mayor sea el horizonte de predicción del modelo, mejor será el impacto predictivo. Cuanto más largo sea el horizonte de previsión, menos preciso será el pronóstico. Por lo tanto, se puede observar a partir de estos datos que el modelo NPAR tenía el error mínimo previsto y el mejor impacto predictivo en comparación con los demás propuestos Modelos de previsión y modelos de referencia. Además, al comparar los errores previstos de los cuatro horizontes de los tres modelos revelaron que la información reciente fue más útil para estimar el ozono.

4. Referencias.

1. Ghorani-Azam, A., Riahi-Zanjani, B. y Balali-Mood, M. Efectos de la contaminación del aire en la salud humana y medidas prácticas para la prevención en Irán. *Revista de investigaciones en ciencias médicas: la oficial. Revista de la Universidad de Ciencias Médicas de Isfahan* 2016, 21.
2. Hailstone, J. Hailstone, J. <https://www.forbes.com/sites/jamiehailstone/2023/03/07/nearly-nowhere-on-earth-safe-from-air-pollution-study-finds/?sh=1e0d9fd9da1d/> (consultado el 25 de julio de 2023).
3. Ordoñez, C., Garrido-Perez, JM, & García-Herrera, R. Ozono cercano a la superficie a principios de primavera en Europa Durante el cierre por COVID-19: los efectos meteorológicos superan los cambios en las emisiones. *ciencia de la entomo total* 2020, 747, 141322.
4. Mostafa, MK, Gamal, G. y Wafiq, A. El impacto de COVID 19 en los niveles de contaminación del aire y otros Indicadores ambientales: un estudio de caso de Egipto. *Revista de gestión ambiental* 2021, 277, 111496.
5. Gagliardi, RV y Andenna, C. Un enfoque de aprendizaje automático para investigar el ozono superficial comportamiento. *Atmósfera* 2020, 11(11), 1173.
6. Jaffe, DA, Cooper, OR, Fiore, AM, Henderson, BH, Tonnesen, GS, Russell, AG, Henze, DK, Langford, AO, Lin, M. & Moore, T. Evaluación científica del ozono de fondo en el EE. UU.: Implicaciones para la gestión de la calidad del aire. *Elem Sci Anth* 2018, 6, 56.
7. Lu, H., Lyu, X., Cheng, H., Ling, Z. y Guo, H. Descripción general de las características espacio-temporales del régimen de formación de ozono en China. *Ciencias ambientales: procesos e impactos* 2019, 21 (6), 916-929.
8. Ka"ffer, MI, Domingos, M., Lieske, I. y Vargas, VM Predicción de los niveles de ozono a partir del clima. Parámetros y características foliares de la variedad de tabaco Bel-W3. *Contaminación Ambiental* 2019, 248, 471-477.
9. Li, Y., Xue, Y., Guang, J., She, L., Fan, C. y Chen, G. PM2 a nivel del suelo. 5 estimación de concentración a partir de datos satelitales en el área de Beijing utilizando una masa de extinción de enjambre de partículas específica Algoritmo de conversión. *Teledetección* 2018, 10(12), 1906.

10. Velasco, E., & Retama, A. La amenaza del ozono contraataca a la Ciudad de México. *Ciudades y sociedad sostenibles* 2017, 31, 260-263.
11. Carbo-Bustanza, N., Belmonte, M., Jiménez, V., Montalban, P., Rivera, M., Martínez, FG, Mo-hamed, MMH, De La Cruz, ARH, da Costa, K. & Lopez-Gonzales, JL Un enfoque de aprendizaje automático para analizar la concentración de ozono en el área metropolitana de Lima, Perú. *Informes científicos* 2022, 12(1), 22084.
12. Cohen, AJ, Brauer, M., Burnett, R., Anderson, HR, Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R. & Feigin, V. Estimaciones y tendencias a 25 años de la carga global de enfermedades atribuibles a la contaminación del aire ambiente: un análisis de datos del Estudio de la carga global de enfermedades 2015. *The lancet* 2017, 389 (10082), 1907-1918.
13. Jakovlev, AR y Smyshlyaev, SP Simulación numérica de los efectos de los océanos mundiales sobre la temperatura y el ozono en la atmósfera media y baja. *Meteorología e Hidrología de Rusia* 2019, 44, 594-602.
14. Gaudel, A., Cooper, OR, Ancellet, G., Barret, B., Boynard, A., Burrows, JP, Clerbaux, C., Co-heur, PF, Cuesta, J., Cuevas, E. & Doniki, S. Informe de evaluación: Distribución actual y tendencias del ozono troposférico relevantes para el clima y la evaluación del modelo de química atmosférica global. *Elem Sci Anth* 2018, 6, 39.
15. Rodríguez-Urrego, D., & Rodríguez-Urrego, L. Calidad del aire durante el COVID-19: PM2. 5 análisis en las 50 capitales más contaminadas del mundo. *Contaminación Ambiental* 2020, 266, 115042.
16. Rybarczyk, Y. y Zalakeviciute, R. Enfoques de aprendizaje automático para el modelado de la calidad del aire exterior: Una revisión sistemática. *Ciencias Aplicadas* 2018, 8(12), 2570.
17. Comrie, AC Comparación de redes neuronales y modelos de regresión para el pronóstico del ozono. *Diario de la Asociación de Gestión del Aire y Residuos* 1997, 47(6), 653-663.
18. Harrou, F., Fillatre, L., Bobbia, M. y Nikiforov, I. Detección estadística de mediciones anormales de ozono basada en una prueba de índice de probabilidad generalizada restringida. En la 52.ª Conferencia del IEEE sobre Decisión y Control, Florencia, Italia, del 10 al 13 de diciembre de 2013.
19. Dueñas, C., Fernández, MC, Cañete, S., Carretero, J., & Liger, E. Modelo estocástico para pronosticar la concentración de ozono a nivel del suelo en áreas urbanas y rurales. *Quimiosfera* 2005, 61(10), 1379-1389.
20. Petetin, H., Bowdalo, D., Soret, A., Guevara, M., Jorba, O., Serradell, K., & Pérez García-Pando, C. Meteorología- impacto normalizado del COVID- 19 confinamiento por la contaminación por NO₂ en España. *Química y Física Atmosférica* 2020, 20(18), 11119-11141.
21. Aljanabi, M., Shkoukani, M. y Hijawi, M. Predicción del ozono a nivel del suelo mediante técnicas de aprendizaje automático: un estudio de caso en Ammán, Jordania. *Revista Internacional de Automatización y Computación* 2020, 17, 667-677.
22. Sousa, SIV, Martins, FG, Alvim-Ferraz, MC, & Pereira, MC Regresión lineal múltiple y redes neuronales artificiales basadas en componentes principales para predecir concentraciones de ozono. *Software y modelado ambiental* 2007, 22(1), 97-103.
23. Chelani, AB Predicción de la concentración máxima diaria de ozono terrestre utilizando una máquina de vectores de soporte. *Monitoreo y evaluación ambiental* 2010, 162, 169-176.

24. Ren, X., Mi, Z. y Georgopoulos, PG Comparación del aprendizaje automático y la regresión del uso de la tierra para la estimación espaciotemporal a escala fina de la contaminación del aire ambiente: modelado de concentraciones de ozono en los Estados Unidos contiguos. *Medio ambiente internacional* 2020, 142, 105827.
25. Yafouz, A., AlDahoul, N., Birima, AH, Ahmed, AN, Sherif, M., Sefelnasr, A., Allawi, MF & Elshafie, A. Comparación completa de varios algoritmos de aprendizaje automático para la predicción de la concentración de ozono a corto plazo. *Revista de ingeniería de Alejandría* 2022, 61(6), 4607-4622.
26. Pan, Q., Harrou, F. y Sun, Y. Una comparación de métodos de aprendizaje automático para la predicción de la contaminación por ozono. *Revista de Big Data* 2023, 10(1), 63.
27. Iftikhar, H.; Zafar, A.; Turpo-Chaparro, JE; Rodríguez, PC; López-Gonzales, JL Pronóstico de los precios del petróleo crudo Brent para el día siguiente utilizando combinaciones híbridas de modelos de series temporales. *Matemáticas* 2023, 16, 3548.
28. Iftikhar, H.; Bibi, N.; Cañas Rodrigues, P.; López-Gonzales, JL Nuevas técnicas de descomposición múltiple para el pronóstico de series temporales: aplicación al pronóstico mensual del consumo de electricidad en Pakistán. *Energías* 2023, 16, 2579.
29. Shah, I.; Iftikhar, H.; Ali, S. Modelización y previsión de la demanda y los precios de la electricidad: una comparación de enfoques alternativos. *J. Matemáticas*. 2022, 2022, 3581037.
30. Shah, I.; Iftikhar, H.; Ali, S. Modelado y previsión del consumo de electricidad a medio plazo utilizando técnica de estimación de componentes. *Previsión para* 2020, 2(2), 163–179.
31. Iftikhar, H. Modelado y pronóstico de series temporales complejas: un caso de demanda de electricidad. Tesis de maestría, Universidad Quaidi-Azam, Islamabad, Pakistán, 2018, 1–94.
32. Shah, I.; Iftikhar, H.; Ali, S.; Wang, D. Previsión de la demanda de electricidad a corto plazo utilizando la técnica de estimación de componentes. *Energías* 2019, 12, 2532.
33. Van Buuren, S.; Oudshoorn, CG Imputación multivariada mediante ecuaciones encadenadas, 2000.
34. Diebold, FX; Mariano, RS Comparación de la precisión predictiva. *J. Autobús. Economía. Estadística*. 2002, 20, 134–144.
35. Dickey, fiscal del distrito; Fuller, WA Distribución de los estimadores para series temporales autorregresivas con raíz unitaria. *J. Am.Stat. Asociación*. 1979, 74, 427–431.
36. Romero, Y.; Díaz, C.; Meldrum, I.; Velásquez, RA; Noel, J. Análisis temporal y espacial de contaminantes relacionados con el tráfico bajo la influencia de la estacionalidad y variables meteorológicas en una ciudad urbana del Perú. *Helio* 2020.
37. León, CAM; Félix, MFM; Olivera, CAC Influencia del Confinamiento Social por COVID-19 en la Calidad del Aire en el Distrito de San 503 Juan de Lurigancho en Lima, Perú. *Transacciones de Ingeniería Química* 2022, 91, 475–480.

5. Figuras y Tablas

Tabla 5. Resultados de Diebold y Marino para la estación ATE: los valores estadísticos de DM para todos Los modelos se dan en la Tabla 1.

Con un día de antelación (24 horas de antelación)									
Modelos	AR	NPAR	ARMA	NNA	NA'VE	SNA'VE			
ARMA	0.0000	-9,0887	-9,0935	-8,8935	-9,0899	5.6142			
NPAR 9.0887 ARMA	0,0000	2,8016	7,8766	7,0433	7,9687				
9.0935 NNA 8.8935	-2,8016	0,0000	7,8831	6,9920	7,9692				
NA'VE 9.0899 SNA'VE	-7,8766	-7,8831	0,0000	-7,8895	7,9761				
-5.6142	-7,0433	-6,9920	7,8895	0,0000	7,9698				
	-7,9687	-7,9692	-7,9761	-7,9698	0,0000				
Con dos días de antelación (48 horas de antelación)									
Modelos	AR	NPAR	ARMA	NNA	NA'VE	SNA'VE			
AR 0.0000 NPAR	-9,5751	-9,5679	-8,5622	0,0000	30,6075	5,5041 -7,6242	5.3401		
9.5751 ARMA 9.5682	-7,0475	-3,0507	7,6242	-7,6183	7,7160		7.7160		
NNA 9.4378 NA'VE	0,0000	-7,7160	7,6015	7,7141	7,7234		7.7141		
9.5679 SNA'VE -5.3401		-7,6015	0,0000	7,7156	0,0000		7.7234		
		-6.5041	7,6183	7,7156			7.7156		
		-7,7141	-7,7234	0,0000			0.0000		
Con tres días de antelación (72 horas)									
Modelos	AR	NPAR	ARMA	NNA	NA'VE	SNA'VE			
ARMA	0.0000	-8,7263	-8,7151	-9,0370	-8,6979	4.5748			
NPAR	8.7263	0,0000	-2,0191	6,1792	18,9793	6.2607			
ARMA 8.7151 NNA	2,0191	0,0000	6,1427	11,0496	6.2578				
9.0370 NA'VE 8.6979	-6,1792	-6,1427	0,0000	-6,0526	6.2674				
SNA'VE -4.5748	-18,9793	-11,0496	6,0526	0,0000	6.2511				
	-6.2607	-6.2578	-6.2674	-6.2511	0,0000				
Con cuatro días de antelación (96 horas de antelación)									
Modelos	AR	NPAR	ARMA	NNA	NA'VE	SNA'VE			
AR 0.0000 NPAR	-8,8059	-8,4630	-8,7729	-8,8052	0,0000	11,6282	19,9019	14.7468	
8.7952 ARMA 8.7952	8,7952	-2,9659	11,5031	19,9553	11,5031	11,6283	-11,5031	0,0000	11.3933
8.7952 ARMA 8.7952	8,7952	-2,9659	11,5031	19,9553	11,5031	11,6283	-11,5031	0,0000	11.3933
-11.3278 NA'VE 8.7729	-19,9019	-19,9553	11,5031	11,6283	-11,5031	0,0000	-14,7468	-11,3933	11.3841
-11.3841	-11,3743	-11,3712							11.3743
									11.3712
									0.0000

Tabla 6. Resultados de Diebold y Marino para la estación Campo de Marte: el estadístico DM Los valores para todos los modelos se dan en la Tabla 2.

Con un día de antelación (24 horas de antelación)											
Modelos	AR	ARMA	NPAR	ARMA	NNA	NA'VE	SNA'VE				
	0.0000	NPAR	2.8199	-2,8199	-2,8217	-2,8216	-2,8222	0,0000	1,2125	2,6957	2.5844
	ARMA	2.8217	NNA	1,7028	-1,2125	0,0000	2,8074	2,1681	-2,6957	-2,8074	2.6888
	2.8216	NA'VE	2.8222	0,0000	-2,8480	-1,7028	-2,16 81	2,8480	0,0000	-2,6888	2.6894
	SNA'VE	-2.5844		-2,6894	-2,6887	-2,6896	Con dos días de antelación (48 horas de antelación)				2.6887
											2.6896
											0.0000
Con tres días de antelación (72 horas de antelación)											
Modelos	AR	ARMA	NPAR	ARMA	NNA	NA'VE	SNA'VE				
	AR	0.0000	NPAR	-4,6521	-4,6491	-4,6552	-1,9907	-4,6501			4.0207
	4.6521	ARMA	4.6491	0,0000	4,4231	0,0000	4,2369	7,3595			4.2838
	NNA	4.6552	NA'VE	1,9907	-4,2369	0,0000	-4,0521	4,0521			4.2827
	4.6501	SNA'VE		-4,4231	4,2581	-4,2827	-4,2830	-4,2581			4.2830
	-4.0207			-7,3595				0,0000			4.2828
				-4,2838				-4,2828			0.0000
Con tres días de antelación (72 horas por delante)											
Modelos	AR	ARMA	NPAR	ARMA	NNA	NA'VE	SNA'VE				
	0.0000	NPAR	6.1758	-6.1758	-6,1735	-6,1677	-3,0788	-6,1719			6.8057
	ARMA	6.1735	NNA	0.0000	6,8135	0,0000	6,5635	10,4177			6.5771
	6.1677	NA'VE	6.1719	3,0788	-6,5635	0,0000	-6,7891	6,7891			6.5758
	SNA'VE	-6.8057		-6.8135	6,5119	-6,5758	-6,5758	-6,5119			6.5758
				-10,4177				0,0000			6.5755
				-6.5771				-6,5755			0.0000
Siete días por delante (168 horas por delante)											
Modelos	AR	ARMA	NPAR	ARMA	NNA	NA'VE	SNA'VE				
	ARMA	0.00000		-6.26121	-6,26481	-6,24572	-6,25654	7,71715			8.29033
	NPAR	6.26121	ARMA	0,00000	-10,89898	10,43512	8,04551				8.01439
	6.26481	NNA	6.24572	10,89898	0,00000	11,12267	0,00000				8.01600
	NA'VE	6.25654		-7,71715	-8,04551	-7,32966	7,32966	0,00000			8.01583
	SNA'VE	-8.29033		-10,43512	-11,12267						8.01278
				-8,01439	-8,01600	-8,01583	-8,01278				0.00000

Tabla 7. Resultados de Diebold y Marino para la estación San Borja: los valores estadísticos de DM para Todos los modelos se dan en la Tabla 3.

		Con un día de antelación (24 horas de antelación)				
Modelos	AR	NP	AR	NNA	NA	SNA
0.0000	NP	AR	3.7110	-3,7110	-3,7439	-3,7089
ARMA	3.7439	NNA	2,9423	-1,6320	0,0000	5,4509
3.7089	NA	3.7168	0,0000	-3,9880	-2,9423	0,36
SNA	-4.0280	-3,8175	-3,8394	-3,8207	Con dos días de antelación	3.8175
		(48 horas de antelación)				3.8207
						0.0000
		Con tres días de antelación				
Modelos	AR	NP	AR	NNA	NA	SNA
0.0000	NP	AR	5.8812	-5,8247	-5,8246	-5,8486
NP	5.8812	ARMA	-1,4931	2,4739	0,0000	5,8244
5.8247	NNA	5.8246	0,0000	-6,6795	1,4931	-3,25
NA	5.8486	-5,8805	-5,8819	-5,8965	Con tres días de antelación	5.8819
SNA	-5.9755	(72 horas de antelación)				5.8965
						0.0000
		Con seis días de antelación				
Modelos	AR	NP	AR	NNA	NA	SNA
0.0000	NP	AR	-6,0736	-6,0326	0,0000	-6,0078
NP	6.0736	ARMA	-3,1012	3,1012	0,0000	8,7842
6.0326	NNA	6.0078	-8,7842	-6,7054	0,4728	6,7054
NA	6.0385	-5,2996	-6,6660	-6,6326	0,0000	-7,1164
SNA	-7.8289	Siete días por delante				7,1164
		(168 horas por delante)				-6,6308
						-6,6399
						0.0000
		Con doce días de antelación				
Modelos	\$	NP	AR	NNA	NA	SNA
AR	0.0000	-13.1648	-13.0593	-13.0974	-13.1023	10.7537
NP	13.1648	ARMA	0.0000	-4.0693	11.9360	-1,0432
13.0594	4.0693	0.0000	12.1186	6.6877	12.2212	12.2212
NNA	13.0974	-11.9360	-12.1186	0.0000	NA	12.8686
NA	12.2237	-12.8686	12.2237	12.2237	12.2237	12.2237
SNA	-10.7537	-12.2667	-12.2212	-12.2043	-12.2429	6.6877
						0.0000

Tabla 8. Resultados de Diebold y Marino para la estación Santa Anita: los valores estadísticos de DM para Todos los modelos se dan en la Tabla 4.

		Con un día de antelación (24 horas de antelación)				
Modelos	AR	NP	AR	NNA	NA	SNA
0.0000	NP	AR	-5,2875	-5,3083	-5,2203	-5,2884
NP	5.2875	ARMA	4,9491	-2,7675	0,0000	6,7317
5.3083	NNA	5.2203	-6,7317	0,0000	-6,3213	-4,9491
NA	5.2884	SNA	-6,1872	-6,2071	-6,1852	-6,1909
-8.1294		antelación (48 horas de antelación)				6.1909
						0.0000
		Con tres días de antelación				
Modelos	AR	NP	AR	NNA	NA	SNA
0.0000	NP	AR	-5,8369	-5,8138	-5,7600	-5,8183
NP	5.8369	ARMA	0,0000	0,0000	6,5466	-7,0218
5.8138	NNA	5.7600	3,0381	-6,5466	0,0000	6,6721
NA	5.8183	SNA	-7.8662	0,2926	-4,8688	0,0000
-7.8662		antelación (72 horas de antelación)				6.5245
						0.0000
		Con seis días de antelación				
Modelos	AR	NP	AR	NNA	NA	SNA
0.0000	NP	AR	-7,3310	-7,2985	-7,2848	-3,4977
NP	7.3310	ARMA	0,0000	8,0225	0,0000	7,4779
7.2985	NNA	7.2848	3,4977	-7,4779	0,0000	-5,7668
NA	7.3063	SNA	-8,0225	7,6111	-7,5678	-7,5720
-8.0393		antelación (168 horas por delante)				0,0000
						7.5737
						0.0000
		Con doce días de antelación				
Modelos	AR	NP	AR	NNA	NA	SNA
0.0000	NP	AR	-6,3528	-6,3311	-6,2905	-6,3270
NP	6.3528	ARMA	0,0000	-3,4125	7,3201	6,4540
6.2905	NA	6.3365	3,4125	0,0000	6,8884	-6,8547
6.3270	SNA	-8.3365	-7,3201	-6,8884	0,0000	-6,4540
-8.3365		antelación (168 horas por delante)				6,8547
						0,0000
						6.9785
						0.0000

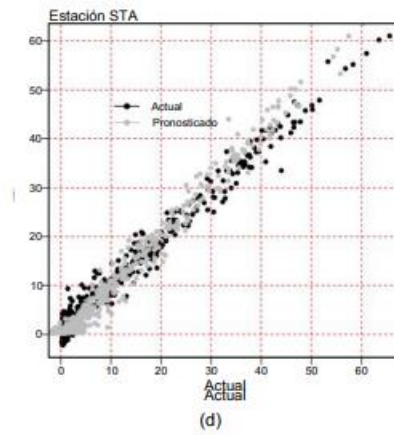
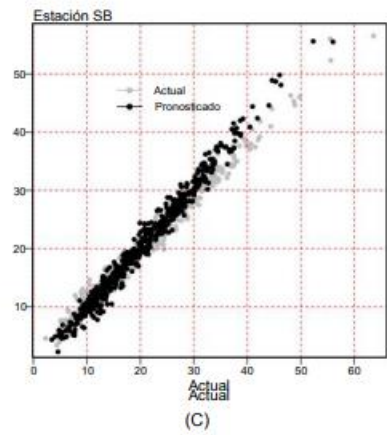
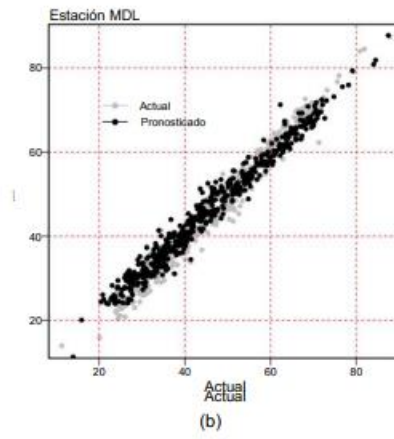
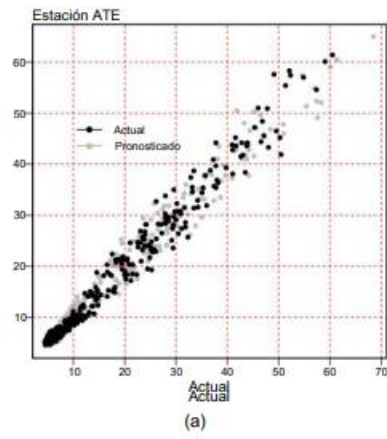


Figura 1. Gráfico de correlación que muestra la concentración de ozono (μgm^3) para todas las estaciones. Muestra los mejores modelos de predicción de la concentración de ozono, que son NPAR (ubicado en arriba a la izquierda), AR (arriba a la derecha), ARMA (abajo a la izquierda) y NPAR (abajo a la derecha).

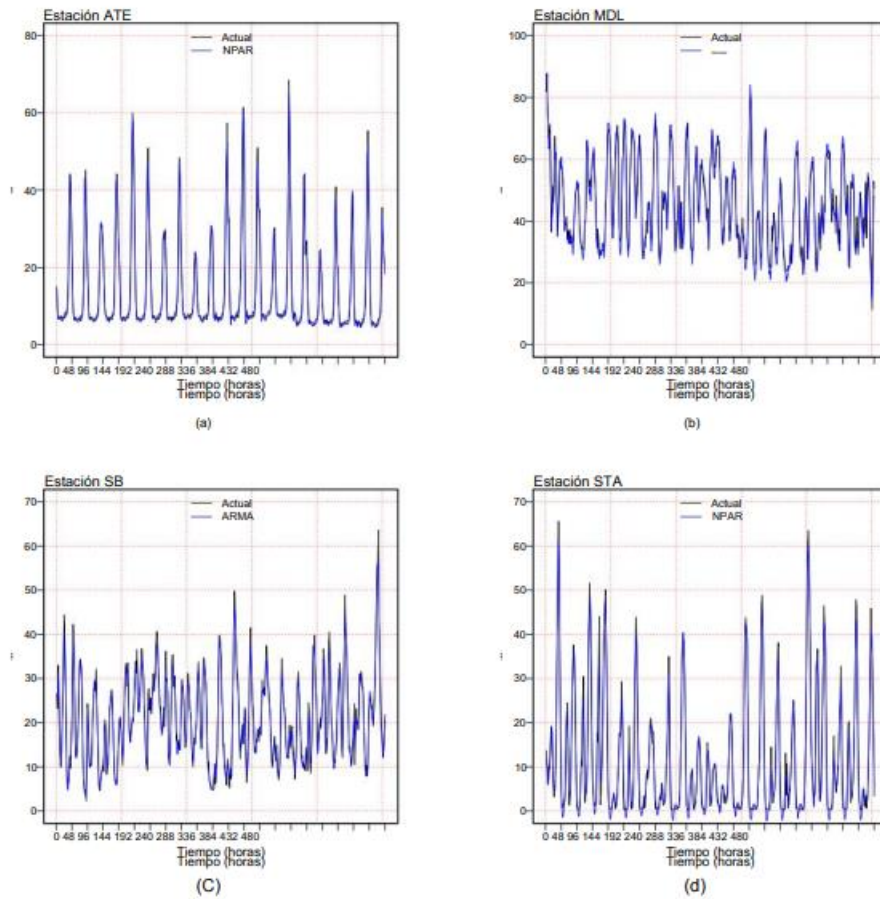


Figura 2. Este gráfico muestra los puntos de datos reales y pronosticados para la concentración de ozono (μgm^3) durante 504 horas en cuatro estaciones diferentes: Estación Ate (a), Estación Campo de Marte (b), Estación San Borja (c) y Santa Anita. Estación (d). Se utilizó el mejor modelo para cada estación.

6. Evidencia de Sumisión.

Manuscript Status

Manuscript Status				
Journal	Manuscript ID	Special Issue	Title	Status
Environ	Environ-715		Multi-step ahead ozone level forecasting using a component-based technique: A case study in Lima, Peru	Revised version review

Gmail - Submission Received—Environ-715



Javier Linkolk López Gonzales <javierlinkolk@gmail.com>

Submission Received—Environ-715

2 mensajes

environ@aimspress.org <environ@aimspress.org> 30 de enero de 2024, 1:21
Responder a: hongzhu.cao@aimsciences.org
Para: Javier Linkolk López-Gonzales <javierlinkolk@gmail.com>

Dear Dr. López-Gonzales,

Thank you for submitting the following manuscript for publication in AIMS Environmental Science:

Manuscript ID: Environ-715
Type of manuscript: Research article
Title: Multi-step ahead ozone level forecasting using a component-based technique: A case study in Lima, Peru
Authors: Flor Quispe, Eddy Salcedo, Javier Linkolk López-Gonzales *
Received: 30 January 2024
E-mails: vickyquispe@upeu.edu.pe, eddy.salcedo@archean.com.pe, javierlinkolk@gmail.com

7. Resolución de inscripción de perfil de proyecto de tesis en formato artículo.



“AÑO DE LA UNIDAD, LA PAZ Y EL DESARROLLO”

RESOLUCIÓN N° 0515-2023/UPeU-FIA-CF-T

Lima, Ñaña 12 de setiembre de 2023

VISTO:

El expediente de **Flor Vicky Quispe Sanchez**, identificado(a) con Código Universitario N° 201320326 y **Eddy Albino Salcedo Esquen**, identificado(a) con Código Universitario N° 201322886, de la Escuela Profesional de Ingeniería Ambiental de la Facultad de Ingeniería y Arquitectura de la Universidad Peruana Unión;

CONSIDERANDO

Que la Universidad Peruana Unión tiene autonomía académica, administrativa y normativa, dentro del ámbito establecido por la Ley Universitaria N° 30220 y el Estatuto de la Universidad;

Que la Facultad de Ingeniería y Arquitectura de la Universidad Peruana Unión, mediante sus reglamentos académicos y administrativos, ha establecido las formas y procedimientos para la aprobación e inscripción del perfil de proyecto de tesis en formato artículo y la designación o nombramiento del asesor para la obtención del título profesional;

Que **Flor Vicky Quispe Sanchez** y **Eddy Albino Salcedo Esquen**, han solicitado: la inscripción del perfil de proyecto de tesis titulado "Pronóstico a corto plazo de la concentración de ozono en Lima metropolitana utilizando una combinación híbrida de modelos de series de tiempo" y la designación del Asesor, encargado de orientar y asesorar la ejecución del perfil de proyecto de tesis en formato artículo;

Estando a lo acordado en la sesión del Consejo de la Facultad de Ingeniería y Arquitectura de la Universidad Peruana Unión, celebrada el 12 de setiembre de 2023, y en aplicación del Estatuto y el Reglamento General de Investigación de la Universidad;

SE RESUELVE:

Aprobar el perfil de proyecto de tesis en formato artículo titulado "**Pronóstico a corto plazo de la concentración de ozono en Lima metropolitana utilizando una combinación híbrida de modelos de series de tiempo**" y disponer su inscripción en el registro correspondiente, designar a **Dr. Javier Linkolk López Gonzales** como ASESOR para que oriente y asesore la ejecución del perfil de proyecto de tesis en formato artículo el cual fue dictaminado por: **Mg. Jackson Edgardo Perez Carpio** y **Ing. Orlando Alan Poma Porras**, otorgándoles un plazo máximo de doce (12) meses para la ejecución.

Regístrese, comuníquese y archívese.




Dra. Erika Inés Acuña Salinas
DECANA




Mg. Milda Amparo Cruz Huaranga
SECRETARIA ACADÉMICA (e)

cc:
-Interesado
-Asesor
-Dirección General de Investigación
-Archivo