



Ciencia e Ingeniería Neogranadina

ISSN: 0124-8170

ISSN: 1909-7735

Universidad Militar Nueva Granada

Parody Muñoz, Alexander Elías; Castillo Ramírez, Margarita; Mendoza Hernández, Martha; Torres Garces, Alejandra; Torrijos Espitia, María; Sanmartín Mendoza, Paul
Inteligencia artificial y la estadística multivariada en el análisis
del comportamiento de PM2.5 en la ciudad de Barranquilla*
Ciencia e Ingeniería Neogranadina, vol. 33, núm. 2, 2023, Julio-Diciembre, pp. 51-64
Universidad Militar Nueva Granada

DOI: <https://doi.org/10.18359/rcin.6789>

Disponible en: <https://www.redalyc.org/articulo.oa?id=91178599005>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](https://www.redalyc.org)



Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

Inteligencia artificial y la estadística multivariada en el análisis del comportamiento de PM_{2.5} en la ciudad de Barranquilla*

Alexander Elías Parody Muñoz^a ■ Margarita Castillo Ramírez^b ■ Martha Mendoza Hernández^c ■ Alejandra Torres Garces^d ■ María Torrijos Espitia^e ■ Paul Sanmartín Mendoza^f

Resumen: Predecir el comportamiento de los contaminantes atmosféricos se ha convertido en un importante insumo para el diseño de estrategias gubernamentales de contingencia que buscan mitigar los niveles de contaminación, por tal motivo esta investigación presenta una metodología para predecir el incumplimiento de los índices máximos permitidos para el material particulado contaminante PM_{2.5}, a partir de modelos estadísticos multivariados que pueden ser utilizados para entender cómo diversas variables meteorológicas influyen en el comportamiento del contaminante, al igual que como forma de seleccionar las variables que serán usadas durante el entrenamiento de un clasificador bayesiano de redes neuronales. Además del modelo multivariado, se aplicó un análisis de varianza multifactorial para evaluar el efecto del día, hora y mes en la media del PM_{2.5}. Como resultados principales, se encontró que las variables: dirección del viento, temperatura, humedad, presión atmosférica, precipitación, radiación solar, hora del día, día de la semana y mes, fueron estadísticamente significativas y permitieron entrenar una red neuronal con una capacidad de pronóstico

-
- * Artículo de investigación producto del proyecto “Estadística en la toma de decisiones” de la Universidad Libre.
- a Magíster en estadística aplicada, ingeniero industrial, Universidad Libre, Barranquilla, Colombia. Correo electrónico: alexandere.parodym@unilibre.edu.co ORCID: <https://orcid.org/0000-0003-0155-266X>
- b Magíster en ingeniería ambiental, bióloga, Universidad Reformada, Barranquilla, Colombia. Correo electrónico: m.castillo@unireformada.edu.co ORCID: <https://orcid.org/0000-0002-6839-3133>
- c Magíster en ingeniería ambiental, ingeniera ambiental, Universidad Reformada, Barranquilla, Colombia. Correo electrónico: m.mendoza@unireformada.edu.co ORCID: <https://orcid.org/0000-0001-5535-8868>
- d Ingeniera industrial, Universidad Libre, Barranquilla, Colombia. Correo electrónico: alejandras-torres@unilibre.edu.co ORCID: <https://orcid.org/0009-0008-5198-0586>
- e Ingeniera industrial, Universidad Libre, Barranquilla, Colombia. Correo electrónico: mariaf-torrijose@unilibre.edu.co ORCID: <https://orcid.org/0009-0004-2513-1012>
- f Ph. D. en ingeniería de sistemas y computación, magíster en ingeniería de sistemas y computación, ingeniero de sistemas, Universidad Libre, Barranquilla, Colombia. Correo electrónico: paul.sanmartinm@unilibre.edu.co ORCID: <https://orcid.org/0000-0002-5748-2333>

del 78,2 % de manera global y del 81 % en los casos de incumplimiento, demostrando así que con variables factibles de manejar (variables meteorológicas, hora, día y mes) se puede diseñar una herramienta de pronóstico muy útil como insumo para la gestión sobre la calidad del aire.

Palabras clave: contaminación atmosférica; análisis estadístico; redes neuronales

Recibido: 12/06/2023 **Aceptado:** 04/09/2023 **Disponible en línea:** 27/12/2023

Cómo citar: A. E. Parody Muñoz, M. Castillo Ramírez, M. Mendoza Hernández, A. Torres Garcés, M. Torrijos Espitia, y P. Sanmartín Mendoza, «Inteligencia artificial y la estadística multivariada en el análisis del comportamiento de PM2.5 en la ciudad de Barranquilla», Cien.Ing.Neogranadina, vol. 33, n.º 2, pp. 51-64.Diciembre 2023.

Artificial Intelligence and Multivariate Statistics in the Analysis of PM2.5 Behavior in the City of Barranquilla

Abstract: Predicting the behavior of atmospheric pollutants has become a crucial input for the design of government contingency strategies aiming to mitigate pollution levels. For this reason, this research presents a methodology to predict the non-compliance with the maximum allowed indices for the particulate matter pollutant PM2.5. The approach involves multivariate statistical models that can be used to understand how various meteorological variables influence the pollutant's behavior. It also serves as a means of selecting variables for training a Bayesian neural network classifier. In addition to the multivariate model, a multifactorial analysis of variance was applied to assess the effect of day, hour, and month on the mean PM2.5 levels. The main results revealed that variables such as wind direction, temperature, humidity, atmospheric pressure, precipitation, solar radiation, time of day, day of the week, and month were statistically significant. These variables allowed the training of a neural network with an overall predictive capacity of 78.2%, reaching 81% accuracy in cases of non-compliance. This demonstrates that with manageable variables (meteorological variables, time, day, and month), a highly useful forecasting tool can be designed to contribute to air quality management.

Keywords: Atmospheric Pollution; Statistical Analysis; Neural Networks

Introducción

Según la Organización Mundial de la Salud (OMS), la contaminación atmosférica es una de las problemáticas ambientales más importantes a nivel mundial, dado que cerca del 90 % de las personas respiran aire con niveles de contaminación por encima de los límites internacionales sugeridos. La contaminación atmosférica, tanto en interiores como en exteriores, genera al año 7 millones de muertes tempranas, de las cuales, aproximadamente 4,2 millones son debidas a mala calidad de aire en exteriores, lo que deriva en accidentes cerebrovasculares, enfermedades coronarias o enfermedades respiratorias agudas [1].

De acuerdo con estudios del Instituto Nacional de Salud y del Observatorio Nacional de Salud, en el país se registran más de 15 mil muertes anuales debido a la exposición a ambientes con aire de mala calidad, y las afecciones relatadas con mayor nivel de mortalidad son la enfermedad pulmonar obstructiva (17,6 %) y la enfermedad isquémica del corazón (13,9 %) [2].

Dentro de los contaminantes atmosféricos que generan mayor impacto a la salud se encuentra el material particulado, siendo el PM_{2.5} el que cobra gran importancia respecto al daño a la salud humana y su contribución al cambio climático [3]. Es por esto que es necesario comprender su comportamiento y distribución, considerando su relación con las variables meteorológicas que juegan un papel fundamental de transporte y dispersión.

Herramientas estadísticas como las redes neuronales y los modelos de regresión pueden ayudar a comprender la relación específica existente entre el comportamiento de PM_{2.5} y las variables meteorológicas, permitiendo generar un pronóstico de la probabilidad de incumplimiento en los niveles máximos de contaminante. Es por esto que el objetivo de esta investigación es generar una herramienta para el pronóstico del cumplimiento de los niveles de PM_{2.5}, según los estándares establecidos por la OMS (2021), la cual ayudará a la gestión del recurso atmosférico en la ciudad de Barranquilla.

Es así como por medio de la entidad Barranquilla Verde se han realizado estudios del

comportamiento de los contaminantes PM₁₀ y PM_{2.5}, evidenciando en el primer semestre del 2021 un comportamiento homogéneo entre ambos, con medias por debajo de los límites establecidos. Cabe aclarar que durante los meses de abril y junio se registraron picos en el contaminante PM₁₀ en las estaciones Escuela de Policía y Unidad Móvil, si bien las concentraciones disminuyeron considerablemente durante las siguientes semanas [4].

A su vez, el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) realiza anualmente un informe del estado de la calidad del aire en Colombia, dadas las concentraciones de PM₁₀ y PM_{2.5} reportadas por las once estaciones de monitoreo administradas y operadas por la autoridad ambiental en Bogotá, según el cual se ha podido evidenciar que en los meses más secos de la ciudad de Barranquilla (noviembre y febrero) se presentaron las concentraciones más altas de los contaminantes criterio, muy seguramente debido a las bajas temperaturas que ocurren en horas de la noche, las cuales dificultan el ascenso y la dispersión de los mismos [5].

El interés por el estudio del comportamiento del material particulado en la ciudad de Barranquilla ha tenido un incremento en los últimos años, gracias a la disponibilidad de información [6-9]. No obstante, los artículos publicados sobre material particulado en la ciudad no han explorado el desarrollo de herramientas que intenten pronosticar comportamientos futuros de PM_{2.5}, analizando el efecto de las variables meteorológicas sobre este contaminante, como una ayuda para el desarrollo de políticas públicas asociadas a la gestión de la calidad del aire.

Materiales y métodos

Para el logro de los objetivos del presente estudio se tomaron las bases de datos generadas por la red de monitoreo de la calidad del aire y manejadas por Barranquilla Verde, la autoridad ambiental de la ciudad. Los años analizados fueron 2018 y 2019, para evitar el sesgo relacionado con la afectación que generó la pandemia sobre el comportamiento del PM_{2.5}, además de que se tomaron registros

de los parámetros meteorológicos: dirección del viento, temperatura, velocidad del viento, humedad, presión atmosférica, radiación solar y precipitación, así como del mes, el día y la hora, esto último con el propósito de representar las fluctuaciones de las actividades económicas que generan emisiones atmosféricas e impactan en la concentración del $PM_{2.5}$, conformando finalmente una matriz con 17 521 filas de información. Por último, en la base de datos se generó una nueva variable de naturaleza dicotómica, para determinar si la concentración promedio del contaminante a lo largo de 24 horas superaba o no el límite máximo sugerido por la OMS.

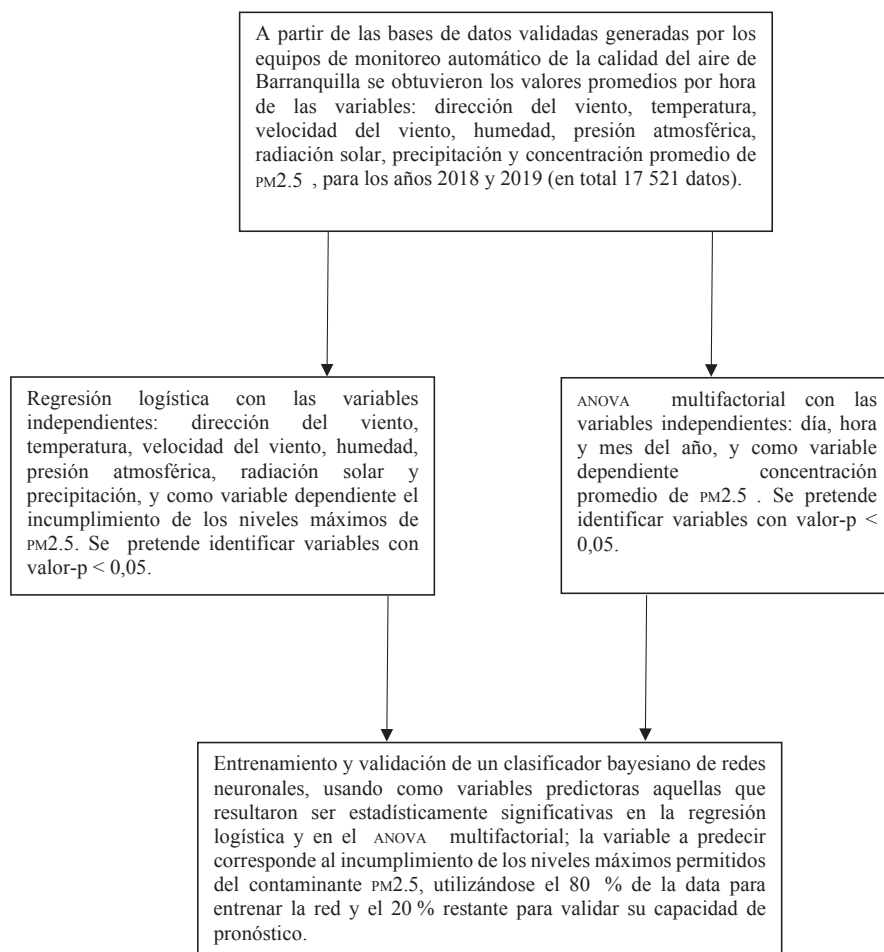
Para identificar el efecto de los parámetros meteorológicos mencionados en la concentración del $PM_{2.5}$ se utilizó un modelo de regresión logística, en el que los mismos fueron tomados como variables independientes, y como variable dependiente se estableció el incumplimiento de la concentración máxima de $PM_{2.5}$, según la OMS. El modelo de regresión logística es útil para determinar si existe alguna relación estadísticamente significativa entre una o más variables cuantitativas o cualitativas, con respecto a una variable cualitativa dicotómica que representa la presencia o ausencia del fenómeno estudiado [10]. Por tanto, es adecuado para identificar con claridad la influencia de las variables meteorológicas en el incumplimiento de los niveles máximos de $PM_{2.5}$ en la ciudad de Barranquilla, de acuerdo con los parámetros sugeridos de la OMS. Para estudiar el efecto de las variables temporales: día de la semana, mes del año y hora del día, se aplicó un ANOVA (análisis de varianza) multifactorial, el cual permite evidenciar si hay diferencias en la concentración promedio del contaminante según el momento. El propósito de incluir estos factores es poder caracterizar el comportamiento de los ciclos productivos y comerciales, de manera que el modelo final resultante cuente con un número de variables suficientes que garanticen un pronóstico con un aceptable nivel de confiabilidad.

Como fase final del apartado estadístico se entrenó un clasificador bayesiano de redes neuronales, utilizando como parámetros de entrada las variables que resultaron estadísticamente significativas tanto en el modelo logístico como en el ANOVA. El propósito de la red neuronal es predecir con un buen nivel de exactitud el incumplimiento de la concentración máxima del $PM_{2.5}$ (según la OMS), para probar que se puede construir una herramienta de pronóstico con variables que sean factibles de establecer (día, hora y mes) y de pronosticar (meteorológicas), de forma que la administración pública local pueda contar con información que le permita actuar de manera preventiva y establecer posibles contingencias por mala calidad del aire en la ciudad. Cabe destacar que la red fue entrenada con el 80 % de la información recopilada en las bases de datos, mientras que el 20 % restante de los datos se usó para validar la red y medir su capacidad de predicción. Los resultados del desempeño de la red, tanto en su fase de entrenamiento como en la fase de validación, se presentaron en la forma de matrices de confusión.

Por otro lado, se estableció la influencia de la hora de la medición, así como del día de la semana y el mes del año en el comportamiento de los contaminantes atmosféricos en estudio, por medio de un ANOVA multifactorial, en el cual se analizaron las interacciones de segundo orden para determinar si la interacción hora/día, hora/mes y día/mes ejercía alguna influencia significativa en el comportamiento del $PM_{2.5}$. En este caso, el ANOVA multifactorial es una herramienta estadística adecuada que permite explicar la posible relación estadística entre variables cualitativas (hora, día y mes) con variables cuantitativas de interés (concentración de los contaminantes) [11].

En la figura 1 se puede visualizar la metodología estadística, haciendo especial énfasis tanto en las técnicas de inferencia estadística, como modelo de regresión logística y el ANOVA multifactorial, como en la herramienta de *machine learning*, en este caso la red neuronal:

Figura 1. Desglose de la metodología estadística aplicada en la investigación



Fuente. elaboración propia.

Como propósito principal del artículo y aporte tanto al campo ingenieril como a la gestión público/administrativa, se pretende mostrar el uso concatenado de modelos estadísticos multivariados, técnicas de inferencia estadística y técnicas propias de la inteligencia artificial. En este caso, a través de un modelo multivariado de regresión logística y ANOVA multifactorial, se escogieron las variables que resultaran ser estadísticamente significativas en el comportamiento del PM_{2.5}, las cuales posteriormente se usaron para el entrenamiento de un clasificador bayesiano de redes neuronales que permitiese pronosticar cuándo se superan los límites máximos sugeridos por la OMS en la concentración promedio de PM_{2.5}, convirtiendo así al presente artículo en una guía teórico/práctica que

pudiera ser replicada por autoridades ambientales territoriales y otros tomadores de decisión para la programación de acciones de contingencia ante futuras emergencias por mala calidad del aire, así como para el diseño de sistemas de alertas temprana para la calidad del aire en las ciudades.

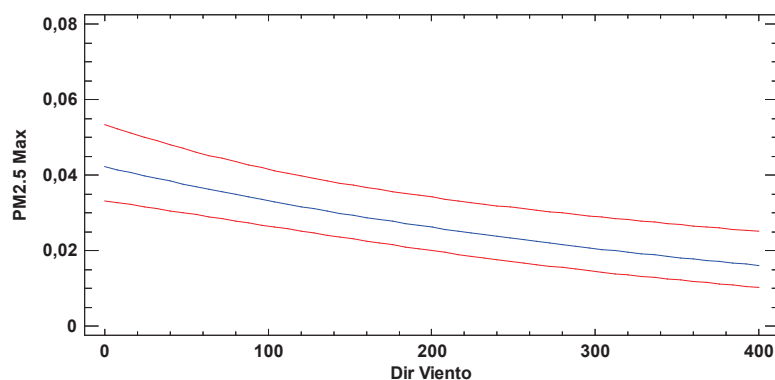
Resultados y discusión

Una vez aplicado el modelo de regresión logística, las variables que resultaron ser estadísticamente significativas con respecto al incumplimiento del nivel máximo de concentración de PM_{2.5} fueron: dirección del viento, temperatura, humedad, presión, radiación solar y precipitación (todas con valor-p de 0,0).

La velocidad del viento presentó un valor-p mayor que 0,05 (0,1110), por tanto, no tiene una influencia significativa en la probabilidad de incumplimiento de los límites de la oms. Para analizar la

influencia de cada variable atmosférica, se presentan los gráficos del modelo logístico, donde se puede ver la variación en la probabilidad de incumplimiento con intervalos del 95 % de confianza:

Figura 2. Gráfico del modelo ajustado para dirección del viento

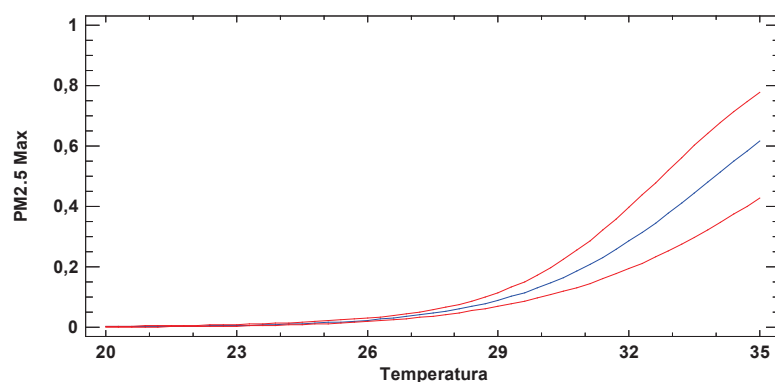


Fuente: elaboración propia.

A partir de la figura 2, que relaciona la dirección del viento con el incumplimiento del nivel máximo permitido de $PM_{2.5}$, puede concluirse que la probabilidad más alta de incumplir los niveles de contaminación se encuentra cuando la dirección

del viento se ubica entre los 0° y los 90° , es decir, hacia el noreste, indicando esto que posiblemente la contaminación proveniente desde el interior del país sea uno de los principales generadores de incremento de $PM_{2.5}$ en Barranquilla.

Figura 3. Gráfico del modelo ajustado para temperatura

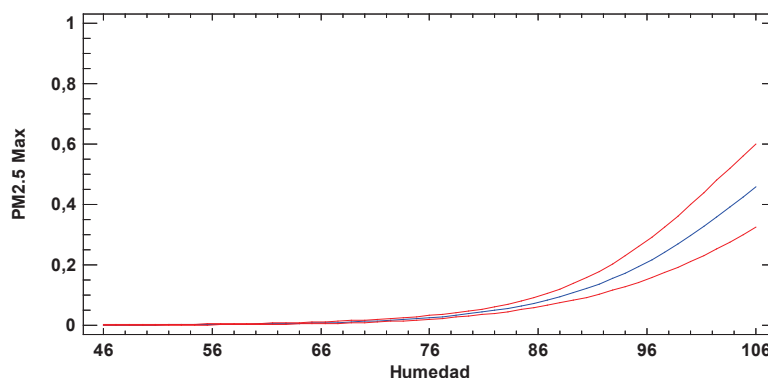


Fuente: elaboración propia.

Con respecto a la figura 3, que relaciona temperatura y el incumplimiento del nivel máximo permitido de $PM_{2.5}$, podemos evidenciar que

temperaturas mayores a los 30 °C están asociadas a mayor probabilidad de incumplimiento.

Figura 4. Gráfico del modelo ajustado para humedad

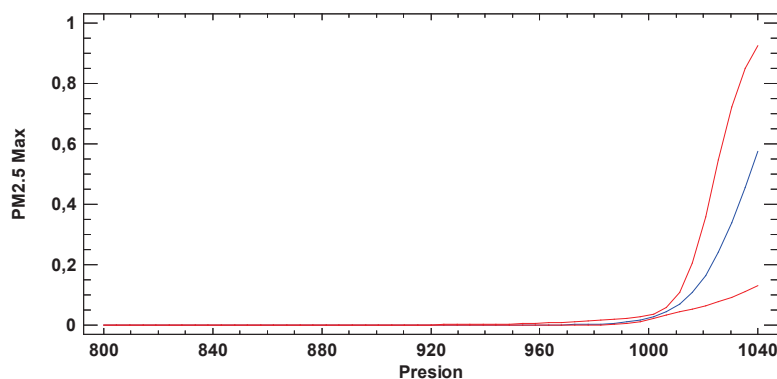


Fuente: elaboración propia.

La figura 4 relaciona la probabilidad de incumplimiento y el porcentaje de humedad, evidenciando un incremento en los niveles de

incumplimiento del nivel máximo permitido de $PM_{2.5}$ cuando el porcentaje de humedad está por encima del 86 %.

Figura 5. Gráfico del modelo ajustado para presión

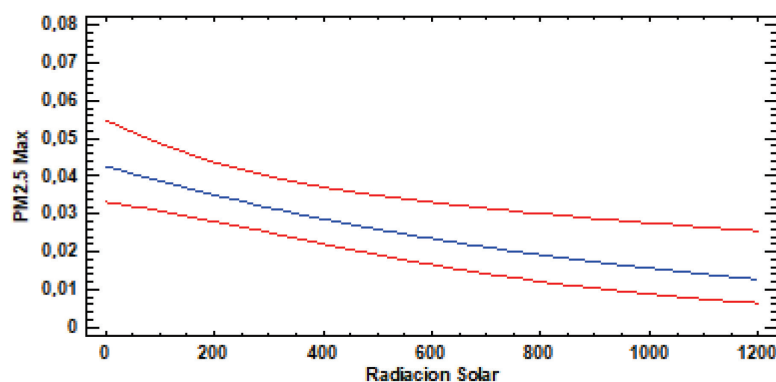


Fuente: elaboración propia.

En la figura 5 se puede evidenciar el efecto de la presión atmosférica sobre el comportamiento de PM 2.5, encontrándose que a partir de los 1000

mm Hg se incrementa la probabilidad de incumplimiento de los niveles máximos establecidos por la OMS.

Figura 6. Gráfico del modelo ajustado para radiación solar

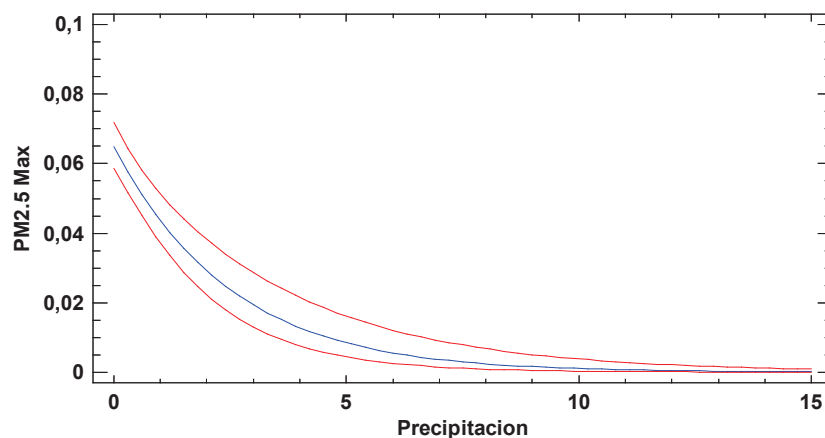


Fuente: elaboración propia.

La figura 6 muestra un comportamiento inverso, por lo que, a mayor radiación, menor probabilidad

de incumplir los niveles máximos de PM2.5, aunque las probabilidades iniciales son igualmente bajas.

Figura 7. Gráfico del modelo ajustado para precipitación



Fuente: elaboración propia.

Por último, la figura 7 señala también un comportamiento inverso para el nivel de precipitación, por lo que a mayor precipitación menor

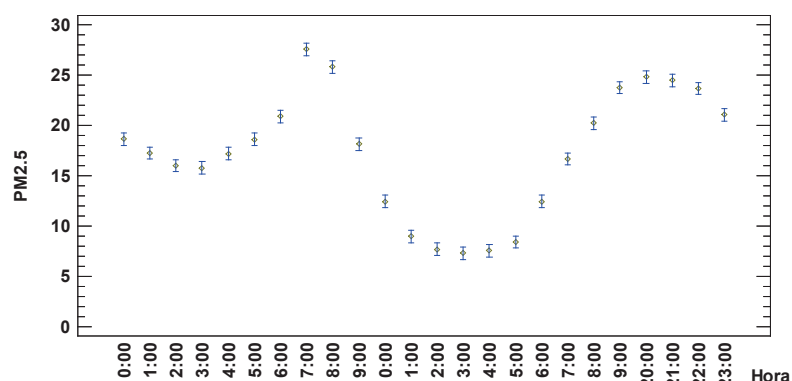
probabilidad de incumplir con la concentración máxima de contaminante. Esto podría atribuirse a que en épocas de lluvia no es frecuente la ocurrencia

de quemas en zonas aledañas a la ciudad, que son las que incrementan los niveles de concentración de $PM_{2.5}$ en el medio ambiente.

A través de ANOVA multivariado se estudió la influencia del mes, el día y la hora en la concentración promedio de $PM_{2.5}$, obteniéndose que

las variables mes (valor-p de 0,0), día (valor-p de 0,0126), hora (valor-p de 0,0) fueron estadísticamente significativas, así como las interacciones de mes y día (valor-p de 0,0), y de mes y hora (valor-p de 0,0), a diferencia de la interacción de día y hora (valor-p de 0,9902), que no fue significativa.

Figura 8. Gráfico de la prueba LSD de Fisher para hora del día

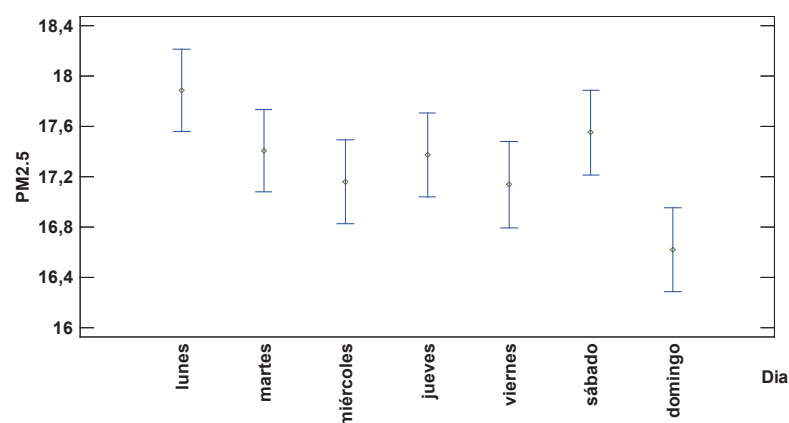


Fuente: elaboración propia.

La figura 8 representa la prueba LSD de Fisher para hora, evidenciándose dos picos a lo largo del día con mayor concentración de $PM_{2.5}$, entre las 7:00 y las 8:00 a.m., y entre las 7:00 y las 10:00 p.m.; en

cambio, los niveles más bajos de concentración de $PM_{2.5}$ ocurren entre las 11:00 a.m. y las 3:00 p.m., lo cual se puede relacionar con las horas pico del tráfico, cuando hay más afluencia de vehículos en la ciudad.

Figura 9. Gráfico de la prueba LSD de Fisher para día de la semana

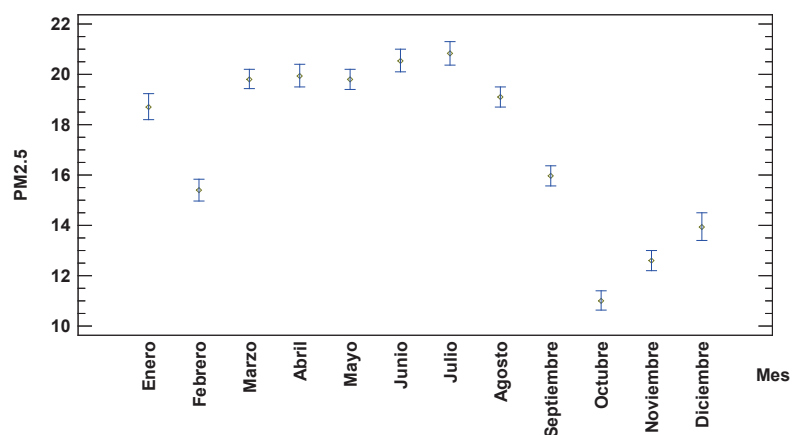


Fuente: elaboración propia.

En la figura 9, se puede observar con respecto al día de la semana que no se evidencian diferencias significativas de martes a sábado, en marcado

contraste con los lunes y los domingos, días de inicio y finalización de las jornadas laborales y de estudio de la mayoría de los habitantes.

Figura 10. Gráfico de la prueba LSD de Fisher para mes del año

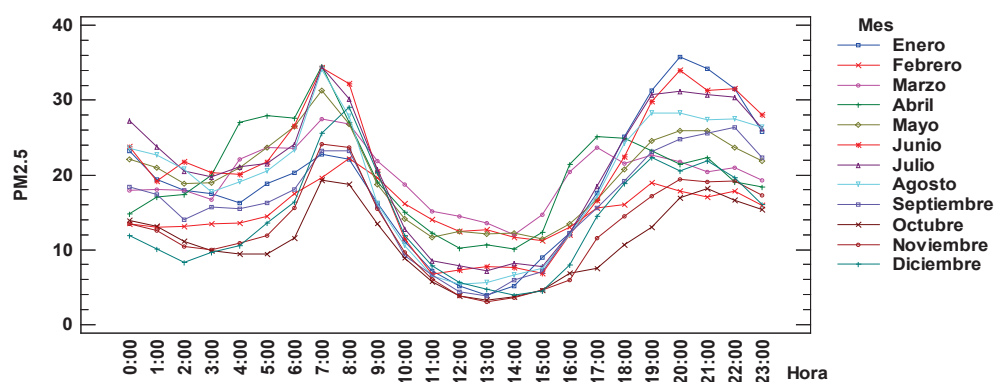


Fuente: elaboración propia.

La figura 10 representa la prueba LSD de Fisher para mes del año, pudiéndose evidenciar que los meses con mayor concentración promedio de PM2.5 son junio y julio, lo cual se encuentra asociado a la preparación de las tierras para cultivo en la Vía Parque Isla de Salamanca, cuando se hacen

quemadas no controladas que impactan la calidad del aire en la ciudad. En cambio, el mes con más baja concentración en el año de 2019 fue octubre, que corresponde al periodo de lluvias, las cuales provocan un lavado en la atmósfera que disminuye los niveles de material particulado.

Figura 11. Gráfico de interacción entre hora y mes con respecto a la concentración promedio de PM2.5

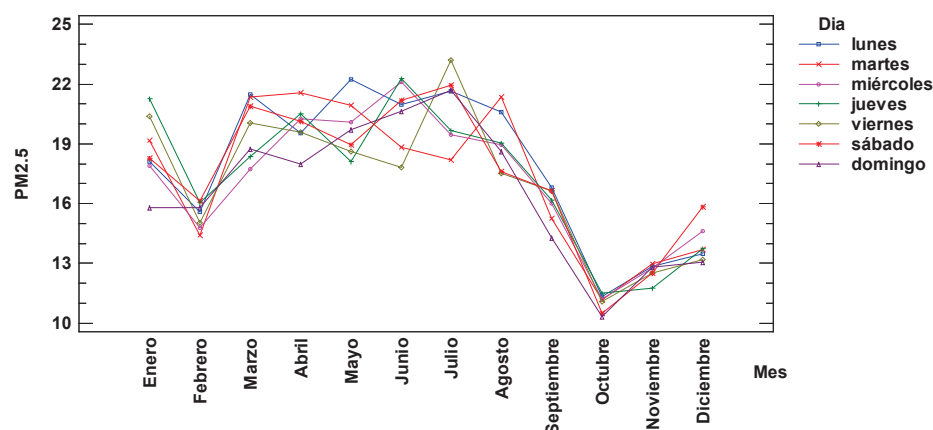


Fuente: elaboración propia.

Sobre la interacción de la hora y el mes, la figura 11 arrojó que en las horas de la mañana (entre 7:00 y 8:00 a.m.) es más alta la concentración de $PM_{2.5}$ para los meses de abril, junio, julio y

agosto en comparación con el resto del año, coincidiendo con las condiciones meteorológicas que contribuyen al incremento de la concentración del contaminante.

Figura 12. Gráfico de interacción entre día y mes con respecto a la concentración promedio de $PM_{2.5}$



Fuente: elaboración propia.

Por último, el gráfico 12 representa la relación del mes y del día del año, pudiéndose evidenciar que el periodo en el que existe una mayor afectación sobre la concentración del día corresponde a los meses entre marzo y agosto. También se puede resaltar cómo en los viernes del mes de julio se obtiene la mayor concentración promedio de $PM_{2.5}$, contrario a los viernes en junio, cuando se logra una de las concentraciones promedio más bajas.

Posteriormente, se generó un clasificador bayesiano de redes neuronales para medir la capacidad de las variables en estudio a la hora de predecir el incumplimiento del nivel máximo de concentración permitido para $PM_{2.5}$ según la OMS. Vale la pena aclarar que, según la normativa nacional contemplada en la Resolución 2254 de 2017, nunca ocurrió un incumplimiento de los límites

de contaminación atmosférica en la ciudad de Barranquilla, motivo por el cual el análisis se realizó conforme a los valores establecidos por la OMS. Para la construcción de la red neuronal, se decidió trabajar con una versión más reducida de la base de datos, debido a que solo el 7 % de los datos registrados para $PM_{2.5}$ presentó incumplimiento del nivel máximo de concentración, lo cual generaba que la capacidad de predecir un incumplimiento en la concentración promedio de media móvil por parte de la red no fuera muy alta, debido al desbalanceo de la base de datos. Por tanto, se generó una nueva base de datos con 5378 filas de información, donde el 18 % correspondió a datos de incumplimiento en el nivel máximo de $PM_{2.5}$. Esto permitió entrenar una red con mejores resultados en su capacidad de predicción del incumplimiento del nivel máximo, como puede verse a continuación.

Tabla 1. Set de entrenamiento

¿Supera el nivel máximo de PM _{2.5} ? (real)	No. de casos	¿Supera el nivel máximo de PM _{2.5} ? (proyectado)	
		No	Sí
No	767	559 (73 %)	208 (27 %)
Sí	430	102 (24 %)	328 (76 %)
Porcentaje de casos correctamente clasificados: 74,10 %			

Fuente: elaboración propia.

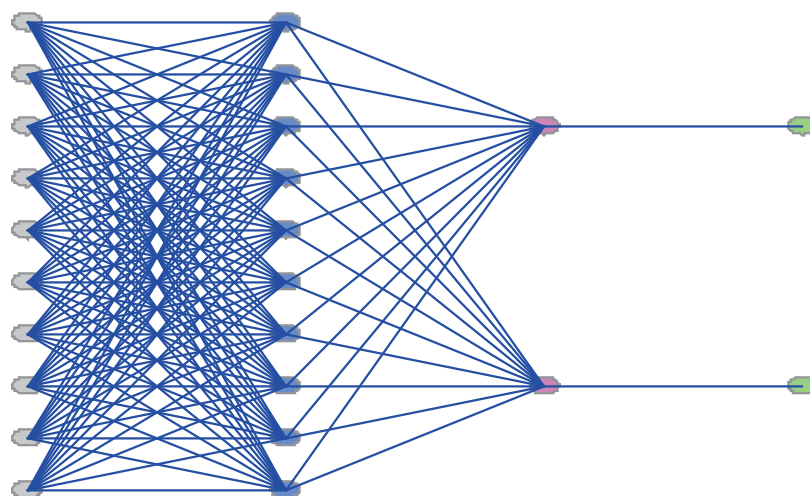
Tabla 2. Set de validación

¿Supera el nivel máximo de PM _{2.5} ? (real)	No. de casos	¿Supera el nivel máximo de PM _{2.5} ? (proyectado)	
		No	Sí
No	185	142 (77 %)	43 (23 %)
Sí	104	20 (19 %)	84 (81 %)
Porcentaje de casos correctamente clasificados: 78,20 %			

Fuente: elaboración propia.

La red neuronal generada presentó buenos resultados en su capacidad de predicción, dado que en el set de entrenamiento se logró un 74,10 % de casos correctamente clasificados, mientras que en el set de validación esta cifra se incrementó hasta un 78,20 %, siendo notable también que la capacidad de predecir un incumplimiento de la concentración máxima de PM_{2.5} fuese del 80,77 %

(ver tabla 2), lo cual permite concluir que con una información confiable de los valores en las variables meteorológicas, sumado al día de la semana, el mes y la hora del día, se puede obtener un pronóstico con buen nivel de confiabilidad sobre el incumplimiento del nivel máximo de concentración para PM_{2.5} conforme a los límites de la OMS de 5 µg/m³.

Figura 13. Representación gráfica de la red neuronal generada

Fuente: elaboración propia.

La figura 13 describe la estructura de la red neuronal entrenada con el propósito de predecir el incumplimiento de los niveles máximos de PM_{2.5}.

Conclusiones

Los resultados de la investigación permitieron establecer relaciones estadísticamente significativas entre la concentración de PM_{2.5} frente al mes, el día de la semana y la hora del día, encontrándose que los meses de mayor concentración fueron junio y julio, debido a la época seca y las quemas para la preparación de terrenos en el Parque Natural Isla de Salamanca que está al frente de la ciudad, alterando de este modo la calidad del aire. La relación frente al día de la semana demostró que los lunes presentaron una mayor concentración asociada al inicio de actividades y el incremento de la movilidad en la ciudad, lo cual a su vez se confirmó a través del relacionamiento con la hora del día, que mostró un comportamiento acorde a la actividad socioeconómica de la ciudad, en el que las horas pico de 7:00-10:00 a.m. y de 7:00-9:00 p.m. registraron la mayor concentración de PM_{2.5}. En relación con las variables meteorológicas, se evidenció una relación estadísticamente significativa de la concentración de PM_{2.5} con la temperatura, la dirección del viento, la presión atmosférica, la humedad, la radiación solar y la precipitación, lo cual con ayuda de las variables de tiempo ayudó a desarrollar un modelo de pronóstico basado en redes neuronales, que permitiría establecer cuando el contaminante no cumple con los límites sugeridos por la oms en la ciudad de Barranquilla. Por ello, esta investigación sirve como prueba de una herramienta de pronóstico necesaria para la gestión de la calidad del aire en Barranquilla y establece una metodología de análisis del comportamiento de la contaminación del aire para otras ciudades.

Cabe destacar que este estudio corresponde a un primer ejercicio sobre elaboración y construcción de un modelo de pronóstico del incumplimiento en los niveles máximos de PM_{2.5}, y describe de buena manera las relaciones y los comportamientos de las variables en los años 2018 y 2019, siendo esto una limitante si se desea aplicar hoy día

en el año 2023, debido a que pueden presentarse fluctuaciones en la naturaleza de las relaciones estadísticas encontradas. Por tal motivo, se recomienda seguir entrenando la red con más datos a medida que el sistema de monitoreo de la calidad del aire genere más lecturas, de modo que sea posible no solo sostener sino incrementar la capacidad de pronóstico de la red.

Referencias

- [1] oms, Contaminación del aire ambiente (exterior), 2022. [https://www.who.int/es/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health).
- [2] S. A. P. F. A. M. G. Valdivia, "Evaluación espaciotemporal del material particulado PM_{2.5} y su relación con las variables meteorológicas en la Universidad Nacional Agraria La Molina", *Anales Científicos*, vol. 79, no. 2, pp. 334-340, 2018. <https://doi.org/10.21704/ac.v79i2.992>
- [3] Universidad de los Andes, "Ministerio de Ambiente y Desarrollo Sostenible", <https://www.minambiente.gov.co/index.php/component/content/article/1801-plantilla->.
- [4] Alcaldía de Barranquilla, "Reporte semestral calidad de aire", 2021. <https://barranquillaverde.gov.co/storage/app/media/calidad-aire/Reporte%20Calidad%20Aire%20I%20Semestre%202021.pdf>
- [5] IDEAM, "Informe del estado de la calidad del aire en Colombia 2019", 2021. <http://documentacion.ideam.gov.co/openbiblio/bvirtual/023898/023898.pdf>
- [6] G. B. Rolong, A. R., Padilla, y D. A. Castañeda, "Aporte de PM 10 y PM 2.5 en la calidad del aire de Barranquilla por quemas en el Vía Parque Isla de Salamanca", presentada en Congreso Colombiano y Conferencia Internacional de Calidad de Aire y Salud Pública IEEE, pp. 1-6, noviembre, 2021. <https://doi.org/10.1109/CASAP54985.2021.9703404>
- [7] J. Luna-Carrascal, J. Olivero-Verbel, A. J. Acosta-Hoyos, A. J., y M. Quintana-Sosa, "Auto repair workers exposed to PM_{2.5} particulate matter in Barranquilla, Colombia: Telomere length and hematological parameters", *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, vol. 887, pp. 1-10, 2023. <https://doi.org/10.1016/j.mrgentox.2023.503597>
- [8] A. L. Duarte, I. L. Schneider, P. Artaxo, y M. L. Oliveira, "Spatiotemporal assessment of particulate matter (PM₁₀ and PM_{2.5}) and ozone in a Caribbean urban coastal city", *Geoscience Frontiers*, vol. 13, no. 1, pp. 1-9, 2022. <https://doi.org/10.1016/j.gsf.2021.101168>

- [9] D. Agudelo-Castañeda, y D. Prieto, “Estimation of atmospheric emissions from ships in the port of Barranquilla”, publicada en Congreso Colombiano y Conferencia Internacional de Calidad de Aire y Salud Pública (CASP) (pp. 1-8). IEEE. Barranquilla, Colombia, 14-16 agosto de 2019. <https://doi.org/10.1109/CASAP48673.2019.9364030>
- [10] M. Fiuza Pérez, y J. C. Rodríguez Pérez, “La regresión logística: una herramienta versátil” Nefrología, vol. 20, no. 6, pp.495-500, 2000. <https://www.revistanefrologia.com/es-la-regresion-logistica-una-herramienta-articulo-X0211699500035664>
- [11] H. Gutiérrez Pulido, *Análisis y diseño de experimentos*. Ciudad de Mexico, Mexico: McGrawHil, 2012 pp. 62-72.