



**UNIVERSIDAD
DEL AZUAY**

DEPARTAMENTO DE POSGRADOS

**ANÁLISIS DEL COMPORTAMIENTO DE LOS ÍNDICES DE
CALIDAD DEL AIRE EN LA CIUDAD DE CUENCA A TRAVÉS DE
PRONÓSTICOS, UTILIZANDO TÉCNICAS DE APRENDIZAJE
COMPUTACIONAL**

Trabajo de graduación previo a la obtención del título de:

Magister en Matemática Aplicada

Autor:

Ruth Mariela Coronel Alvarado

Director:

Avilés Gonzales Jonnatan Fernando

Cuenca – Ecuador

2023

[

AGRADECIMIENTOS

Quiero expresar mi más sincero agradecimiento a Dios, quien ha sido mi fortaleza y ayuda para superar los momentos difíciles y culminar este proceso. A mi familia por creer en mí, por darme su apoyo incondicional en esta importante etapa de mi formación académica. Ellos han sido mi fuente de inspiración para seguir adelante y trabajar sin descanso.

Asimismo, quiero agradecer a mi tutor quien ha sido una guía indispensable en el proceso de elaboración de este trabajo. Gracias por su experiencia, conocimientos y por su apoyo en todo momento.

DEDICATORIA

Con mucho cariño y agradecimiento, dedico este trabajo a mis padres y hermanos por su amor, por su apoyo en cada etapa de este proceso y por su confianza.

ÍNDICE

| | |
|---|----|
| RESUMEN | 7 |
| INTRODUCCIÓN | 9 |
| Breve Revisión Bibliográfica | 10 |
| Contaminantes Ambientales..... | 10 |
| <i>Tipos de Contaminantes.</i> | 10 |
| <i>Fuentes de Contaminación Atmosférica.</i> | 12 |
| <i>Variables Meteorológicas.</i> | 12 |
| Métodos de Predicción..... | 13 |
| <i>Método ARIMA.</i> | 16 |
| <i>Redes Neuronales Artificiales.</i> | 17 |
| METODOLOGÍA | 18 |
| Preprocesamiento de la Base de Datos..... | 20 |
| Análisis Exploratorio. | 20 |
| Implementación de los Modelos de Predicción..... | 21 |
| Evaluación de los Modelos | 21 |
| RESULTADOS | 22 |
| Procesamiento de la base de datos | 22 |
| Pronóstico de la Concentración del Contaminante O3..... | 23 |
| <i>Estación BALL</i> | 24 |
| <i>Modelo Con Red Neuronal</i> | 31 |
| <i>Resultados, Estación CCA</i> | 35 |
| <i>Resultados, Estación CEB</i> | 35 |
| <i>Resultados, Estación CRB</i> | 35 |
| <i>Resultados, Estación ECC</i> | 36 |
| DISCUSIÓN Y CONCLUSIÓN. | 38 |
| Discusión..... | 38 |
| Conclusiones. | 40 |
| REFERENCIAS. | 41 |
| APÉNDICE | 46 |
| Apéndice A. Gráficas de SCF y PCAF de la Serie Diferencia, Estación BALL..... | 46 |
| Apéndice B. Gráficas y Resultados del Modelo ARIMA, Estación BALL. | 47 |
| Apéndice C. Gráficas y Resultados del Modelo ARIMA y Redes Neuronales, Estación CCA | 49 |

| | |
|---|----|
| Apéndice D. Gráficas y Resultados del Modelo ARIMA y Redes Neuronales, Estación CEB | 55 |
| Apéndice E. Gráficas y Resultados del Modelo ARIMA y Redes Neuronales, Estación CRB | 62 |
| Apéndice F. Gráficas y Resultados del Modelo ARIMA y Redes Neuronales, Estación ECC | 69 |
| Apéndice G. Códigos: Modelo ARIMA | 76 |
| Apéndice H. Código RNA | 80 |

TABLAS Y FIGURAS.

Tablas.

| | |
|--|----|
| Tabla 1. Descripción de los principales contaminantes del aire y sus efectos..... | 11 |
| Tabla 2. Resumen de estudios realizados sobre contaminación atmosférica | 13 |
| Tabla 3. Nombre y código de estaciones | 19 |
| Tabla 4. El número de meses que se registran información en cada estación y de cada contaminante. | 22 |
| Tabla 5. Número de datos en cada estación y fechas en las que no se registran datos. | 23 |
| Tabla 6. En la tabla se presenta los datos pronosticados junto a los datos reales | 29 |
| Tabla 7. Resumen de métricas, modelos ARIMA. | 30 |
| Tabla 8. Datos reales - datos pronosticados | 32 |
| Tabla 9. Resumen de métricas de modelos aplicados. | 33 |
| Tabla 10. Datos pronosticados junto datos reales del mejor modelo ARIMA y el modelo NNA | 34 |
| Tabla No 11. Valor de las métricas por estaciones | 36 |
| Tabla 12. Resumen de los valores pronosticados en cada estación analizada. Contaminante O3 | 37 |
| Tabla 13. Muestra de valores pronosticados. | 38 |
| Tabla. 14. Valores máximos, mínimos y medias del pronóstico por estaciones..... | 39 |
| Tabla No 15. Valor del MASE por estaciones | 39 |

Figuras.

| | |
|---|----|
| Figura 1. Fuentes de contaminación. | 12 |
| Figura 2. Síntesis de resultados de la aplicación de métodos en predicción de contaminantes del aire..... | 15 |
| Figura 3. Localización de las estaciones de la Red de Monitoreo de la Calidad del aire. Cuenca | 18 |
| Figura 4. Resumen de la metodología propuesta | 20 |
| Figura 5. Gráfica data y tendencia. | 24 |
| Figura 6. Gráfico de la medias por año | 25 |
| Figura 7. Descomposición de la serie de tiempo. | 26 |
| Figura 8. Gráfica de ACF y PACF | 27 |
| Figura 9. Serie diferenciada (primera diferencia) | 27 |
| Figura 10. Diagnóstico del modelo..... | 28 |
| Figura 11. Residuales..... | 31 |
| Figura 12. Pronóstico..... | 33 |

ANÁLISIS DEL COMPORTAMIENTO DE LOS ÍNDICES DE CALIDAD DEL AIRE EN LA CIUDAD DE CUENCA A TRAVÉS DE PRONÓSTICOS, UTILIZANDO TÉCNICAS DE APRENDIZAJE COMPUTACIONAL

Ruth M. Coronel A. ruthcoralvalive@com - Jonnatan F. Avilés G. javiles@uazuay.edu.ec

RESUMEN

La gran preocupación existente por la contaminación del aire y sus graves efectos ha llevado a una búsqueda de soluciones para mitigar dichos efectos. Una forma de hacerlo es a través de información que permita prever los comportamientos futuros de los contaminantes y facilite la toma de decisiones oportunas. Con este fin han sido utilizado ampliamente modelos de pronóstico. En el presente estudio, se propone un modelo de predicción para estimar la concentración de uno de los contaminantes del aire. Los modelos se basan en información histórica, y utiliza técnicas de aprendizaje computacional. Se desarrolla un modelo basado en estadística tradicional y otro utilizando algoritmos de Machine Learning. Ambos modelos fueron entrenados con valores de una base de datos proporcionada por el Instituto de Estudios de Régimen Seccional del Ecuador (IERSE). Los resultados muestran que el modelo basado en Redes Neuronales obtuvo las mejores métricas de desempeño en comparación al modelo tradicional (ARIMA)

Palabras Clave: modelos de pronósticos, contaminación del aire, aprendizaje computacional, Machine Learning, Red Neuronal Artificial.

ABSTRACT

The great concern about air pollution and its serious effects has led to search solutions to mitigate these effects. One way to do so is through information that allows forecasting future pollutant behaviors for sound decision making. To this end, forecasting models have been widely used. In this study, a prediction model is proposed to estimate the concentration of one of the air pollutants. The models are based on historical information, and use computational learning techniques. One model was developed based on traditional statistics and another using Machine Learning algorithms. Both models were trained with values from a database provided by the Instituto de Estudios de Régimen Seccional del Ecuador (IERSE). The results show that the model based on Neural Networks obtained the best performance metrics compared to the traditional model (ARIMA).

Keywords: forecasting models, air pollution, computational learning, Machine Learning, Artificial Neural Network.

