

Implementación de una red neuronal y un modelo de forrester para la predicción entre factores demográficos y agentes contaminantes

Implementation of a neural network and a forrester model for the prediction between demographic factors and pollutants

G. Suárez- Guerrero  ; J. A. Arias-Franco  ; E. Vallejo-Morales  ; L.J Cruz-Riaño 

DOI: <https://doi.org/10.22517/23447214.25177>

Artículo de investigación científica y tecnológica

Abstract— The greatest problems that has been presented in the world, especially in Colombia, is due to an accelerated population growth generated by the overwhelming increase in migrant circulation or mobility of people due to other causes or eventual situations of nature. These social situations have been significantly affecting the urban order of cities, particularly due to the confinement of the population and the increase in environmental pollution that this increase brings, in addition to unhealthiness, poverty, among other factors. To study population growth based on demographic variables and environmental factors, an Artificial Neural Network was built to carry out a data analysis that incorporated related variables such as CO₂ and other pollutants. The purpose was to identify the influence of these toxic agents with the growth of a population. Once the dynamic prediction was verified through the neural network, a mathematical model was built to study a specific case of demographic behavior for a certain Colombian region by means of a Forrester Diagram, where the migratory conditions, birth rate, morbidity, mortality, migratory flow (or mobility of people due to other natural events) and pollutants that affect human health mainly in respiratory diseases. The results delivered predictions and the Forrester Diagram confirming the relationship between toxic agents and demographic aspects.

Index Terms— CO₂, demographic behavior, environmental pollution, Forrester diagram, mathematical model, migration flow, neural network, pm2.5, Soacha.

Resumen— Uno de los grandes problemas que se vienen presentando en el mundo, especialmente Colombia, es debido a un acelerado crecimiento poblacional generado por el aumento desbordado de la circulación migrante o movilidad de personas por otras causas o situaciones eventuales de la naturaleza. Estas situaciones sociales vienen afectando sensiblemente el orden

urbanístico de las ciudades, particularmente por el confinamiento de la población y el aumento de la contaminación ambiental que aporta este incremento, además de la insalubridad, pobreza, entre otros factores. Para estudiar el crecimiento poblacional en función de variables demográficas y factores ambientales, se construyó una Red Neuronal Artificial para llevar a cabo un análisis de datos que incorporó variables relacionadas tales como CO₂ y otros agentes contaminantes. El propósito fue identificar la influencia de estos agentes tóxicos con el crecimiento de una población. Una vez verificada predicción dinámica mediante la red neuronal, se construyó un modelo matemático para estudiar un caso específico del comportamiento demográfico para una región colombiana por medio de un Diagrama de Forrester bajo condiciones migratorias, natalidad, morbilidad, mortalidad, flujo migratorio (o movilidad de personas debido a otros eventos de la naturaleza) y agentes contaminantes que afectan la salud humana principalmente en enfermedades respiratorias. Los resultados entregaron predicciones y del Diagrama de Forrester constatando relación entre los agentes tóxicos y en los aspectos demográficos.

Palabras claves— CO₂, comportamiento demográfico, contaminación ambiental, diagrama de Forrester, flujo migracional, modelo matemático, pm2.5, red neuronal, Soacha.

I. INTRODUCCIÓN

EN esta investigación, se presenta una estructura científica para el desarrollo problema demografía-contaminación conformada por una breve descripción del estado del arte, marco teórico y una metodología que incorpora una amplia descripción matemática y computacional basada en el desarrollo de una red neuronal y una solución de ecuaciones diferenciales por técnicas numéricas.

Este manuscrito fue enviado el 02 de junio del 2021 y aceptado el 09 de marzo de 2023.

G. Suárez- Guerrero Grupo de Investigación e Innovación en Energía, Facultad de Ingeniería, Institución Universitaria Pascual Bravo, Medellín, Colombia. e-mail: gustavo.suarez@pascualbravo.edu.co.

J. A. Arias-Franco Grupo de Investigación Sobre Nuevos Materiales, Escuela de Ingeniería, Universidad Pontificia Bolivariana, Medellín, Colombia. E-mail: julian.ariasf@upb.edu.co

E.Vallejo-Morales Grupo de Investigación Sobre Nuevos Materiales, Escuela de Ingeniería, Universidad Pontificia Bolivariana, Medellín, Colombia. E-mail: esteban.vallejomo@upb.edu.co

L.J Cruz-Riaño Grupo de Investigación Sobre Nuevos Materiales, Escuela de Ingeniería, Universidad Pontificia Bolivariana, Medellín, Colombia. E-mail: luis.cruz@upb.edu.co.



A nivel mundial, es indiscutible la aceleración del crecimiento demográfico que han venido presentando las ciudades en las últimas cinco décadas. Comunicados oficiales por las Naciones Unidas afirman que este fenómeno es debido al aumento del número de personas que sobreviven hasta alcanzar una etapa reproductiva, influyendo en cambios marcados en las tasas de natalidad. Esto sin duda alguna, produce flujos migratorios en los distintos continentes que desencadenan problemáticas en los planes de ordenamiento territorial de cada nación [1].

Como consecuencia del aumento en la población, los niveles de agentes contaminantes como las emisiones de: CO₂ + pm2.5. Estos agentes contaminantes han venido presentando un aumento de toxicidad en la atmosfera. Además es un gas que interviene en el fenómeno de efecto invernadero antropogénico y que se encuentra en abundante concentración en la atmosfera [1]. Son producidos principalmente por emisiones fósiles afectando la salud de la población humana y generando como consecuencia las Infecciones Respiratorias Agudas (IRA).

La principal afectación de la salud por CO₂ es la asfixia por la reducción de oxígeno por debajo del 20%. Puede generar dolores de cabeza, falta de concentración, somnolencia, mareos y problemas respiratorios [2], afectando primordialmente a personas de la tercera edad y niños en sus primeros años de vida.

Por su parte, el agente contaminante pm2.5, viene generando altas tasas de morbilidad y mortalidad en las ciudades, incluidas ciudades de China, especialmente por enfermedades de tipo respiratorio y cáncer pulmonar [2, 3, 4].

En china, se están midiendo las concentraciones de carbono y partículas (pm1, pm2.5 y pm10) con exposición en niños durante sus viajes multimodales (caminar, automóviles privados y bicicletas eléctricas), mostrando altos niveles de absorción y efectos perjudiciales en sus sistemas respiratorios [3]. Además los agentes contaminantes CO₂ y pm2.5, interactúan entre sí ya que comparten el mismo ambiente con la población, acrecentado la problemática [4]. Un estudio desarrollado por la *United States Environmental Protection Agency* (EPA), obtuvo unos niveles de contaminación 5 veces mayores en zonas interiores que exteriores debido al confinamiento de gases en espacio cerrados donde las personas permanecen un 90% mayor de su tiempo [5].

Por otro lado, el flujo migratorio o movilidad de personas por otras causas o situaciones eventuales de la naturaleza, es un fenómeno que viene intensificando las problemáticas sociales y medioambientales de las ciudades en el mundo, ya que estas interactúan con crecimiento de la población y la salubridad.

Colombia es uno de los principales países del mundo donde se viene presentando grandes movilizaciones de personas que afectan los cambios ambientales de las zonas ocupadas y las condiciones sociales.

En esta investigación, Se construyó una *Red Neuronal Artificial* para predecir interacción entre los agentes contaminantes y el crecimiento imprevisto de la población. Se aplicó el concepto

de *aprendizaje automático* para elaborar un análisis de datos utilizando una *Red Neuronal Artificial* (ANN) [6, 7, 8]. Además, desarrolló un *Diagrama de Forrester de la Dinámica de Sistemas*, con el propósito de determinar los valores de incidencia de los agentes contaminantes y población, especialmente el pm2.5. que es un toxico más mortal en un caso específico de una región colombiana.

II. MATERIALES Y MÉTODOS

Se elaboró una metodología para la implementación de la Red Neuronal Artificial y la elaboración del *Diagrama Forrester de la Dinámica de Sistemas*, como se muestra a continuación Fig.1.

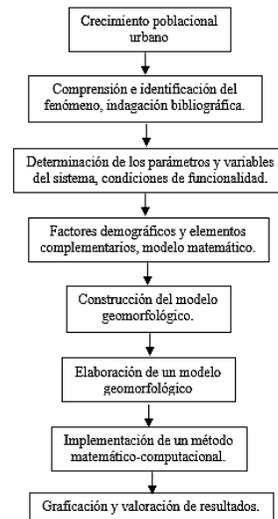


Fig. 1. Metodología para el estudio de la interacción entre agentes contaminantes, crecimiento demográfico y movilidad de personas (migración).

El primer desarrollo de la Red Neuronal validó una predicción generalizada mediante un algoritmo entrenado por medio un conjunto de datos determinados de registros públicos sobre valores de agentes contaminantes (CO₂, pm2.5) y crecimiento demográfico (tasas de natalidad, tasas de mortalidad y migración neta), utilizando un proceso de entrenamiento supervisado [6, 7, 8], y la construcción de un *Diagrama Forrester de la Dinámica de Sistemas* permite el estudio de un caso especial (municipio de Soacha) incorporando las variables consideradas en la Red Neuronal, una vez comprobada predicción relacional entre las variables poblacionales.

A. Análisis de la relación existente entre variables de contaminación ambiental y el crecimiento poblacional

Se comienza con la identificación si las variables de los agentes contaminantes, principalmente de mayor cantidad (CO₂), tienen relación con el incremento o decremento poblacional.

Se desarrolló el entrenamiento de la red neuronal habiendo realizado las correspondientes comparaciones entre los valores predichos y los valores reales. De esta forma, se pretendió constatar la efectiva relación entre los agentes contaminantes y los niveles de población en un territorio. A partir de esta

premisa, Se construyó el Diagrama Forrester cuyo modelo matemático estudió el comportamiento del crecimiento demográfico (entre las cuales está la movilidad migrante) y la interacción con la contaminación, considerando como caso particular de análisis el municipio de Soacha.

B. Modelamiento matemático de una red neuronal para la predicción del incremento o decremento poblacional

Para construir una red neuronal se implementaron los modelos matemáticos (1), (2), (3) y (4) los cuales se describen a continuación [6, 7, 8]:

$$\delta^L = (a^L - t^L) \odot \sigma'(Z^L) \tag{1}$$

$$\delta^l = (W^{l+1} \delta^{l+1}) \odot \sigma'(Z^l) \tag{2}$$

$$\frac{\partial J}{\partial w^l} = a^{l-1} (\delta^l)^T \tag{3}$$

$$\frac{\partial J}{\partial b^l} = \delta^l \tag{4}$$

Donde el superíndice L hace referencia a la última capa y el superíndice l a cualquier capa anterior a L .

En cuanto a la actualización de los factores de peso y modelo bias, se describen respectivamente mediante por medio de (5) y (6) [6, 7, 8]:

$$W^l = W^l - \sum \frac{\partial J}{\partial w^l} \tag{5}$$

$$b^l = b^l - \sum \frac{\partial J}{\partial b^l} \tag{6}$$

El modelo (1) entregó los cálculos correspondientes a la última capa de la red neuronal, por su parte el modelo (2) entregó los resultados obtenidos en las capas anteriores a la última. Con los modelos (3) y (4), se determina el ajuste necesario en los factores de peso y el sesgo respectivamente.

Se entreno la red neuronal minimizando la función de error o costo mediante el ajuste de los factores de peso que conectan las neuronas del modelo [6, 7, 8]. Se aplicó como función de costo el error cuadrático medio el modelo (7) [6, 7, 8]:

$$J = \sum_{j=1}^n \frac{1}{2} (t_j - y_j)^2 \tag{7}$$

Donde J es el error cuadrático medio, t_j corresponde al valor de salida conocido para la neurona j de la última capa de la red neuronal y y_j es el valor de salida calculado por la red neuronal para la neurona j , Fig. 2:

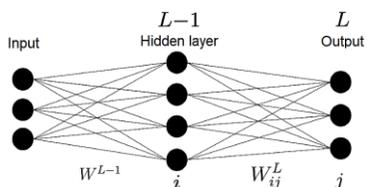


Fig. 2. Arquitectura de la red neuronal.

De acuerdo con la Fig. 1, los factores de peso entre la capa $L - 1$ y L son representados a través de la siguiente matriz (8):

$$W_{ij}^L = \begin{bmatrix} W_{11} & W_{12} & W_{13} \\ W_{21} & W_{22} & W_{23} \\ W_{31} & W_{32} & W_{33} \\ W_{41} & W_{42} & W_{43} \end{bmatrix} \tag{8}$$

Mientras que los valores de activación las neuronas de la capa $L - 1$ y los valores de bias para la capa L se representan de la siguiente manera en el modelo (9) y (10):

$$a_i^{L-1} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \tag{9}$$

$$b_j^L = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \tag{10}$$

Con los modelos (1), (2), (3) y (4), se obtuvieron los cambios de los diferentes Factores de peso establecidos en la red neuronal de tal manera que, tras pasar por las correspondientes funciones de activación en cada capa de la red, la función de costo sea minimizada y, por lo tanto, el valor predicho se acerque al valor de salida real.

C. Implementación

Con base en la descripción matemática descrita anteriormente, se construyó un modelo predictivo del crecimiento poblacional de Colombia con base en una historia de 61 años a través de una Red Neuronal Artificial (ANN) con aprendizaje supervisado, con el propósito de predecir la población total en función de variables ambientales y demográficas.

Se construyó una base de datos utilizando la información disponible en los sitios web de DANE [9] y el Banco Mundial [10]. Con esta información fue posible depurar un conjunto de datos que incluyeran tanto variables demográficas como indicadores de emisiones de CO₂ en Colombia entre los años 1960 y 2020.

Variables de entrada del modelo

En la Tabla I, se describen las variables que fueron tenidas en cuenta para el correspondiente entrenamiento del modelo. Se hace un énfasis especial en las variables que conciernen a los indicadores de emisión principalmente de CO₂ que está en mayor cantidad.

Como variable dependiente, es decir, variable a predecir, se ha seleccionado “Población, total” mientras que las variables

restantes son consideradas predictoras y constituyen el conjunto de campos utilizado como entrada a la red neuronal. Cada registro en la base de datos corresponde a un año entre 1960 y 2020, lo que da como resultado una muestra de 60 registros para realizar el estudio.

TABLA I
VARIABLES DEMOGRÁFICAS Y DE CONTAMINACIÓN POR CO₂

Variable	Descripción
Emisiones de CO ₂ por consumo de combustibles gaseosos (kt)	Las emisiones de dióxido de carbono por el consumo de combustibles líquidos se refieren principalmente a las emisiones por el uso de gas natural como fuente de energía.
Emisiones de CO ₂ por consumo de combustibles sólidos (kt)	Las emisiones de dióxido de carbono por el consumo de combustibles sólidos se refieren principalmente a emisiones derivadas del uso del carbón como fuente de energía.
Emisiones de CO ₂ del transporte (% de la quema total de combustible)	Las emisiones de CO ₂ del transporte contienen las emisiones de la combustión de combustible para toda la actividad de transporte, independientemente del sector, excepto para los búnkeres marítimos internacionales y la aviación internacional.
Emisiones de CO ₂ del transporte (% de la quema total de combustible)	Las emisiones de CO ₂ de la producción de electricidad y calor son la suma de tres categorías de emisiones de CO ₂ : (1) Electricidad y calor del principal productor de la actividad, (2) auto productores no asignados que contienen emisiones de la generación de electricidad y/o calor por parte de los auto productores. (3) Otras industrias energéticas que contienen emisiones de combustibles quemados en refinerías de petróleo, para la fabricación de combustibles sólidos, extracción de carbón, extracción de petróleo y gas y otras industrias productoras de energía.
Emisiones de CO ₂ por consumo de combustibles sólidos (% del total)	Las emisiones de dióxido de carbono por el consumo de combustibles sólidos se refieren principalmente a las emisiones por el uso del carbón como fuente de energía.
Emisiones de CO ₂ de otros sectores, excluyendo edificios residenciales y servicios comerciales y públicos (% de la quema total de combustibles)	Las emisiones de CO ₂ de otros sectores, menos los edificios residenciales y los servicios comerciales y públicos, contienen las emisiones de actividades comerciales/institucionales, residenciales, agrícolas/silvícolas, pesca y otras emisiones no especificadas en otra parte que están incluidas en las categorías de fuente/sumidero del IPCC 1 A 4 y 1 A 5. En las Directrices del IPCC de 1996, la categoría también incluye emisiones de auto productores en los sectores comercial/residencial/agrícola que generan electricidad y/o calor. Los datos de la IEA no se recopilan de manera que permitan dividir el consumo de energía por uso final específico y, por lo tanto, los auto productores se muestran como un elemento separado (auto productores no asignados).
Tasa de natalidad, bruta (por 1.000 personas)	La tasa bruta de natalidad indica el número de nacidos vivos que ocurren durante el año, por cada 1.000 habitantes estimados a mediados de año. Al restar la tasa bruta de mortalidad de la tasa bruta de natalidad se obtiene la tasa de aumento natural, que es igual a la tasa de cambio de población en ausencia de migración.
Tasa de mortalidad, bruta (por cada 1.000 habitantes)	La tasa bruta de mortalidad indica el número de muertes que ocurren durante el año, por cada 1.000 habitantes estimado a mediados de año. Al restar la tasa bruta de mortalidad de la tasa bruta de natalidad se obtiene la tasa de aumento natural, que es igual a la tasa de cambio de población en ausencia de migración.
Migración neta	La migración neta es el total neto de migrantes durante el período, es decir, el número total de inmigrantes menos el número anual de emigrantes, tanto ciudadanos como no ciudadanos.
Población, total	La población total cuenta a todos los residentes independientemente de su estatus legal o ciudadanía.

Los valores que se muestran son estimaciones de mitad de año

Bibliotecas aplicadas

Para la implementación de la solución se utilizaron un conjunto de bibliotecas de Python para facilitar la manipulación, visualización y tratamiento de los datos que se presentan en la Tabla II.

TABLA II
BIBLIOTECAS UTILIZADAS

Nombre de la librería	Abreviatura	Uso
Pandas	pandas	Manipulación y análisis de datos
Numeric Python	numpy	Manipulación de matrices y análisis numérico
Scikit-learn	sklearn	Herramientas para programar redes neuronales
Keras Tensorflow	keras	Herramienta para el manejo de redes neuronales de aprendizaje profundo
Matplotlib	matplotlib	Generación de gráficas

La biblioteca Keras Tensorflow permite crear instancias de redes neuronales que utilizan como principio matemático el modelo explicado anteriormente.

Estructura de la red neuronal

La arquitectura utilizada para la ANN se construyó para un total de nueve neuronas en la capa de entrada, una capa oculta de cuatro neuronas y una capa de salida con el valor predicho de población total. La disposición de estos elementos se puede evidenciar en la Fig. 3.

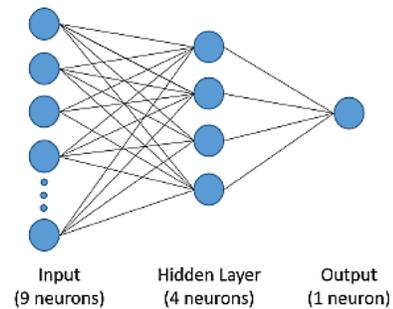


Fig. 3. Estructura de la red neuronal.

En la primera y segunda capa se configuró una función de activación ReLU, mientras que en la capa de salida se implementó una función lineal. El 80% de los datos se utilizó para entrenar la red neuronal y el 20% restante para evaluar el modelo.

D. Resultados de la predicción del total de población a través de la red neuronal entrenada.

En la Fig. 4 se pueden apreciar las diferentes superposiciones entre el modelo entrenado y los datos reales obtenidos de las

fuentes mencionadas anteriormente. De acuerdo con estas superposiciones que se ajustan en gran medida a los datos reales se comprueba que las variables estudiadas tienen relevancia como elementos predictivos y valiosos para modelar el crecimiento poblacional a través de un análisis que utilice como herramienta principal la dinámica de sistemas.

El error absoluto medio obtenido tras la evaluación del modelo indica una favorable capacidad de la ANN entrenada para ajustarse al comportamiento de los datos recolectados; este valor corresponde a 0,038.

La ausencia de una superposición perfecta genera confianza en que el modelo no fue sobreentrenado, lo que le permite recibir nuevos datos y predecir si sesgo el correspondiente valor de salida.

Los resultados de este estudio fueron el punto de partida para comenzar a estudiar el modelo de dinámica (*Diagrama Forrester*) del sistema de crecimiento poblacional con la confianza de que las variables de emisión de CO₂ (como agente principal) tenían un valor predictivo y un impacto en el comportamiento poblacional, esto debido al ajuste entre los valores reales y los valores predichos con una discrepancia o error cuadrático medio con un valor bajo.

A continuación, se presentan los elementos tenidos en cuenta en el análisis de la relación existente entre variables de contaminación ambiental y el crecimiento poblacional a través de la construcción de una red neuronal artificial.

III. ESTUDIO DEL CRECIMIENTO POBLACIONAL A TRAVÉS DE LA DINÁMICA DE SISTEMAS

Una vez validada la relevancia de los factores de contaminación tales como el CO₂, como agente principal, a continuación, se procedió con la implementación del estudio del crecimiento poblacional en función de un conjunto de nuevas variables (incorporando la pm2.5) relacionadas con más detalle en la Fig. 5.

El Diagrama de Forrester fue implementado con base en conceptos de dinámica de sistemas, el cual permite dar cuenta de su estructura y comportamiento a través de una serie de ecuaciones diferenciales que describen comportamiento y que pueden ser procesadas de manera computacional con el método de Euler [11].

Este modelo está basado en métodos de la dinámica de sistemas, que asume información con respecto de cómo se producen las interacciones en el seno del sistema. Esta información permite conceptualizar la realidad del problema en estudio, con el fin de lograr la construcción del diagrama de Forrester [12] descrito en la Fig. 5.

Se estudió un caso especial en Colombia (Soacha), ya que se cuentan con todos los registros públicos necesarios [13, 14].

En este análisis se incluyen factores de flujos migratorios y ambientales, entre otros, donde estos últimos son determinantes para predecir las muertes que se producen por la contaminación ambiental producto de agentes contaminantes como el pm2.5 (principalmente por mayor mortalidad) y CO₂ [15, 16, 17].

Se construyó el Diagrama de Forrester con los flujos de entrada y salida de datos, como se presenta a continuación la Fig. 5.

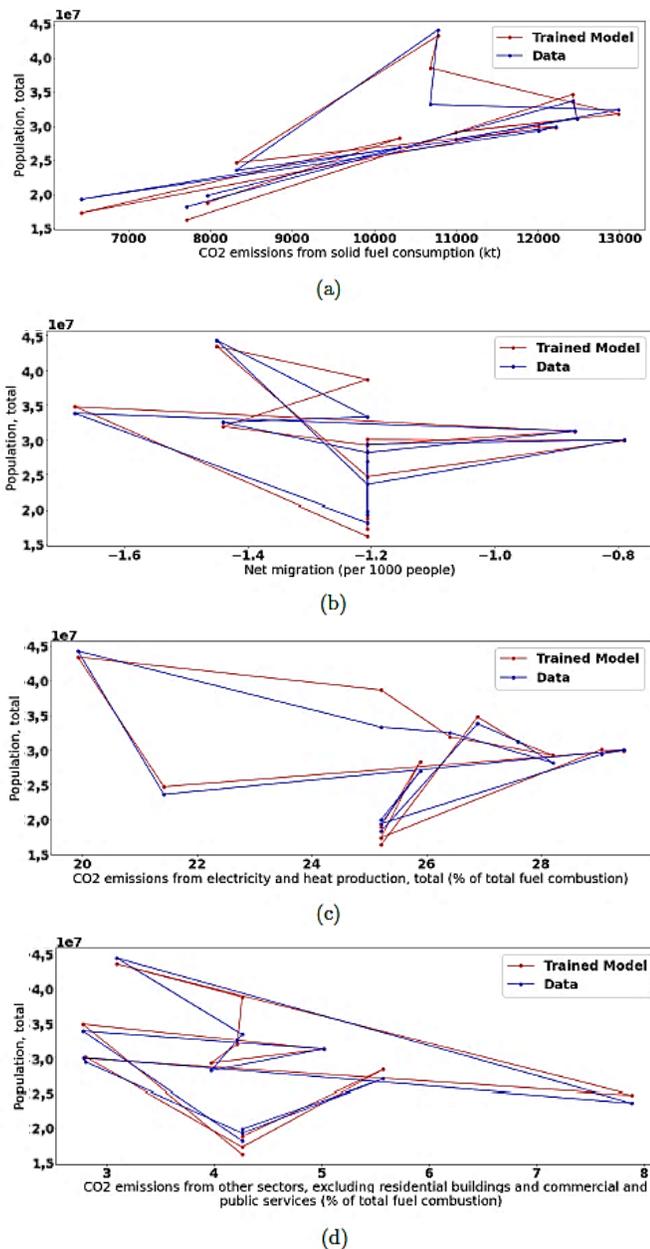


Fig. 4. Relación ANN entre Población, total y a) Emisiones de CO₂ de combustibles sólidos (kt), b) Migración neta, c) Emisiones de CO₂ de la producción de electricidad y calor, total (% de la quema total de combustible), d) emisiones de CO₂ de otros sectores, excluyendo residencial edificios y servicios comerciales y públicos (% sobre la quema total de combustibles).

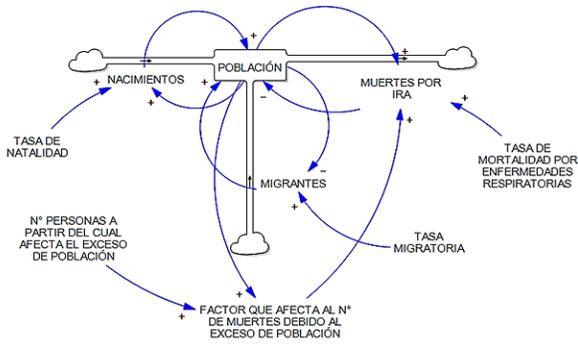


Fig. 5. Diagrama de Forrester del modelo poblacional.

A. Comportamiento Matemático del Modelo

La dinámica de sistemas permite expresar matemáticamente la variación de una variable de nivel, determinada por la siguiente ecuación diferencial en (11) [18, 19, 20]:

$$\frac{dN(t)}{dt} = \sum_i FE_i(t) - \sum_j FS_j(t) \quad i, j = 1, 2, 3, \dots \quad (11)$$

Donde,

t = tiempo.

FE_i = flujo de entrada de i .

FS_j = flujo de salida de j .

Este mismo sistema de ecuaciones puede expresarse como:

$$N(t) = N(0) + \int_0^t \left[\sum_i FE_i(t) - \sum_j FS_j(t) \right] dt \quad i, j = 1, 2, 3, \dots \quad (12)$$

Se puede evidenciar que en el modelo (12) las variables de nivel acumulan, a lo largo del tiempo el saldo neto de los valores de sus flujos. Para solucionar analíticamente estos sistemas de ecuaciones, es posible emplear diferentes métodos numéricos, que para el caso de la investigación se empleó el método de Euler y el método de Runge Kutta de cuarto orden, para verificar la existencia de coherencia en su desarrollo matemático.

B. Implementación

Para la implementación del modelo matemático que representa el comportamiento del crecimiento poblacional, se consideraron las variables de nacimientos y muertes de la población por Infección Respiratoria Aguda (IRA). De los registros oficiales del último censo nacional (DANE-2018) [15], se obtuvo el número de habitantes con un registro de 645205. Los datos de la tasa de natalidad (0,04), la tasa de mortalidad por IRA (0,0046) y tasa migratoria (0,009), fueron obtenidos de la plataforma oficial del municipio de Soacha y la institución de migración Colombia.

Se construyó un modelo matemático que describe el comportamiento transitorio demográfico del diferencial entre los nacimientos y muertes, usando (13) [18, 20]:

$$\frac{dPOB(t)}{dt} = NAC(t) - MU(t) \quad (13)$$

El comportamiento de los nacimientos de la población, se definen mediante (14) [18, 20]:

$$NAC(t) = TN * POB(t) \quad (14)$$

El comportamiento de las muertes por Infecciones Respiratorias Agudas, se describe mediante (15) [18, 20]:

$$MU(t) = TM * POB(t) \quad (15)$$

El comportamiento migratorio poblacional, se describe mediante (16) [18, 20]:

$$MI(t) = TMI * POB(t) \quad (16)$$

MI = migración de la población no nativa

TMI = tasa migratoria de la población no nativa

La tasa de mortalidad por Infección Respiratoria Aguda (IRA) puede ser calculada mediante (17) [18, 20]:

$$TM = \frac{\text{número de muertes por IRA}(MU)}{POB} * 1000 \quad (17)$$

La tasa de mortalidad se constituye en un indicador fundamental para la medición de la mortalidad ya que considera la relación entre el “número de defunciones ocurridas en un periodo de tiempo y una estimación de la población expuesta al riesgo de morir” [16].

Para el estudio aplicado al municipio de Soacha, se conocen registros del número de muertes anuales 2018, por IRA de 3 habitantes, siendo la tasa calculada con (18) [18, 20]:

$$TM = \frac{3}{645205} * 1000 = 0,0046 \quad (18)$$

La tasa de natalidad se logra describir mediante (19) [18, 20]:

$$TN = \frac{\text{número de nacimientos}}{POB} * 1000 \quad (19)$$

Sin embargo, el valor de número de muertes de 3 habitantes, no fue un dato verificado directamente del sistema de salud pública del municipio. Otros estudios han indicado que el número de muertes por IRA fue superior a 50 habitantes anuales, esto debido al acelerado crecimiento industrial, comercial y parque automotor, situación que ha incrementado los factores de contaminación ambiental [17, 18, 19].

La tasa de mortalidad se constituye en un indicador fundamental para la medición de la mortalidad ya que considera la relación entre el “número de defunciones ocurridas en un periodo de tiempo y una estimación de la población expuesta al riesgo de morir” [16].

La ecuación (19) describe un modelo para un comportamiento lineal de la variable de muertes por IRA y crecimiento poblacional (MU), sin embargo, este modelo matemático no

considera un aumento significativo de la población que también incide en el incremento de las muertes por IRA [20, 21, 22].

Este fenómeno puede ser descrito de manera más apropiada mediante un comportamiento matemático no lineal justificado a partir de limitaciones que tienen los sistemas de salud pública en los municipios para la atención de IRA, y suscitado por un incremento desbordado de la mala calidad del aire que pueden conllevar a emergencias sanitarias [20, 23, 24].

A continuación, se considera un nuevo término matemático en (19), descrito por el modelo transitorio para (20) [20, 23, 24]:

$$FAMU(t) = f = \left(\frac{POB(t)}{NMAX} \right) \quad (20)$$

Donde,

$NMAX$ = número de personas a partir del cual afecta el exceso de la población al número de muertes

$FAMU$ = factor que afecta al número de muertes debido al exceso de población

La ecuación mejorada para el cálculo del número de muertes por IRA, se pueden describir así en (21) [20, 23, 24].:

$$MU(t) = TM * POB(t) * FAMU(t) \quad (21)$$

Se realizaron los cálculos considerando una predicción de población con 950.000 habitantes, considerando valores iniciales para $f = \left(\frac{POB(t)}{NMAX} \right)$ y para $FAMU$, y se implementó un modelo de regresión polinomial no lineal para obtener los demás valores.

El modelo se describe en (22) de la siguiente manera [20, 25, 26, 28]:

$$\frac{\partial S_t}{\partial a_0} = -2 \sum (y_i - a_0 - a_1 x_i - a_2 x_i^2) \quad (22)$$

$$\frac{\partial S_t}{\partial a_1} = -2 \sum x_i (y_i - a_0 - a_1 x_i - a_2 x_i^2) \quad (23)$$

$$\frac{\partial S_t}{\partial a_2} = -2 \sum x_i^2 (y_i - a_0 - a_1 x_i - a_2 x_i^2) \quad (24)$$

Donde:

$$y_i = FAMU_i$$

$$x_i = f_i$$

S_i = la función de la ecuación i .

Este método permitió corroborar la precisión el método de Taylor de orden cuatro y puede deducirse emparejando los coeficientes anteriores con los del método de Taylor de orden cuatro de manera que el error local sea de orden $O(h^5)$ [27, 28].

Las simulaciones fueron mediante los registros oficiales los cuales se pueden observar de la siguiente tabla III y IV [3]:

TABLA III
VARIABLES Y PARÁMETROS

Simulación	Tiempo inicial	Tiempo final	Población inicial	N. personas a partir del cual afecta el exceso de población	Tasa natalidad	Tasa mortalidad por enfermedades respiratorias	Tasa migratoria
1	2018	2050	645205	950000	0,04	0,0046	0,009
2	2018	2050	671155	950000	0,045	0,0059	0,009
3	2020	2050	398295	950000	0,04	0,007	0,009
4	2018	2050	645205	950000	0,11	0,08	0,009
5	2018	2060	671155	950000	0,10	0,077	0,009

Datos numéricos de las variables endógenas y exógenas, empleados en cada una de las simulaciones.

TABLA IV
RANGO DE CONCENTRACIÓN POR GASES CONTAMINANTES

Contaminante	Tiempo de exposición	Prevención	Alerta	Emergencia
pm2.5	24 horas	38-55	56-150	≥ 151
CO ₂	8 horas	10820-14254	14255-17688	≥ 17689

Rango de concentración [$\mu\text{g}/\text{m}^3$] para la declaratoria de los niveles de prevención, alerta o emergencia para pm2.5 y CO₂.

IV. RESULTADOS DE LA IMPLEMENTACIÓN, ANÁLISIS Y DISCUSIÓN

A continuación, se muestran los gráficos de los resultados en las fig. 6, 7, 8 y 9:

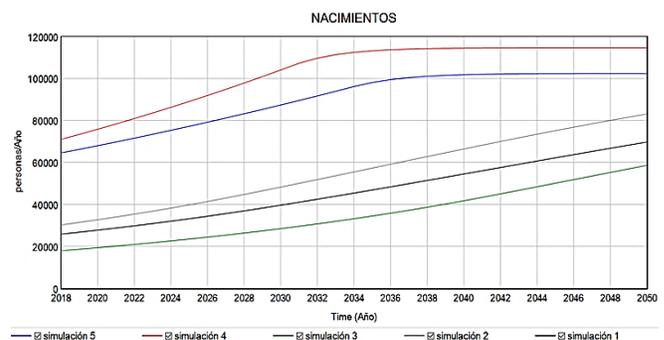


Fig. 6. Gráficas del número de nacimientos/año, incorporando en la simulación 4 y 5 factores ambientales e incremento subido poblacional.

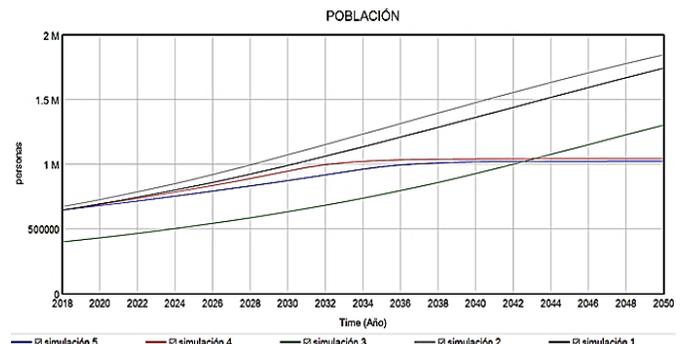


Fig. 7. Gráficas del número total de habitantes, incorporando en la simulación 4 y 5 factores ambientales e incremento subido poblacional.

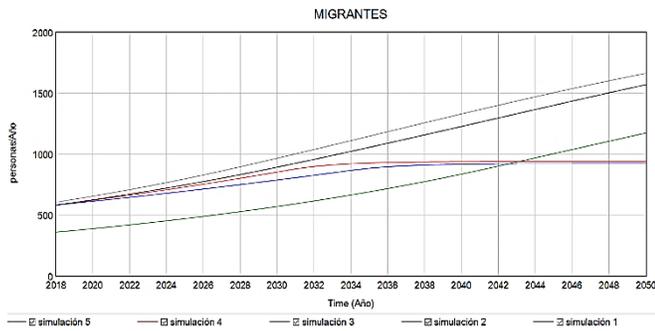


Fig. 8. Gráficas del número de movilidad de personas, incorporando en las simulaciones 4 y 5 factores ambientales.



Fig. 9. Gráficas del número de muertes por Infecciones Respiratorias Agudas, incorporando en las simulaciones 4 y 5 factores ambientales.

A. Análisis de Resultados

Con base en la predicción que, valido la relación entre los agentes contaminantes y el crecimiento imprevisto de la población, se resolvieron las ecuaciones diferenciales del *Diagrama de Forrester* que representó la dinámica del sistema.

En la Fig. 6, se obtuvieron los resultados graficas (sim.1, sim.2, sim.3) que indicaron un comportamiento de la natalidad con tendencia lineal ascendente para una proyección entre año 2018 y 2050. Sin embargo, al incorporar factores ambientales (pm2.5 y Co2), y un crecimiento súbito poblacional las los resultados gráficos (sim.4 y sim.5) no mostraron un aumento ascendente significativo de la natalidad. Esto demuestra lo obtenido de la red neuronal, que es la incidencia del factor ambiental en el sistema. El fenómeno puede ser producto del mismo problema de aumento poblacional que genera mayores índices contaminantes y se puede reflejar de igual manera en [25, 26].

En la Fig. 7, se obtuvieron los resultados graficas (sim.1, sim.2, sim.3) que indicaron un comportamiento del crecimiento del número de habitantes con tendencia ascendente para una proyección entre año 2018 y 2050. Sin embargo, al incorporar factores ambientales (pm2.5 y Co2), y el crecimiento se estabilizó, especialmente entre el año 2032 a 2050 (sim.4 y sim.5). Nuevamente se ve reflejado en los resultados la incidencia del factor ambiental en el crecimiento de la población.

En la Fig. 8, se obtuvieron los resultados graficas (sim.1, sim.2, sim.3) que indicaron un comportamiento del número de movilidad de personas con tendencia ascendente para una proyección entre año 2018 y 2050. Sin embargo, al incorporar factores ambientales (pm2.5 y Co2), y el crecimiento se estabilizó hasta el año 2034 y luego tuvo un leve descenso, especialmente entre el año 2032 a 2050 (sim.4 y sim.5). Se ve reflejado a incidencia de los factores tóxicos y otros factores de morbilidad como consecuencia del aumento de los problemas respiratorios, que puede verse reflejado en [27].

En la Fig. 9, se obtuvieron los resultados graficas (sim.1, sim.2, sim.3) que indicaron un comportamiento de muertes por IRA con tendencia leve y una afectación poblacional baja (máximo 40.000/año 2050), sin embargo, al incorporar agentes contaminantes, especialmente el pm2.5 que generan mayor toxicidad, la proyección de personas fallecidas por causas IRA Entre año 2018 y 2050 se aumenta significativamente hasta alcanzar valores extremos de 120.000 personas en el año 2050. Este último análisis del desarrollo del modelo dinámico entre agentes contaminantes, demografía y movilidad imprevista de flujo de personas demuestra efectivamente una alta afectación de la salud humana por causas ambientales y que puede verse incrementado por los mismos flujos migratorios, a través del análisis de estos factores críticos [28].

En Colombia el Departamento Nacional de Planeación (DNP), presentó que 10527 muertes y 67,8 millones de síntomas y enfermedades estuvieron relacionadas a este fenómeno ambiental. Además, los últimos informes del estado de la calidad del aire en Colombia, informa que el contaminante con mayor potencial de afectación en la salud es el Material Particulado Menor a 2,5 micras (pm2.5).

El modelo fue implementado con las dos técnicas de Euler y Runge Kutta, donde se logra verificar la coherencia de los resultados en cada una de las simulaciones, pero se toma como principal técnica el método de Euler en la solución del sistema, en donde la Dinámica de Sistemas permitió la modificación de valores específicos principalmente de las tasas de natalidad, mortalidad y migratoria, concluyendo que el modelo fue numéricamente insensible a las variaciones de los parámetros en ambas técnicas.

Estos resultados obtenidos de esta investigación fueron comparados con otros tratados que también estudiaron la incidencia de la contaminación ambiental en la demografía. Se encontraron otros desarrollos, pero utilizando modelos de estudio estadísticos. Se obtuvieron similitudes en los resultados obtenidos con la región de Tarento, Italia [29], con lo cual esta comparación científica le brinda una validez a la modelación e implementación desarrollada en esta investigación.

V. CONCLUSIONES

Los resultados obtenidos de las simulaciones demostraron una relación entre las variables dinámicas del sistema debido al crecimiento desbordado de las poblaciones, producto de movilizaciones imprevistas por causas sociales, ambientales o

económicas. Los sistemas de salud no se encuentran preparados para este tipo de circunstancias en el mundo, lo que ocasiona mayores enfermedades de tipo respiratorio.

Los modelos predictivos permiten anticipar comportamientos sociales, ambientales y de salud que pueden afectar las economías de los países, el orden urbano, la seguridad, entre otros factores demográficos. Las herramientas de aprendizaje automático podrían contribuir globalmente a determinar estrategias, normas y leyes que conducen a mejores decisiones ambientales, económicas y de salud de la sociedad.

En el anterior estudio, se evaluó el comportamiento poblacional para un caso particular, donde se confirma un aumento de la población que incide por flujos migratorios y variables ambientales. Se pudieron obtener datos públicos sobre el aumento de la población, tales como la tasa de natalidad, la tasa migratoria, la mortalidad y registros sobre Infecciones Respiratorias Agudas.

Los últimos informes públicos sobre la calidad del aire (Colombia), informan que el contaminante con mayor potencial de afectación en la salud humana viene siendo el pm2.5, siguiendo en lista el Co2.

La movilidad de personas y el cambio climático viene siendo ampliamente estudiando por la comunidad científica, ya que se ha convertido en una causa de pobreza en el mundo. Los distintos conflictos en el mundo vienen acrecentando el problema social, económico y ambiental.

La investigación permitió observar comportamientos demográficos en el municipio, de acuerdo a factores migratorios y las consecuencias ambientales que se pueden presentar, por ende, se sugiere: El flujo migratorio debe presentar un mayor control por las entidades encargadas de la municipalidad, donde se mitigue la inseguridad, el trabajo informal y el crecimiento de migrantes; Además, realizar investigaciones que permitan evaluar la incidencia que tienen los contaminantes que se encuentran suspendidos en el aire y que producen enfermedades de tipo respiratorio para disminuir la mortalidad causado por estos factores.

En este trabajo científico no se consideraron factores topográficos, clase de migrantes (raza) y las otras fuentes de contaminación fija y móvil. Estas variables que pueden ser consideradas a futuro para complementar este estudio desarrollado.

REFERENCIAS

- [1] K. Tovar, Z. Gómez, and A. Zamorategui, "Monitoreo de contaminantes atmosféricos (PM2.5, PM10 y CO2) y variables meteorológicas (temperatura, humedad relativa, velocidad y dirección del viento) en la ciudad de Guanajuato," *Jóvenes en la Cienc. Rev. Divulg. Científica*, vol. 5, no. 1, pp. 347–352, 2017, [Online]. Available: <https://bit.ly/3PqjsE1>
- [2] Instituto para la salud geoambiental, "Medición Gas Radón," 2022. <https://www.saludgeoambiental.org/dioxido-carbono-co2/>
- [3] J. Gao, Z. Qiu, W. Cheng, and H. O. Gao, "Children's exposure to BC and PM pollution, and respiratory tract deposits during commuting trips to school," *Ecotoxicol. Environ. Saf.*, vol. 232, p. 11, 2022, DOI: 10.1016/j.ecoenv.2022.113253.
- [4] Y. H. Cheng, Z. S. Liu, and J. W. Yan, "Comparisons of PM10, PM2.5, particle number, and CO2 levels inside metro trains traveling in underground tunnels and on elevated tracks," *Aerosol Air Qual. Res.*, vol. 12, no. 5, pp. 879–891, 2012, DOI: 10.4209/aaqr.2012.05.0127.
- [5] M. E. Blanco Chávez, I. Gómez Carvajal, and S. Mena Bonilla, "CO2 y PM2.5 en la oficina de docentes del Departamento de Operaciones Unitarias," *Arquit.*, vol. 7, pp. 27–35, 2022, DOI: 10.5377/arquit.2022.v7i13.14438.
- [6] C. Zhang, C. Miao, W. Zhang, and X. Chen, "Spatiotemporal patterns of urban sprawl and its relationship with economic development in China during 1990–2010," *Habitat Int.*, vol. 79, no. July, pp. 51–60, 2018, DOI: 10.1016/j.habitatint.2018.07.003.
- [7] B. A. Garro, K. Rodríguez, and R. A. Vázquez, "Classification of DNA microarrays using artificial neural networks and ABC algorithm," *Appl. Soft Comput.*, vol. 38, pp. 548–560, 2015, DOI: 10.1016/j.asoc.2015.10.002.
- [8] F. E. Ahmed, "Artificial neural networks for diagnosis and survival prediction in colon cancer," *Mol. Cancer*, vol. 4, pp. 1–12, 2005, DOI: 10.1186/1476-4598-4-29.
- [9] Departamento Administrativo Nacional de Estadística DANE, "Dane. Estimaciones del cambio demográfico." <https://www.dane.gov.co/index.php/estadisticas-por-%0Atema/demografia-y-poblacion/estimaciones-del-cambio-demografico>
- [10] "Banco de datos. Indicadores del desarrollo mundial," 2021. <https://databank.bancomundial.org/reports.aspx?%0Asource=2&country=COL>
- [11] G. Suárez Guerrero, E. E. Clavijo Gañan, and Universidad Pontificia Bolivariana, "Una propuesta metodológica para una mayor comprensión e interpretación del comportamiento de los fenómenos naturales: A partir de la observación hasta obtener las respuestas del comportamiento del fenómeno. Una propuesta metodológica para una mayor comp," *Mendomatemática*, vol. 20, pp. 1–9, 2009.
- [12] J. Aracil, *Publicaciones de Ingeniería de Sistemas: Dinámica de sistemas*. 1995. [Online]. Available: http://s3.amazonaws.com/academia.edu.documents/30937935/Aracil_I_Gordillo_DS.pdf?AWSAccessKeyId=AKIAJ56TQJRTWSMTNPEA&Expires=1459994585&Signature=RWjAdKFm/D+Aeud+2RUzsgqpCmw=&response-content-disposition=inline;filename=Dinamica_de_sistemas.pdf
- [13] Ministerio De Relaciones Exteriores, "Política Integral Migratoria," 2022. <https://www.cancilleria.gov.co/colombia/migracion/politica>
- [14] N. Ospina Cifuentes and M. C. García Álvarez, "Organizational evaluation using a computational tool for the analysis of critical factors," *LACCEI Int. Multi-Conference Eng. Educ. Technol.*, no. July, pp. 19–21, 2018, DOI: <http://dx.doi.org/10.18687/LACCEI2018.1.1.472>.
- [15] Departamento Administrativo Nacional de Estadística DANE, "Comunicado Oficial," *Comité Nacional para la Vigilancia Epidemiológica (CONAVE)*, 2020. https://www.gob.mx/cms/uploads/attachment/file/573732/Comunicado_Oficial_DOC_sospechoso_ERV_240820.pdf
- [16] A. Beltrán and A. R. Grippa, "Políticas efectivas para reducir la mortalidad infantil en el Perú: ¿Cómo reducir la mortalidad infantil en las zonas más pobres del país?," *Cent. Investig. la Univ. del Pacífico*, pp. 1–45, 2004.
- [17] F. Díez Ballester, J. M. Tenías, and S. Pérez Hoyos, "Efectos de la contaminación atmosférica sobre la salud: Una introducción," *Esp Salud Pública*, vol. 73, pp. 109–121, 1999, DOI: 10.1590/S1135-57271999000200002
- [18] L. Liu, J. Fang, M. Li, M. A. Hossain, and Y. Shao, "The effect of air pollution on consumer decision making: A review," *Clean. Eng. Technol.*, vol. 9, no. December 2020, p.100514, 2022, DOI: 10.1016/j.clet.2022.100514.
- [19] G. He and T. Lin, "Does air pollution impair investment efficiency?," *Econ. Lett.*, vol. 215, p. 110490, 2022, DOI: 10.1016/j.econlet.2022.110490.
- [20] M. Timothy, "Demographic models for projections of social sector demand," *CEPAL-población y Desarrollo*, vol. 66, no. June, pp. 1–63, 2006.
- [21] M. Parves and I. N. Ilina, "Climate change and migration impacts on cities: Lessons from Bangladesh," *Environ. Challenges*, vol. 5, no. May, p. 100242, 2021, DOI: 10.1016/j.envc.2021.100242.

- [22] S. Ayoub, A. Adnan, A. Abdullah, O. Mohammed, W. Sami, and D.C. Klonoff, "Effect of environmental pollutants PM-2.5, carbon monoxide, and ozone on the incidence and mortality of SARS-COV-2 infection in ten wild fire affected counties in California," *Sci. Total Environ.*, vol.757, p.143948, 2021, DOI: 10.1016/j.scitotenv.2020.143948.
- [23] S. Dong, C. Wang, Z. Han, and Q. Wang, "Projecting impacts of temperature and population changes on respiratory disease mortality in Yancheng," *Phys. Chem. Earth*, vol. 117, p. 102867, 2020, DOI: 10.1016/j.pce.2020.102867.
- [24] H. Yin *et al.*, "Population ageing and deaths attributable to ambient PM 2.5 pollution: a global analysis of economic cost," *Lancet Planet. Heal.*, vol. 5, no. 6, pp. e356–e367, 2021, DOI: 10.1016/S2542-5196(21)00131-5.
- [25] J. M. Rodríguez Navarrete, "Los factores ambientales como determinantes del estado de salud de la población en el municipio de Soacha 2006-2016," Pontificia Universidad Javeriana, 2018. [Online]. Available: <https://repository.javeriana.edu.co/handle/10554/46829?show=full>
- [26] J. A. Ezquerro Fernández, *Iniciación a los métodos numéricos*, Universida. 2012.
- [27] Cámara de Comercio de Bogotá, *Plan económico para la competitividad del municipio de Soacha*. 2009.
- [28] N. OspinaCifuentes and M. C. García Álvarez, "Organizational evaluation using a computational tool for the analysis of critical factors," *LACCEI Int. Multi-Conference Eng. Educ. Technol.*, no. July, pp. 19–21, 2018, DOI: 10.18687/LACCEI2018.1.1.472.
- [29] G. Valerio *et al.* "Use of official municipal demographics for the estimation of mortality in cities suffering from heavy environmental pollution: Results of the first study on all the neighborhoods of Taranto from 2011 to 2020". *Environmental Research*, Volume 204, Part B, March 2022, 112007. DOI: 10.1016/j.envres.2021.112007.



Esteban Vallejo Morales es Ingeniero en Informática y Sistemas y recibió la maestría en Ingeniería de la Universidad Pontificia Bolivariana en 2021. Su investigación se centra en el diagnóstico temprano y tratamiento del cáncer colorrectal con simulación por computadora e IA.

ORCID: <https://orcid.org/0000-0002-9916-5766>.



Luis Javier Cruz Riaño, es un investigador del Grupo de Investigación Sobre Nuevos Materiales. Obtuvo su grado en Ingeniería Mecánica en la Universidad Pontificia Bolivariana Medellín, Colombia, en 1985. Luego recibió su Doctorado en Ingeniería en Materiales no Metálicos de la Universidad Politécnica de Madrid, España. Es profesor y director del Grupo de Investigación Sobre Nuevos Materiales de la Universidad Pontificia Bolivariana.

ORCID: <https://orcid.org/0000-0001-7650-5441>.



Julián Andrés Arias Franco obtuvo su grado de Licenciatura en Educación Básica con énfasis en Matemáticas, en 2015, por la Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Es Magister en Ciencias Naturales y Matemática de la Universidad Pontificia Bolivariana, Medellín, Colombia. Sus áreas de interés son:

Pensamiento matemático computacional, Inferencia y regresión de experimentos, métodos numéricos, y optimización matemática.

ORCID: <https://orcid.org/0000-0003-3809-3524>.



Gustavo Suárez Guerrero es un investigador del Grupo de Investigación e Innovación en Energía de la Institución Universitaria Pascual Bravo. Obtuvo su grado en Ingeniería Mecánica por la Universidad Pontificia Bolivariana, Medellín, Colombia, en 1995. Luego recibió su Doctorado en Ingeniería de la Universidad Pontificia Bolivariana,

Medellín, Colombia, en 2011, y además otro título de Doctor en Ciencias Matemáticas por la Universidad Pierre et Marie Curie, París, Francia, en 2012. Sus áreas de interés son los sistemas de biomecánica, modelado y simulación, las matemáticas y los métodos numéricos aplicados.

ORCID: <https://orcid.org/0000-0001-7595-7366>.