



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS
CÓMPUTO CIENTÍFICO

MODELADO DE TRÁFICO VEHICULAR

TESIS

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

PRESENTA:
MIGUEL ANGEL PÉREZ LEÓN

TUTOR:
DR. LUIS MIGUEL DE LA CRUZ SALAS
INSTITUTO DE GEOFÍSICA

COTUTOR:
DR. CARLOS GERSHENSON GARCÍA
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS APLICADAS Y EN SISTEMAS



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Agradezco principalmente a mis padres Sara y Miguel, ya que gracias a ellos y en gran medida por ellos he llegado hasta este punto de mi vida profesional y de mi formación universitaria.

De igual manera agradezco a mis familiares, debido a que sus comentarios han servido de formas inesperadas a esta tesis.

No menos importante, agradezco a mis profesores del posgrado, en especial a mis tutores Dr. Luis Miguel de la Cruz y Dr. Carlos Gershenson por su apoyo y guía.

A mis sinodales, ya que sin sus observaciones esta tesis no habría sido posible.

También a mis compañeros del posgrado, en especial a Hugo Herrera, debido a que su ayuda fue invaluable para llevar a cabo este proyecto.

Finalmente agradezco al personal administrativo del posgrado, así como al CONACyT, ya que gracias a ellos la estancia en el posgrado fue muy grata y gracias a ellos pude dedicarme de lleno a mis estudios.

Pérez León Miguel Angel, Ciudad Universitaria

Resumen

En este trabajo se analizan datos históricos del parque vehicular en la Zona Metropolitana del Valle de México y mediante programación genética se propone un modelo matemático para predecir cómo se comportará el crecimiento del parque vehicular con el paso de los años. La misma técnica es empleada en ciudades con el mismo problema como Río de Janeiro o Beijing y también para ciudades que hayan logrado disminuir el parque vehicular. Posteriormente se realiza una comparativa de los resultados obtenidos por los modelos propuestos para plantear alternativas de solución al creciente problema de tráfico vehicular en la ciudad de México. Finalmente se lleva a cabo la simulación de los datos arrojados por el modelo para la ciudad de México en una muestra representativa.

Contenido

Agradecimientos	1
Resumen	2
Contenido	3
Índice de Figuras	6
Índice de Tablas	7

1. Introducción

1.1. Planteamiento del problema.....	9
1.2. Organización de la Tesis.....	11
1.3. Estado del Arte.....	12
1.3.1. Modelos Macroscópicos.....	13
1.3.2. Modelos Microscópicos.....	13
1.3.2.1. Autómatas Celulares.....	14
1.3.3. Modelos Cinéticos.....	15
1.3.3.1. Modelos tipo Car-following.....	15
1.4. Programación Genética.....	16
1.5. Objetivos	18
1.5.1. Objetivos Generales.....	18
1.5.2. Objetivos Particulares.....	19
1.5.3. Justificación de la Tesis.....	19

2. Marco Teórico

2.1. Programación Genética.....	20
2.1.1. Espacio de Búsqueda.....	21
2.1.2. Primitivas.....	21
2.1.3. Propiedades importantes.....	22
2.1.4. Población Inicial.....	22
2.1.5. Función Objetivo.....	23
2.1.5.1. Aptitud de un individuo.....	23
2.1.6. Operadores Genéticos.....	25
2.1.7. Método de Selección (con elitismo).....	29
2.1.7.1. Torneo.....	29
2.2. Simulación de Tráfico Vehicular.....	31
2.3. Resumen.....	32

3. Modelo Matemático

3.1. Diccionario de Operaciones y Terminales.....	33
3.1.1. Importancia de la Función Exponencial.....	34
3.1.2. Protección de las Operaciones.....	35
3.2. Parámetros del Algoritmo de PG.....	36

Modelado de tráfico vehicular

3.2.1.	Población Inicial.....	36
3.2.2.	Número de Generaciones.....	37
3.2.2.1.	Cruce.....	37
3.2.2.2.	Método de Selección.....	37
3.2.3.	Función Objetivo.....	37
3.3.	Datos Históricos del Parque Vehicular.....	39
3.3.1.	ZMVM.....	39
3.3.2.	Otras ciudades.....	40
3.4.	Modelo Matemático.....	41
3.4.1.	Validación del modelo.....	44
3.5.	Detalles de la implementación.....	48
3.5.1.	Hardware.....	49
3.5.2.	Arquitectura.....	49
3.6.	Resumen.....	46

4. Resultados

4.1.	Población.....	52
4.2.	Parque Vehicular.....	54
4.2.1.	Beijing.....	55
4.2.2.	Los Angeles.....	55
4.2.3.	Londres.....	56
4.2.4.	Nueva York y Río de Janeiro.....	56
4.3.	Proyección para cada ciudad.....	57
4.3.1.	Beijing.....	57
4.3.2.	Nueva York.....	58
4.3.3.	Londres.....	59
4.3.4.	Los Angeles.....	60
4.3.5.	Río de Janeiro.....	60
4.3.6.	Delegación Cuauhtémoc.....	62
4.4.	Simulación.....	63
4.4.1.	Conceptos importantes.....	63
4.4.1.1.	Velocidad.....	63
4.4.1.2.	Tiempo de viaje.....	64
4.4.1.3.	Velocidad Promedio de Viaje.....	64
4.4.1.4.	Densidad.....	64
4.4.1.5.	Flujo.....	65
4.5.	Simulador.....	65
4.5.1.	Construcción del mapa.....	66
4.6.	Simulaciones.....	66
4.6.1.	Contaminante (CO ₂ , NO _x).....	66
4.6.2.	Consideraciones de la Simulación.....	68
4.6.3.	Resultados de la Simulación.....	69
4.7.	Resumen.....	70

5. Conclusiones

5.1.	Ventajas de emplear PG en el contexto del parque vehicular.....	71
5.2.	Colapso Vial.....	71
5.3.	Contaminación Ambiental.....	72
5.4.	Inseguridad Vial.....	72
5.5.	Propuestas.....	73
5.5.1.	Uso de la Bicicleta.....	73
5.5.2.	Compartir Vehículos.....	73
5.5.3.	Gravámenes a la compra de vehículos.....	74
5.5.4.	Distribución de horarios laborales y escolares por zonas.....	74
5.6.	Trabajo a Futuro.....	74
5.6.1.	Enfoque Multifactorial.....	74
5.6.2.	Programación en paralelo.....	75
5.6.3.	Estimación y análisis del índice de asaltos a automovilistas en vialidades conflictivas.....	75
	Referencias.....	76

Índice de Figuras

Figura 1: Arbol sintáctico básico.....	21
Figura 2: Cruce entre individuos.....	26
Figura 3: Resultado del cruce de individuos.....	26
Figura 4: Operador de mutación.....	27
Figura 5: Resultado de la mutación.....	27
Figura 6: Proceso general de la Programación Genética.....	28
Figura 7: Función Objetivo ZMVM.....	38
Figura 8: Grafica de los 50 modelos generados ZMVM.....	41
Figura 9: Mejor modelo ZMVM.....	42
Figura 10: Modelos generados con datos hasta el año 2011.....	45
Figura 11: Mejor modelo generado con datos hasta el año 2011.....	45
Figura 12: Diagrama de clases.....	50
Figura 13: Comparativa de población por años, de las ciudades analizadas.....	53
Figura 14: Comparativa del parque vehicular por años, ciudades analizadas.....	54
Figura 15: Modelo para la ciudad de Beijing.....	57
Figura 16: Grafica del modelo promedio de la ciudad de Beijing.....	58
Figura 17: Mejor modelo para la ciudad de Nueva York.....	58
Figura 18: Mejor modelo para la ciudad de Londres.....	59
Figura 19: Mejor modelo para la ciudad de Los Angeles.....	60
Figura 20: Mejor modelo para la ciudad de Río de Janeiro.....	61
Figura 21: Modelo promedio Río de Janeiro.....	61
Figura 22: Mejor modelo delegación Cuauhtémoc.....	62
Figura 23: Mapa de la delegación Cuauhtémoc.....	69
Figura 24: Proyección de las emisiones de contaminantes por años.....	70

Índice de Tablas

Tabla 1: Ejemplo de dominio y codominio.....	23
Tabla 2: Función objetivo de ejemplo.....	24
Tabla 3: Resultado de la evaluación del individuo x.....	24
Tabla 4: Población de individuos relacionados con su aptitud.....	29
Tabla 5: Organización del torneo.....	30
Tabla 6: Resultado del torneo.....	31
Tabla 7: Municipios que conforman la ZMVM.....	40
Tabla 8: Datos históricos del parque vehicular de la ZMVM.....	44
Tabla 9: Aptitudes de individuos generados para la validación.....	46
Tabla 10: Resultado del modelo de la ZMVM con datos hasta el 2011.....	48
Tabla 11: Población por años (2000, 2010, 2014) de ciudades analizadas.....	53
Tabla 12: Población de las ciudades analizadas.....	54
Tabla 13: Cantidad de vehículos por año, empleados en la simulación.....	68

Capítulo 1

INTRODUCCIÓN

El tráfico vehicular se ha convertido en un área de estudio y de interés para las ciencias, ya que en gran medida el desempeño de una ciudad puede verse afectado por dicho fenómeno. Las consecuencias tanto negativas como positivas del tráfico vehicular impactan de múltiples maneras en los habitantes de una ciudad, y se pueden ver reflejadas en temas que van desde la contaminación ambiental y la salud de los habitantes de una ciudad, hasta la cantidad de dinero que se invierte en el transporte cuando el tráfico es excesivo, así como el nivel de estrés que el tráfico vehicular puede provocar.

Es por esto que el estudio del tráfico vehicular puede proveer de herramientas importantes para agilizar y en el mejor de los casos reducir en gran medida el tránsito de vehículos.

Para llevar a cabo el estudio del tráfico vehicular es necesario generar modelos conceptuales, que se transforman en modelos matemáticos y que permitan simular con datos reales los factores que intervienen en este fenómeno. De tal manera que mediante este análisis se puedan tomar medidas que ayuden a minimizar el tráfico vehicular.

Desde hace un par de décadas se han generado modelos que permiten analizar el fenómeno del tráfico vehicular, pero por otro lado es importante dar una predicción adecuada de en qué momento la densidad vehicular será tal que el desplazamiento en un vehículo será prácticamente imposible.

Actualmente, la ciudad de México se encuentra en el primer lugar en el ranking de las ciudades con el peor tráfico en el mundo, la Ciudad de México se encuentra muy cerca de un colapso vial si no se toman las medidas pertinentes para disminuir el tránsito de vehículos, sobre todo en horas pico.

Tomando como base datos reales y recientes del parque vehicular de la Zona Metropolitana del Valle de México en este trabajo se construye un modelo matemático, empleando programación genética y regresión simbólica, para estudiar el comportamiento del crecimiento en el parque vehicular con el paso de los años. Este modelo tiene la intención de dar una predicción del momento en que será imposible el tránsito vehicular si no se toman las medidas adecuadas. De igual manera cómo se construyó el modelo matemático que predice el crecimiento del parque vehicular en la Zona Metropolitana del Valle de México, se propone otro modelo matemático para la predicción del crecimiento del parque vehicular de una zona representativa de la Ciudad de México, en este caso la delegación Cuauhtémoc.

Modelado de tráfico vehicular

El mismo procedimiento empleado para la creación de los modelos matemáticos de la Zona Metropolitana del Valle de México y de la delegación Cuauhtémoc, será llevado a cabo con las ciudades de, Beijing, Los Angeles, Londres, Nueva York y Río de Janeiro, con la intención de realizar una comparativa entre estas ciudades y la Ciudad de México, y también para identificar qué medidas tomaron las ciudades que lograron disminuir el parque vehicular.

Finalmente, el modelo de la delegación Cuauhtémoc cumple con el propósito de alimentar con los datos obtenidos de este modelo, una simulación generada con SUMO, un simulador de movilidad urbana. Esta simulación tiene el objetivo de hacer tangible el impacto que tendrá el crecimiento vehicular en la delegación Cuauhtémoc.

El tema de la movilidad es de los que más afecta la calidad de vida de la población, solo después de la seguridad, aunque en fechas recientes se ha visto que el incremento en el tráfico vehicular es un factor que interviene en el incremento de asaltos a automovilistas, por lo que es importante proponer alternativas para mejorarla. Se presentan al final de este trabajo, algunas de las medidas que otras ciudades han tomado para disminuir el parque vehicular y así mejorar el tránsito.

1.1 Planteamiento del problema

Cada vez es más notable el creciente problema de tráfico vehicular, al hacer memoria se llega a la conclusión de que el tiempo que toma trasladarse del punto *a* al punto *b* en la ciudad de México cada vez es mayor, esto se debe a múltiples factores como lo son:

- El clima.
- Las manifestaciones ciudadanas.
- El mal estado de las calles y avenidas, etc.

Desde mediados del siglo pasado el estudio del tráfico vehicular[1] ha sido analizado para tratar de encontrar una solución al congestionamiento de las vías. Sin embargo, es fácil notar que el problema principal por el cual se genera tráfico en la ciudad de México es la cantidad de vehículos que en ella se desplazan.

En este sentido tenemos dos problemas a analizar, la simulación del tráfico vehicular y la modelación en el crecimiento del parque vehicular. Actualmente las técnicas de planeación de tráfico vehicular desempeñan un rol muy importante para tratar de optimizar los recursos de los que se disponen.

Por otro lado, tratar de predecir un comportamiento, cualquiera que este sea, por ejemplo; la trayectoria de un huracán, el comportamiento de la bolsa de valores, o simplemente la tasa de crecimiento poblacional siempre ha sido tema de interés para el área de las ciencias. Desde hace mucho tiempo se han hecho esfuerzos por

Modelado de tráfico vehicular

tratar de encontrar métodos eficientes que proporcionen resultados confiables y así poder predecir un determinado comportamiento. Respecto a la modelación del crecimiento del parque vehicular, no se tiene un problema trivial ya que es una labor compleja tratar de identificar relaciones analíticas en comportamientos sociales como lo es crecimiento del parque vehicular.

En Nueva Zelanda en el año 2002[2] se llevó a cabo una revisión de los métodos para la predicción del crecimiento en el parque vehicular con los se contaba en ese entonces. Pero la programación genética aún no tenía la relevancia con la que cuenta actualmente, así que en el documento antes mencionado no se hace uso de esta técnica tan poderosa para predecir el crecimiento del parque vehicular en Nueva Zelanda.

Se han llevado a cabo trabajos empleando programación genética respecto a la predicción en el comportamiento de fenómenos sociales como los que se mencionan en [3]. En este artículo se emplea programación genética y regresión simbólica para construir el modelo que trate de predecir el crecimiento en la confiabilidad del desarrollo de software.

En fechas más recientes (2009) se han hecho esfuerzos por tratar de modelar de manera más eficiente cualquier tipo de comportamiento a partir de datos experimentales, en [4] se muestra la manera en la cual se ha intentado llevar a cabo esta tarea.

Sin embargo en ninguno de los casos antes mencionados se ha empleado la programación genética para poder obtener un modelo confiable que prediga el comportamiento del tráfico vehicular.

En fechas recientes es notable que el tráfico en cualquier punto de la ciudad de México ha ido en aumento, ya no es posible desplazarse a nuestros destinos en el tiempo que habitualmente se tenía contemplado. Esto se debe principalmente al crecimiento del parque vehicular, mismo que se incrementa año tras año. Es fácil notar que llegará un momento en el cual la cantidad de vehículos activos y desplazándose en la ciudad de México será tal que ya no será viable este método de transporte. Con estos datos será posible generar una comparación con ciudades que han tenido el mismo problema y más aún, una simulación que muestre de manera tangible el problema hacia el cual se dirigen los habitantes de esta ciudad. Tener una estimación de cuándo sucederá esto es importante para las autoridades de la ciudad, ya sea para que tomen medidas preventivas o medidas correctivas.

Hasta el momento de redactar este trabajo no existen registros previos de haber empleado programación genética para predecir el crecimiento del parque vehicular. Por otro lado para la parte de simulación existen muchos tipos de modelos empleados, desde modelos que se enfocan en la interacción de los vehículos, hasta

los que consideran el tráfico vehicular como un fluido. En las siguientes secciones se revisará más a detalle el funcionamiento de algunos de estos modelos.

1.2 Organización de la Tesis

- Capítulo 1: Planteamiento del problema del crecimiento del parque vehicular en la ciudad de México, organización de la tesis, así como los objetivos generales, particulares y la justificación de este trabajo. Por otra parte, se muestra el estado del arte de los modelos empleados en la simulación de tráfico vehicular, así como estado del arte de la programación genética.
- Capítulo 2: Marco teórico de la programación genética, en este capítulo se muestran los conceptos teóricos. Se describen los conceptos de programación genética necesarios para el desarrollo de esta tesis.
- Capítulo 3: Creación del modelo matemático, se muestra el proceso mediante el cual fue posible obtener el modelo planteado en esta tesis. También se mencionan las peculiaridades encontradas en el proceso de creación del modelo matemático, así como los experimentos llevados a cabo en este proceso.
- Capítulo 4: Resultados del modelo matemático, así como de la comparativa con las distintas ciudades consideradas en este trabajo. De igual manera se generó la simulación necesaria para hacer tangible el impacto que tendrá el crecimiento del parque vehicular.
- Capítulo 5: Conclusión de los datos obtenidos por el modelo para la ciudad de México y trabajo a futuro.

Al principio de cada uno de los capítulos se incluyen una introducción, que describe de manera breve el contenido de éstos. Para facilitar la comprensión de la información presentada, se incluye al final de cada capítulo un resumen, con lo más relevante del mismo.

Además, se incluyen las referencias que se ha consultado, esto con el objeto de que cualquier persona que quiera consultar de manera detallada alguno de los puntos tratados aquí, tenga las fuentes necesarias.

De manera complementaria, se incluyen dos anexos: manual de usuario y del programador. El manual de usuario muestra de manera clara y concisa cómo emplear el software desarrollado en esta tesis para poder hacer uso del sistema desde la perspectiva del usuario final. En este manual se indica cómo utilizar el software desarrollado para poder crear modelos matemáticos basados en datos históricos, y cómo interpretar y utilizar los datos arrojados por el sistema. En el manual del programador se detalla el desarrollo del sistema, se describen las herramientas empleadas, las versiones de las plataformas empleadas, así como la documentación de cada una de las clases creadas para tal motivo. El manual del programador sirva

para que cualquiera, que así lo desee, pueda actualizar o usar este software para los fines que así le convengan.

1.3 Estado del Arte

Actualmente, la modelación y simulación por computadora tienen una gran importancia en la optimización de problemas, en particular en el problema del flujo de tráfico vehicular. Investigaciones teóricas y experimentales han sido desarrolladas, analizadas y reportadas.

La teoría del flujo de tráfico vehicular puede separarse ampliamente en tres tipos de modelos: Los Modelos Macroscópicos, los Modelos Microscópicos y los Modelos Cinéticos.

De igual manera los modelos de flujo de tráfico se pueden clasificar de varias maneras:

- a) Por su nivel de detalle.
- b) Por su nivel de operación.
- c) Por la representación de los procesos.

Cuando se habla de modelos por nivel de detalle, las aproximaciones son las ya antes mencionadas.

1.3.1 Modelos Macroscópicos

Estos modelos tienen su base en ecuaciones de la dinámica de fluidos, debido a esto los modelos macroscópicos no son idóneos para describir interacción entre individuos de un pelotón de autos, es decir las interacciones individuales que pueden llevarse a cabo entre los distintos vehículos. En este tipo de modelos el enfoque principal es ver a los vehículos como un flujo global y los vehículos son el equivalente a partículas que se desplazan a través de un ducto. Cuando se genera una simulación empleando este tipo de modelos, el tiempo de simulación y los requerimientos de memoria, se basan en la discretización del espacio y tiempo, el número de vehículos no es un factor que sea destacado para este tipo de modelos pero este valor si se encuentra relacionado con la densidad y el flujo. Las variables que si son importantes en este tipo de modelos son la densidad espacial promedio y la velocidad promedio.

Para poder medir la confiabilidad de los resultados que arrojan las simulaciones basadas en este tipo de modelos es necesario evaluar la exactitud de las ecuaciones macroscópicas que fueron empleadas en la generación del modelo

matemático y sobretodo el o los métodos numéricos empleados para la solución de las mismas.

La calidad y confiabilidad de los resultados de la simulación dependen principalmente de la exactitud de las ecuaciones macroscópicas aplicadas y de la elección de un método de integración numérica conveniente.

Uno de los modelos macroscópicos que ha sido ampliamente estudiado es el modelo básico de Lighthill y Whitham[1] publicado en 1955. En este modelo se introdujo por primera vez la ecuación de continuidad como base para crear el modelo matemático, asumiendo que el flujo únicamente depende de la densidad, lo cual es erróneo en el contexto del tráfico vehicular.

Otra ecuación que fue empleada en este tipo de modelos es la ecuación de aceleración Payne[5] y Kühne[6]. Años después Payne realizó cambios al modelo, reemplazando la suposición de la adaptación instantánea del modelo de Lighthill-Whitman por una ecuación de inercia muy similar a la ecuación de Navier-Stokes[7]

Otros modelos macroscópicos que han tenido aportes importantes son los de Euler y el de Helbing. Posteriormente Helbing identificó que la varianza de velocidad era un factor importante en los modelos macroscópicos, es por esto que al tratar de modelar el tráfico vehicular un incremento en esta varianza es un indicador importante de que se puede producir un embotellamiento. Así que Helbing[8] introdujo su modelo en el cual esta varianza es de vital importancia para el desempeño óptimo de la simulación.

1.3.2 Modelos Microscópicos

A diferencia de los modelos macroscópicos, los modelos microscópicos analizan el flujo dándole un tratamiento individual, es decir que los vehículos son analizados de manera individual. Por cada vehículo se tiene una ecuación que describe su comportamiento, en particular el movimiento que este tiene en el modelo. Dado que en los modelos microscópicos el comportamiento individual desempeña un papel importante, los datos probabilísticos son de mayor importancia que en los modelos macroscópicos. En este tipo de modelos la cantidad de vehículos es un factor importante en el modelo matemático, es por esto que al interactuar los vehículos emerge una gran complejidad que tiene efectos importantes en el desempeño del modelo.

Matemáticamente, parte de esta teoría es similar al tratamiento de los movimientos atómicos en cristales y da resultados acerca de la estabilidad de cadenas de vehículos (pelotones). Dentro del marco de los modelos microscópicos también encontramos los modelos basados en la teoría de colas.

1.3.2.1 Autómatas Celulares

También considerado dentro de los modelos microscópicos se tiene el estudio del flujo vehicular tratado mediante modelos generados con Autómatas Celulares (AC). El análisis del tráfico vehicular mediante AC simplifica en gran medida la complejidad que emerge de este tipo de modelos, ya que el comportamiento de cada vehículo es tratado mediante reglas de evolución, estas reglas son en general sencillas, pero esto no decreta en sentido alguno la confiabilidad de los modelos generados con AC. Mediante AC la discretización del dominio se lleva sobre una lattice¹ de nodos y en cada nodo se tiene un vehículo con una velocidad discreta, cada vehículo se desplaza en función de las reglas de transición definidas en el modelo. El modelo insignia para el flujo de tráfico vehicular empleando AC, es el modelo de Kai Nagel y Michael Schreckenberg (modelo NaSch[9]), este es un modelo probabilístico de un solo carril con velocidad máxima de 5. Sin embargo antes del modelo NaSch ya se tenía una primera noción del uso de AC para el modelado de tráfico vehicular, el modelo de Biham[10] es un modelo básico con velocidad máxima igual a 1.

La simplicidad del modelo NaSch es tal que permite describirlo en pocas palabras. Cada celda o nodo de la lattice está ocupada por un vehículo o se encuentra vacía en cada instante de tiempo t . Cada vez que el sistema se actualiza pasa del tiempo t al tiempo $t+1$ y la forma que cada vehículo se comporta está dada por cuatro reglas que rigen el modelo. Las reglas toman en cuenta la velocidad que lleva el vehículo i -ésimo y la distancia que existe entre este vehículo y el vehículo siguiente.

Supongamos que se tiene una autopista de un solo carril y esta se encuentra dividida en celdas iguales de tamaño tal que en cada una de ellas cabe a lo más un vehículo. Se define el espacio entre dos vehículos consecutivos como esp , como el número de celdas vacías que hay entre los dos. Si v es la velocidad de un vehículo cualquiera al inicio de un intervalo de tiempo, entonces su velocidad al inicio del siguiente intervalo de tiempo se calcula de acuerdo a las siguientes reglas:

1. Si $v < v_{max}$ y $v < esp$, entonces la velocidad del vehículo al inicio del siguiente intervalo de tiempo será $v + 1$.
2. Si $v < esp$, entonces la velocidad del vehículo al inicio del siguiente intervalo de tiempo será esp .
3. Se genera un número aleatorio r , entre 0 y 1. Si $r < \rho$, donde ρ es algún número constante también entre 0 y 1, entonces la velocidad del vehículo al inicio del siguiente intervalo de tiempo será $v - 1$.
4. El vehículo avanza v celdas hacia adelante.

¹ En este contexto es sinónimo de retícula, que se emplea en la discretización del modelado del tráfico vehicular.

Como se mencionó anteriormente en este tipo de modelos el parámetro probabilístico está dado por la probabilidad que se tiene de que un vehículo disminuya su velocidad por algún contratiempo, dicho parámetro es conocido como parámetro de frenado.

Este y otros modelos han sido probados con datos de tráfico real y los resultados han sido muy próximos a los resultados esperados. Posteriormente Knospe y algunos otros autores ampliaron el modelo NaSch agregando factores como:

- Carriles extra.
- Carriles independientes.
- Carriles adyacentes.

Para poder simular estos nuevos factores fue necesario extender el modelo NaSch incluyendo un conjunto de reglas que permitan simular el cambio de carril. Fue a partir de este momento que se inicia el estudio de flujo de tráfico multi-carril con AC.

1.3.3 Modelos Cinéticos

Entre los modelos macroscópicos y los modelos microscópicos, se tienen los modelos cinéticos, esto quiere decir que los modelos cinéticos pueden conducir a los modelos microscópicos y de igual manera los modelos macroscópicos pueden derivar en los modelos cinéticos. Algunos de los modelos más importante en el campo de los modelos cinéticos es el modelo de Prigogine y el modelo de Paveri-Fontana, en ambos casos la pieza fundamental es una función de distribución $f(x,v,t)$ que describe el número de vehículos en cierto lugar y velocidad en el tiempo t .

En el modelo de Paveri-Fontana se introdujo una función de distribución que generaliza la distribución de Prigogine. Esta función $g(x,v,t,w)$ describe el número de vehículos en cierto lugar con velocidad v en el tiempo t y con una velocidad deseada w .

1.3.3.1 Modelos de tipo Car-following

Los modelos de tipo vehículo-siguiente también son considerados como modelos microscópicos. A principios de la década de los 50's se comenzó a considerar esta teoría para el estudio del tráfico vehicular. La idea básica de este tipo de modelos se centra en la relación que existe entre un vehículo y su antecesor inmediato. Para esta teoría, se toman en cuenta situaciones de un solo carril, en las

que el conductor de un determinado vehículo reacciona en función del movimiento del vehículo que se encuentra delante de este.

Una de las primeras teorías de flujo vehicular se basó en leyes de estímulo-respuesta (conductismo) donde la respuesta al cambio en la velocidad se mide por la aceleración y el estímulo puede tener diversos orígenes. El común de este tipo de modelos supone una relación lineal entre la velocidad y el espaciamiento o bien entre la aceleración y la diferencia de velocidades entre vehículos consecutivos. Esta teoría se desarrolló principalmente en la compañía norteamericana General Motors que se conoció con el nombre de "car-following"[11]. Como es costumbre en esta teoría las posiciones de los vehículos se denotan de derecha a izquierda comenzando con el líder, así $x_n(t)$, $x_{n-1}(t)$, $x_{n-2}(t)$ y $v_n(t)$, $v_{n-1}(t)$, $v_{n-2}(t)$ son las posiciones y velocidades respectivamente del n-ésimo vehículo, $n = 0$ es el líder y $n = 1, 2, 3, \dots$ son los seguidores. En este caso nos estaremos refiriendo a una caravana (pelotón) de autos. Cuando se hace énfasis en la interacción local, se acostumbra la notación x_f para el seguidor x_l para el líder. Los modelos de tipo car-following de una sola vía suponen que existe una alta correlación entre los vehículos en un rango de separación intervehicular de 100 a 125 metros y dan una forma explícita para ella. Asumen que el conductor es un agente activo que responde a estímulos sensoriales mediante aceleración-desaceleración del vehículo, dependiendo de sus propias habilidades psicomotoras y condiciones físicas del camino y del vehículo.

1.4 Programación Genética

La Programación Genética (PG) es una técnica que ha sido estudiada extensivamente y ha demostrado ser eficiente en múltiples escenarios. Tiene como objetivo construir, a partir de un conjunto de algoritmos, el modelo matemático que se ajuste de la mejor manera al problema que se quiere modelar. Además mejora de manera automática al ir experimentando con los datos con los cuales se está experimentando.

Se puede ubicar la primera referencia hacia la PG basada en árboles jerárquicos y operadores definidos formalmente en 1985. Esta definición fue dada por Michael L. Cramer[12] y es en este punto donde podemos decir que la PG toma la forma que conocemos actualmente.

Más adelante John R. Koza[13], principal exponente de la PG profundizó en el tema y aplicó su teoría en diversos problemas de búsqueda y optimización. Posteriormente Gianna Giavelli estudiante doctoral de Koza fue pionero empleando PG para modelar secuencias de ADN.

Algunos autores consideran que la PG es parte de lo que se conoce como Machine Learning[14]. La PG es un sistema de aprendizaje de máquina que

evoluciona por sí misma empleando estructuras de árbol para encontrar el modelo matemático adecuado para el problema planteado, en este caso el parque vehicular.

En principio se puede decir que la PG pertenece al conjunto de los algoritmos genéticos (AG)[15]. La PG es una técnica evolutiva en la que cada individuo representa un posible modelo, es decir una posible solución. Además cada individuo está conformado por funciones de valores y dicho individuo está codificado empleando un árbol jerárquico que representa el modelo a evaluar. Es debido a esta característica que la PG proporciona una manera de representar las posibles soluciones altamente dinámica que ayuda a construir el posible modelo matemático.

El espacio de búsqueda es el espacio de todos los posibles árboles jerárquicos que representan el modelo matemático que se busca, estos árboles están conformados por argumentos definidos para el problema y las funciones que se hayan definido previamente. Al emplear PG se garantiza que el modelo encontrado está bien definido y por lo tanto los resultados arrojados por dicho modelo son altamente confiables. Por otro lado, es necesario definir de manera clara los elementos que intervienen en el proceso de PG, es decir, que elementos pueden formar parte de un diseño adecuado para generar un modelo empleando PG. En el caso del parque vehicular existen múltiples factores además del tiempo, que determinaran como crecerá el parque vehicular, factores tales como, precio de la gasolina, políticas publicas etc. En este trabajo se consideró únicamente el factor del tiempo para determinar el modelo matemático que describe el crecimiento del parque vehicular, pero para un análisis multifactorial sería necesario considerar más elementos que formen parte de los factores del problema a analizar.

Para poder emplear PG es necesario definir un límite, por ejemplo un número máximo de nodos que componen el árbol o una profundidad finita para el árbol jerárquico. Una vez que ya se tiene bien definido el problema, las funciones y los datos de entrada es necesario definir la función objetivo (función fitness) que en la mayoría de los casos es el error producido por la solución que se evalúa en ese momento, es decir la diferencia entre los valores arrojados por el modelo evaluado y los datos reales. Generalmente cada individuo es evaluado tomando en consideración un conjunto de diferentes casos que permiten determinar que tan 'buena' solución es. De igual manera que en AG existen muchas formas de asignar el valor fitness para cada individuo, que pueden ir desde ranking, pasando por estrategias multi-objetivos y hasta funciones de agregación.

El resto del funcionamiento de la PG es muy similar a los AG, es decir que se tiene una etapa de cruzamiento de los individuos, mutaciones y evaluaciones para determinar qué individuos logran transferir sus genes a la siguiente generación.

Para la gran mayoría de las ciencias, la predicción del comportamiento de un determinado fenómeno siempre ha sido una labor de vital importancia, ya sea en cuestiones climáticas, bursátiles, urbanas o de la índole que sea. El uso de PG para

predecir este tipo de comportamientos es relativamente nuevo. En general, se trata de un campo de reciente aparición. Se puede decir que las bases sobre las cuales se sustenta la PG y la regresión simbólica se tiene a finales de los años 80 ya que fue en el año de 1987 cuando Diday[16][17] define las bases para poder emplear los principios de la PG empleando regresión simbólica.

Respecto a la modelación en el crecimiento del parque vehicular se cuentan con algunos trabajos realizados, por ejemplo el llevado a cabo por G.F. Koorey, y T.J. Mitchell en 2002 en Nueva Zelanda [2]. En este trabajo se muestran los métodos empleados hasta el momento para la proyección del crecimiento en el parque vehicular, sin embargo no se hace uso de PG, lo cual es una gran desventaja ya que posteriormente se ha demostrado la gran confiabilidad de la PG para la predicción de comportamientos.

Más recientemente Wasif Afzal, Richard Torkar en su artículo *Suitability of genetic programming for software reliability growth modeling*[3] discute lo adecuado que puede ser aplicar PG para determinar el crecimiento en la confiabilidad del desarrollo de software. Esto muestra la gran versatilidad que proporciona la PG para la creación de modelos que determinen el comportamiento de un fenómeno, cualquiera que este sea, en este caso algo tan complejo como lo es la calidad del software.

Por último en el año (2009) se publicó el artículo *Distilling Free-Form Natural Laws from Experimental Data*[4] y aquí se discute de manera más formal el problema de modelar de manera más exacta cualquier fenómeno que pueda proporcionar datos experimentales, como los son las mediciones históricas de un comportamiento.

De tal manera que el uso de PG para la predicción en el crecimiento del parque vehicular no se ha llevado a cabo de manera formal, así que puedo concluir que este es el estado del arte de la PG con aplicación en el crecimiento del parque vehicular.

1.5 Objetivos

1.5.1 Objetivo General

- Empleando como base datos históricos del parque vehicular generar un modelo matemático confiable, que ayude a determinar cómo se comportará a futuro el crecimiento del parque vehicular en la Zona Metropolitana del Valle de México y en ciudades que comparten el problema del tráfico vehicular. Realizar una comparativa con estas ciudades y con algunas ciudades que hayan logrado disminuir de manera satisfactoria el parque vehicular.

1.5.2 Objetivos Particulares

- Analizar las distintas técnicas empleadas para la creación de modelos de tráfico vehicular.
- Realizar un estudio minucioso de la teoría de la Programación Genética y regresión simbólica.
- Usando el modelo propuesto, mostrar en qué momento la cantidad de vehículos activos en la ciudad de México sobrepasarán el límite de vehículos que pueden circular por las vialidades de la ZMVM.
- Identificar las estrategias empleadas por las ciudades que lograron disminuir el parque vehicular.
- Proponer estrategias similares por las empleadas por ciudades que lograron superar este problema, para la ciudad de México.
- Empleando datos históricos del parque vehicular y datos arrojados por el modelo propuesto en esta tesis, crear una simulación de una zona representativa de la ciudad de México que muestre de manera tangible el impacto que tendrá el problema del crecimiento vehicular de manera desmedida.

1.5.3 Justificación de la Tesis

Es de vital importancia el conocer cómo es que se comportara el crecimiento del parque vehicular, ya que con esta información las autoridades pueden tomar cartas en el asunto para evitar un posible colapso vial. Existen análisis previos sobre el crecimiento del parque vehicular en la Zona Metropolitana del Valle de México, sin embargo no se ha realizado alguno empleando Programación Genética, ni regresión simbólica. Por otra parte, la Programación Genética y la regresión simbólica son técnicas que no tienen tantas restricciones como las empleadas hoy en día para la predicción de fenómenos tales como el del crecimiento del parque vehicular. Es por esto que el trabajo presentado en esta tesis cuenta con una justificación válida.

En este trabajo se utilizará Programación Genética y regresión simbólica para poder construir un modelo confiable que ayude a predecir de manera confiable el crecimiento del parque vehicular en la ciudad de México. De manera similar se construirán modelos matemáticos para otras ciudades y para una zona representativa de la Ciudad de México. Los datos de salida del modelo de la zona representativa, servirán como datos de entrada para la simulación realizada en esta tesis.

Capítulo 2

MARCO TEÓRICO

Es labor de este capítulo mostrar conceptos fundamentales sobre Programación Genética y Simulación de Tráfico Vehicular, mismos que son necesarios para la creación del modelo matemático que se propone en esta tesis.

2.1 Programación Genética

La Programación Genética es la técnica mediante la cual se construyen modelos matemáticos, mismos que para su mayor facilidad de manejo y comprensión son representados en forma de árbol jerárquico². Cada árbol representa un modelo posible y pueden ser evaluados de manera recursiva para determinar qué tan “buena” solución representa. Un posible modelo (o solución) representado en forma de árbol, está conformado por nodos, donde cada nodo consta de una función como operador y todo nodo terminal contiene un operando que puede ser una constante o un valor aleatorio. De esta forma las expresiones matemáticas resultantes pueden evolucionar fácilmente para ser evaluadas nuevamente y finalmente obtener el modelo que mejor se ajuste al problema planteado.

A este procedimiento comúnmente se le conoce como regresión simbólica, es decir, una aplicación de la PG. La regresión simbólica está enfocada en conseguir el mismo objetivo que la regresión lineal, sin embargo, cuenta con un espacio de búsqueda mucho mayor y sus limitaciones son menos.

En el contexto de esta tesis los datos de entrada que alimentarán el algoritmo de PG será la medición anual del parque vehicular en la ZMVM³, para así obtener el modelo que logre predecir el comportamiento del mismo al paso del tiempo.

Es decir que cada posible modelo o individuo de la población tendrá la siguiente estructura que se muestra en la figura 1.

² Estructura de Datos conformada por nodos y aristas, en teoría de graficas se dice que un árbol jerárquico es un grafo dirigido, conexo y acíclico.

³ Abreviatura para Zona Metropolitana del Valle de México.

Modelado de tráfico vehicular

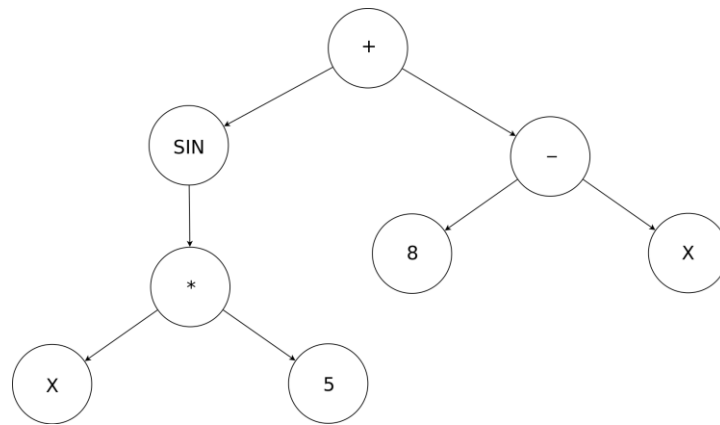


Figura 1: Árbol sintáctico básico, que representa un individuo en el contexto de PG.

Lo que en este ejemplo sencillo se interpreta como:

$$F(x) = \sin(X * 5) + (8 - X)$$

2.1.1 Espacio de Búsqueda

El espacio de búsqueda definido para la PG es el espacio de todos los posibles árboles jerárquicos compuestos de las funciones y argumentos definidos para el problema que se requiere atacar. Cada una de estas estructuras es un posible modelo o una posible solución al problema, en la mayoría de los casos el espacio de búsqueda está definido por factores como la población y las generaciones, estos factores son importantes para la PG y normalmente son factores que deben ser calibrados para obtener óptimos resultados.

2.1.2 Primitivas

Terminales: son todo aquel conjunto que consta de variables, constantes y funciones que no necesitan recibir parámetros. En el contexto de este trabajo los terminales son:

- Datos de entrada: es decir la variable independiente que representa el año en la medición del parque vehicular.
- Valores aleatorios: son valores que ayudan a enriquecer el espacio de búsqueda y se ha demostrado que con estos valores los resultados se ajustan de mejor manera.

Funciones: es el conjunto de funciones que se emplearán para encontrar el modelo matemático que mejor se ajuste al problema del crecimiento del parque vehicular. En el contexto de este trabajo tendremos las operaciones:

- Aritméticas: +, -, *, % (división).

- Trigonómicas⁴: seno, coseno, tangente (versiones hiperbólicas).
- Funciones: exponencial y sigmoide.

Este conjunto de terminales y operadores mostraron ser los ideales para llevar a cabo la creación del modelo matemático que se describe en el siguiente capítulo, ya que con este conjunto de terminales y operadores se obtenían los mejores modelos, es decir modelos matemáticos cuya aptitud era superior a modelos generados empleado un conjunto diferente de terminales y operadores. Una de las operaciones que probó ser vital para crear modelos matemáticos cuyas aptitudes fuera aceptables, fue la función exponencial, pero se hablara de esto más a fondo en el siguiente capítulo.

2.1.3 Propiedades importantes

- **Cerradura:** con esta propiedad se asegura que todo elemento del conjunto de funciones debe recibir datos de un solo tipo y debe devolver valores que puedan ser argumentos que reciba nuevamente otra función. Es decir que cualquier elemento que se encuentre en el conjunto de funciones debe estar bien definida para todo tipo de combinación de argumentos que se encuentre en el espacio de búsqueda.
- **Suficiencia:** esta propiedad garantiza que el modelo obtenido mediante PG puede expresarse como una composición de elementos del conjunto de funciones y terminales.

2.1.4 Población Inicial

Para crear la población inicial es necesario seguir estos pasos:

1. Definir un límite de la búsqueda en el espacio de las posibles soluciones. El límite en este caso en particular está dado por la cantidad de individuos que conforman la población. Por otro lado, también los individuos tienen un límite en la cantidad de nodos que pueden tener, esto con el fin de generar soluciones que puedan ser expresadas con relativa facilidad.
2. Generar de manera aleatoria un nodo (que inicialmente será el nodo raíz), este elemento debe pertenecer al conjunto de funciones previamente definidas.
3. Considerar la aridad⁵ del nodo para poder generar árboles válidos que cumplan con las propiedades de cerradura y suficiencia.

⁴ Ya que lenguaje empleado para el algoritmo de PG, las funciones trigonométricas reciben argumentos en radianes.

⁵ Número de argumentos que recibe una función bien definida.

4. Si se elige un elemento del conjunto de los terminales, ese nodo es una hoja del árbol⁶ y por lo tanto esa rama queda terminada. Si no es un terminal entonces es una función, por lo que se tiene que repetir el paso 2 y 3.
5. Los pasos del 2 al 4 se repiten de manera recursiva hasta que se han generado todos los individuos que conformarán la población inicial.

Mediante este procedimiento se genera la población inicial que sirve de base para llevar a cabo el proceso de búsqueda del modelo matemático para el crecimiento del parque vehicular mediante PG.

2.1.5 Función Objetivo

También conocida como función *fitness*⁷ esta función determina el desempeño de cada individuo del espacio de búsqueda, es decir que esta función nos indica cuál es el potencial del individuo evaluado para ser una posible solución. En este trabajo la función objetivo está determinada por los valores conocidos del parque vehicular. Y para determinar cuál es el fitness (o aptitud) de un individuo, se emplea el error cuadrático medio que produce el individuo al ser evaluado y comparado contra la función objetivo.

Para este trabajo la función objetivo se expresa mediante un dominio y un codominio que representan el año y la cantidad de vehículos en ese determinado año.

AÑO	0	1	2	n
VEHÍCULOS	10	20	30	m

Tabla 1: Ejemplo de dominio y codominio de la función objetivo.

2.1.5.1 Aptitud de un individuo

La aptitud de un individuo es el parámetro que determina que tan buena solución es y qué tanta probabilidad tiene este individuo de pasar a la siguiente generación. De tal manera que la forma en la que se determina la aptitud de un individuo es un aspecto de vital importancia para todo el proceso de PG.

Supongamos que se tiene al individuo *X*, a este individuo se le asocia un árbol jerárquico que representa una posible solución. Sin embargo, una vez analizado este

⁶ En teoría de gráficas y estructuras de datos, una hoja de un árbol es un nodo del cual no derivan más nodos.

⁷ En el contexto de esta tesis la función de evaluación está determinada por el error cuadrático medio que se obtiene al evaluar un individuo contra los datos de la función objetivo.

individuo puede ser visto en su forma final, que es una expresión matemática. Digamos al individuo X, se le asocia la función.

$$F(x) = \sin(X * 5) + (8 - X)$$

El dominio en el cual está función será evaluada, debe ser el mismo que el dominio de la función objetivo, es decir, el mismo rango de datos que se tenga en la función objetivo. Una vez evaluada esta función generará un vector de valores.

Supongamos que este individuo es evaluado empleando la función objetivo.

x	1	2	3	4	5
$\Gamma(x)$	6	5.5	5.7	5	3

Tabla 2: Función objetivo de ejemplo.

El vector que se obtiene al evaluar al individuo X en este dominio es el siguiente.

x	1	2	3	4	5
$F(x)$	6.041	5.455	5.650	4.912	2.867

Tabla 3: Resultado de la evaluación del individuo X.

El error cuadrático medio (ECM) de este individuo está dado por la expresión.

$$ECM = \frac{\sum_{i=1}^5 (F_i - \Gamma_i)^2}{5} = 0.0063278$$

Es decir que la aptitud de este individuo es 0.0063278, lo que significa que es un individuo bastante apto para transmitir sus genes a la próxima generación o incluso para ser un modelo representativo.

De manera general la forma en la que se calcula el ECM de un individuo es mediante la expresión.

$$ECM = \frac{\sum_{i=1}^n (F_i - \Gamma_i)^2}{n}$$

Donde:

- F_i : el resultado de evaluar al individuo en el valor i-ésimo.

- f_i : el i -ésimo valor conocido de la función objetivo.
- n : número total de datos.

2.1.6 Operadores Genéticos

- **Cruzamiento:** este es el operador principal, se aplica a un par de individuos, lo que significa que un individuo intercambia un subárbol⁸ adecuado extraído de el mismo, por otro subárbol adecuado de algún otro individuo de la población. En PG el operador de cruzamiento genera expresiones matemáticas completamente diferentes a las de sus padres, esto añade una gran efectividad al operador de cruzamiento. Una vez que se tiene a todos los individuos juntos en un vector ordenado, se toma al individuo i y al $i+1$ y se realiza la cruce. Los individuos resultantes deben cumplir con el número máximo de nodos permitidos, de no ser así se les asigna una penalización que afecta su aptitud, lo cual los hace individuos no aptos para pasar a la siguiente generación y la selección natural los descarta.
- **Mutación:** este operador se aplica a un solo individuo de la población y lo que sucede cuando un individuo muta es que ese individuo sustituye un nodo completo y por lo tanto un subárbol del mismo o simplemente se cambia el valor contenido en este nodo, ya sea por un elemento adecuado del conjunto de los terminales o de las funciones. Para poder llevar a cabo la mutación es importante fijarse en la aridad de las funciones del nodo en el cual se llevará a cabo la mutación. De igual manera que con la cruce, si el individuo resultante después de la muta excede la cantidad de nodos permitidos, se le asigna una penalización.

En la figura 2 se muestra un ejemplo gráfico de cómo es que la cruce se lleva a cabo.

⁸ Subconjunto de nodos y vértices de un árbol, que también cumplen con la definición de árbol.

Modelado de tráfico vehicular

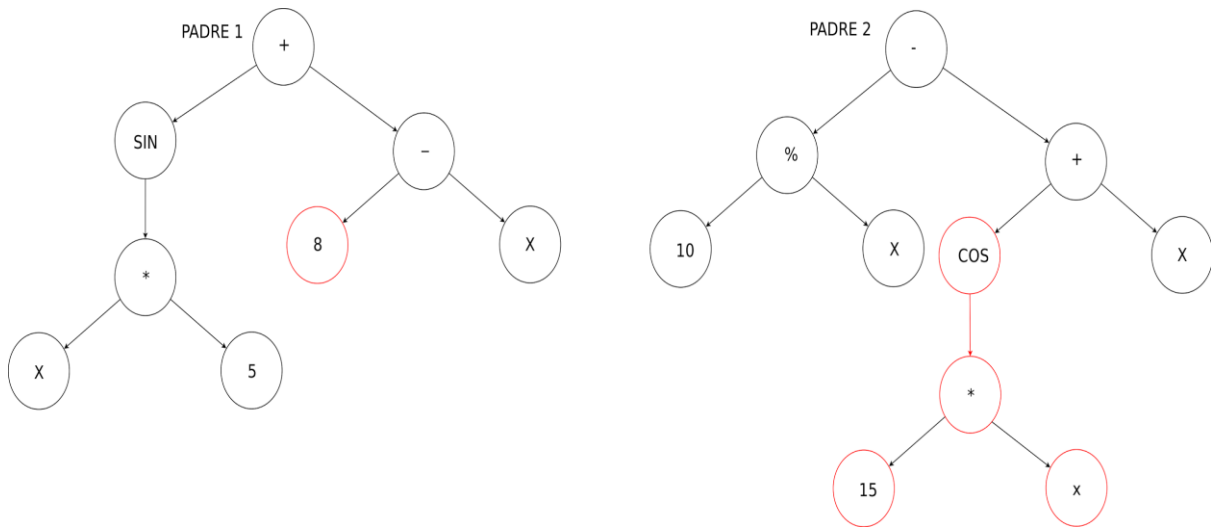


Figura 2: Cruce entre individuos.

- ★ El PADRE 1 representa la solución: $F(X) = \sin(X * 5) + (8 - X)$
 - ★ El nodo candidato a ser intercambiado por el operador de cruzamiento es el nodo cuyo valor es un “8” y está marcado en color rojo.
- ★ El PADRE 2 representa la solución: $F(X) = \left(\frac{10}{X}\right) - ((\cos(15 * X) + X)$
 - ★ El nodo candidato a ser intercambiado por el operador de cruzamiento es el nodo cuyo valor es un “coseno” y está marcado en rojo.

De tal manera que después del cruzamiento los individuos resultantes son

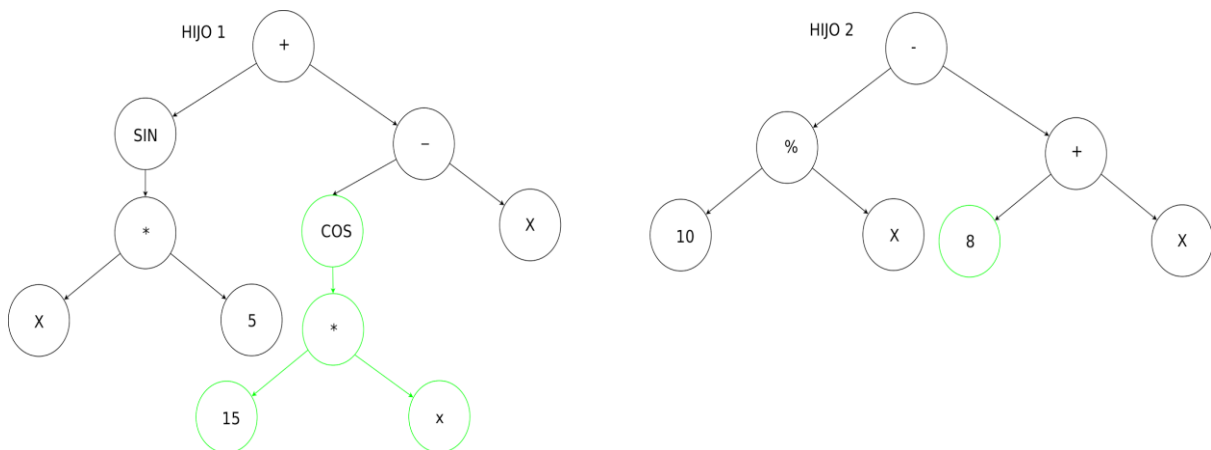


Figura 3: Resultado del cruce de individuos.

- ★ El HIJO 1 representa la solución: $F(X) = \sin(X * 5) + ((\cos(15 * X) - X)$
- ★ El HIJO 2 representa la solución: $F(X) = \left(\frac{10}{X}\right) - (8 + X)$

En la figura 4 se muestra el operador de mutación.

Modelado de tráfico vehicular

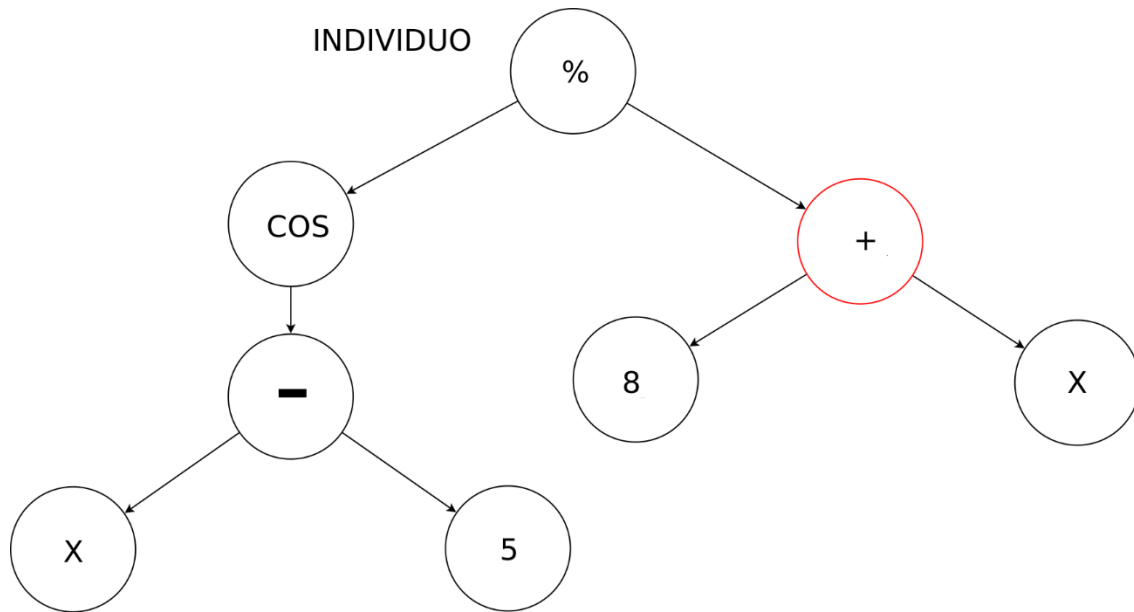


Figura 4: Operador de mutación.

★ Este INDIVIDUO representa la solución: $F(X) = \frac{\text{COS}(X-5)}{(8+X)}$

★ El nodo que mutará es el nodo de la operación “+” que se encuentra en color rojo.

Y después de haberse llevado a cabo la mutación este individuo se ve de la siguiente manera.

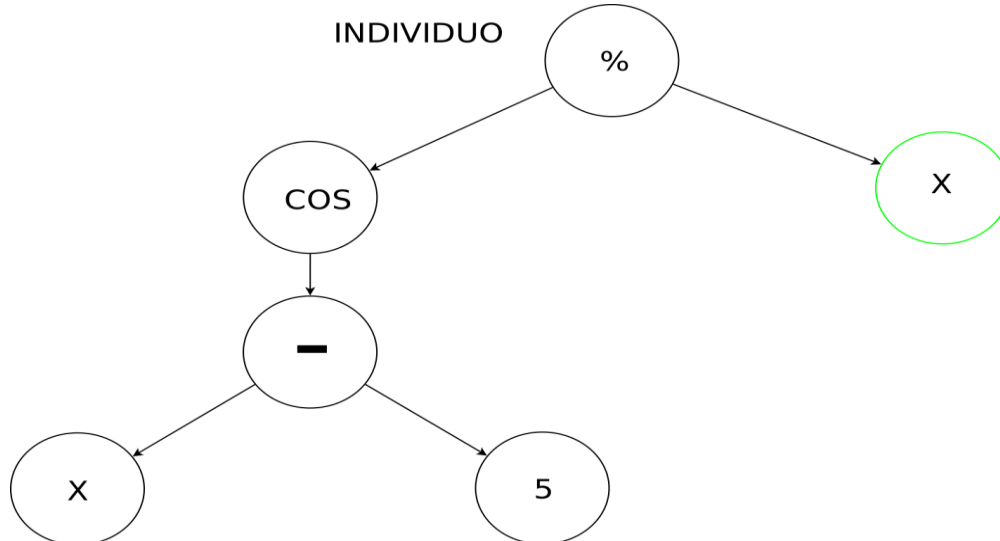


Figura 5: Resultado de la mutación.

Después de haberse llevado a cabo la mutación, el individuo ahora representa la solución.

★ Este INDIVIDUO representa la solución: $F(X) = \frac{\text{COS}(X-5)}{(X)}$

Así que el proceso completo de PG que se llevará a cabo en el siguiente capítulo se puede mostrar de manera gráfica en la siguiente figura.

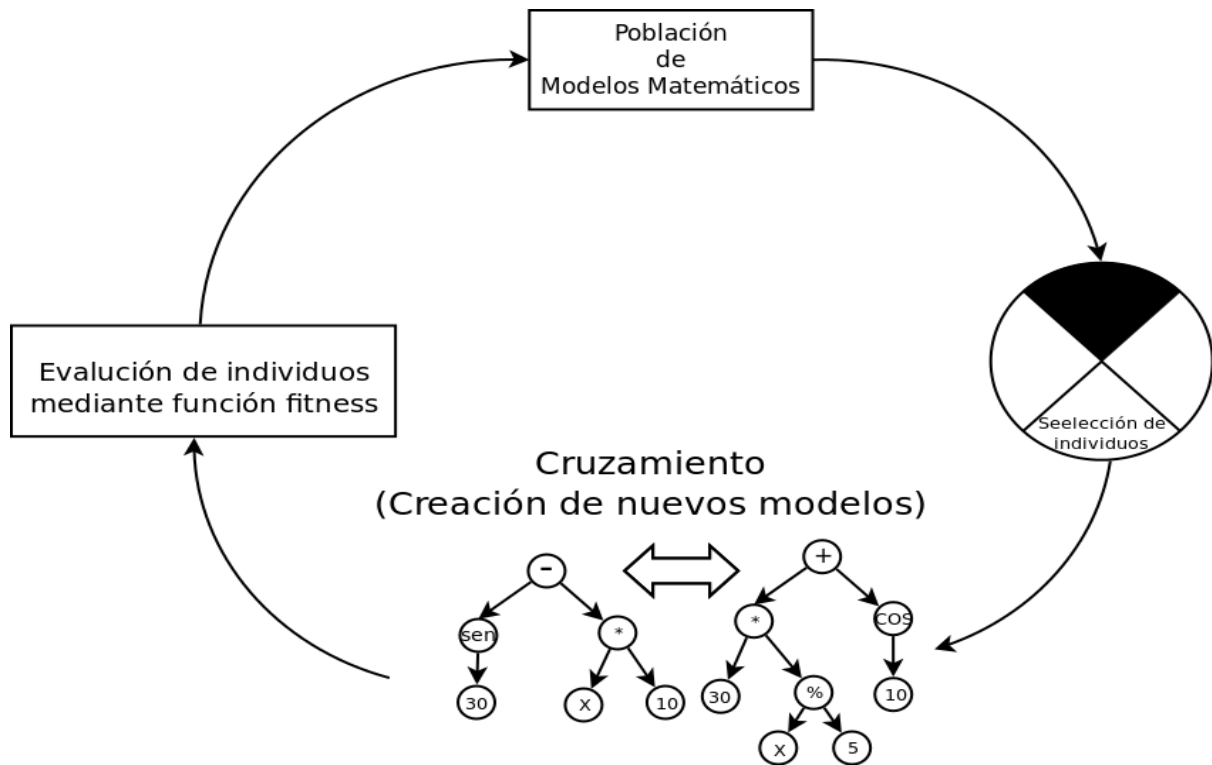


Figura 6: Proceso general de la Programación Genética.

De manera tal que el proceso de PG se puede ver en su forma de algoritmo de la siguiente manera.

1. Crear una población inicial de individuos.
2. Seleccionar los individuos aptos para transmitir sus genes a la siguiente generación. En el contexto de este trabajo, este proceso se lleva a cabo mediante torneo.
3. Realizar el cruce con los individuos seleccionados. Una vez que se llevó a cabo el operador de cruce, se procede a llevar a cabo la mutación con los individuos resultantes del cruce. Esto con el objetivo de formar la nueva generación (o población).
4. Evaluar a los individuos de la nueva generación, para que posteriormente puedan ser seleccionados o en su defecto para seleccionar al mejor individuo.
5. Repetir los pasos del 2 al 4 hasta cumplir con la condición de paro. Generalmente se emplea un número determinado de generaciones que cumple con la finalidad de establecer el número de veces que se llevara a cabo este proceso, aunque no es la única forma de determinar cuándo termina el algoritmo de PG.

2.1.7 Método de Selección (con elitismo)

Una vez que se tiene una población completa, la manera en la que se eligen los individuos que transmitirán sus genes a la siguiente generación es mediante torneo. Se eligió esta forma de selección ya que se garantiza que el mejor individuo de dicha generación transmitirá al menos una vez sus genes al menos una vez a la siguiente generación, a esto se le conoce en el ámbito de la PG como selección con elitismo.

2.1.7.1 Torneo

El torneo funciona de la siguiente manera, supongamos que ya se tiene una población de 5 individuos, cada individuo tiene asociada una aptitud. Por lo tanto, podemos pensar que la población se ve cómo un vector que luce de la siguiente manera.

Individuo	Aptitud
0	0.02548
1	0.12255
2	1.02259
3	5.00056
4	0.99526

Tabla 4: Población de individuos relacionados con su aptitud.

De esta población se eligen 5 individuos al azar que competirán con todos y cada uno de los individuos en la población. Supongamos que los individuos elegidos

Modelado de tráfico vehicular

fueron, el individuo 1, el individuo 3, el individuo 2, el individuo 4 y el individuo 1. De tal manera que el torneo luce de la siguiente forma.

Individuo	Aptitud	V.S.	Individuo	Aptitud
0	5.00056		1	0.02548
1	0.02548		3	0.12255
2	0.99526		2	0.99526
3	0.12255		4	1.02259
4	1.02259		1	0.02548

Tabla 5: Organización del torneo.

No importa que un individuo no haya sido elegido para formar el vector virtual⁹ de contendientes, cómo es el caso del individuo 0. El torneo garantiza que todos y cada uno de los individuos de la población tienen al menos una oportunidad de transmitir sus genes a la siguiente generación, siempre y cuando posean una buena aptitud. Después de haber llevado a cabo la selección mediante torneo, la nueva población luce de la siguiente manera.

Individuo	Aptitud
0	0.02548
1	0.02548
2	0.99526

⁹ En este contexto, virtual significa que no fue necesario emplear recursos computacionales extra para poder llevar a cabo el torneo. Es decir que no fue necesario crear un nuevo vector de individuos para almacenar a los contendientes de la población inicial.

Modelado de tráfico vehicular

3	0.12255
4	0.02548

Tabla 6: Resultado del torneo.

Cómo puede notarse en la nueva generación, cada vez que el individuo más apto (individuo 1 de la población inicial) contenido, resultó victorioso y logró transferirse tal cual a la siguiente generación. En este ejemplo este individuo transferirá 3 veces sus genes a la siguiente generación, sin embargo, el torneo garantiza que al menos una vez este individuo logrará transferir sus genes. Esto propicia que las siguientes generaciones tengan mayor probabilidad de que sean mejores.

2.2 Simulación de Tráfico Vehicular

Respecto a la modelación y simulación de tráfico vehicular se puede decir que hay 3 categorías de modelos:

A) Modelos Macroscópicos: Se enfocan en ver el tráfico como un fluido, analizan a cada vehículo como una partícula y le dan mayor importancia al flujo que a cada vehículo en particular. Están basados en ecuaciones de dinámica de fluidos, en los cuales intervienen variables agregadas por intensidad espacial o temporal y su principal característica es que interpretan el flujo vehicular, como un fluido desplazándose a través de un ducto.

B) Modelos Microscópicos: Se enfocan en la interacción individual de cada vehículo con su entorno, es decir que analizan cuál es la relación del vehículo en cuestión con los vehículos vecinos para cada paso de tiempo. Normalmente se emplean EDO's¹⁰ que emplean las leyes de Newton para describir el movimiento de cada vehículo. En esta categoría también podemos encontrar los modelos basados en Autómatas Celulares o los modelos de tipo Car-Following.

C) Modelos Cinéticos: Este tipo de modelos emplean ecuaciones cinéticas de tipo Boltzmann. Sin embargo, se podría decir que los modelos cinéticos son un punto intermedio entre los modelos macroscópicos y los modelos microscópicos.

¹⁰ Ecuaciones Diferenciales Ordinarias

Modelado de tráfico vehicular

En este trabajo se emplea SUMO para llevar a cabo la simulación de los datos arrojados por el modelo matemático creado mediante programación genética. SUMO utiliza un modelo de tipo Car-Following para realizar las simulaciones vehiculares.

La sección 4.5 aborda más a fondo los conceptos empleados en la simulación de tráfico vehicular.

2.3 Resumen

En este capítulo se dio una introducción a los conceptos fundamentales de la PG y se muestra cómo es que será generado el modelo para la predicción del crecimiento del parque vehicular. También se mencionan los diferentes modelos existentes para la simulación de tráfico vehicular y en particular el modelo empleado por el simulador que se utiliza en esta tesis. En el siguiente capítulo se mostrarán los experimentos con datos reales de Zona Metropolitana del Valle de México y de algunas otras ciudades que llevaron a generar el modelo matemático que se propone en esta tesis.

Capítulo 3

MODELO MATEMÁTICO

En este capítulo se presentará el modelo matemático basado en programación genética y regresión simbólica para estimar el crecimiento del parque vehicular en la ZMVM. Por otra parte, se detallan de manera minuciosa los elementos necesarios para llevar a cabo el proceso de construcción de dicho modelo. También se discutirán las peculiaridades que se encontraron al realizar los experimentos necesarios para la construcción del modelo. Los datos de salida del modelo presentado en este capítulo sirven para tener un estimado del momento en el cual se alcanzará el colapso vial en la ZMVM en caso de no aplicar medidas preventivas y correctivas. Para el caso de la delegación Cuauhtémoc y para las ciudades empleadas en la comparativa del capítulo 4 se empleó el mismo procedimiento, pero para una mejor comprensión del trabajo únicamente se muestran los resultados de dichos modelos en el siguiente capítulo. Los datos de salida del modelo de la delegación Cuauhtémoc son empleados como datos de entrada para la simulación del capítulo 5.

3.1 Diccionario de Operaciones y Terminales

Las operaciones y terminales son un elemento de vital importancia para el desempeño de la programación genética, en este caso en particular son incluso más importantes que otros elementos que analizaré más adelante. Es por esto que este capítulo comienza definiendo en este apartado los operaciones y terminales que se emplearon en este trabajo.

Para la creación del modelo fue necesario definir un diccionario de operaciones y de variables terminales:

- Variables terminales: dado que los datos históricos que se emplean para modelar el crecimiento del parque vehicular están definidos en términos del año como variable independiente y cantidad de vehículos como variable dependiente, Las variables terminales que se emplearon en el diccionario fueron variables aleatorias y el año en cuestión. Los modelos con mejor aptitud en la mayoría de los casos elegían como elemento terminal únicamente a la variable independientes, es decir el año en cuestión, así que en a pesar de que, si se consideran variables aleatorias para el procedimiento de PG, el elemento terminal que se usa en los modelos generados, es la variable independiente X.
- Operaciones:
 - Aridad 2:

- + (suma)
- - (resta)
- * (multiplicación)
- % (división)
- Aridad 1:
 - exponencial..... e^x
 - seno..... $sen(x)$
 - coseno..... $cos(x)$
 - tangente..... $tan(x)$
 - seno hiperbólico..... $sinh(x)$
 - coseno hiperbólico..... $cosh(x)$
 - tangente hiperbólica..... $tanh(x)$
 - sigmoide..... $\frac{1}{1+e^{-x}}$

Después de realizar múltiples pruebas para calibrar¹¹ el algoritmo de PG que genera los modelos matemáticos, se llegó a la conclusión de que este conjunto de operaciones generaba los modelos que se aproximaban de manera más fiel a los datos históricos.

Las funciones trigonométricas fueron incluidas dado que los datos de entrada podían ser descritos empleando estas funciones. Y al evaluar los modelos generados fue notable que la PG elegía en múltiples ocasiones estas operaciones.

En un principio, se incluyó la operación **sigmoide**¹² en el diccionario de operaciones, pero después de realizar las pruebas para calibrar el algoritmo, fue evidente que esta decisión generaba modelos cuya aptitud era pobre, lo cual significa que son modelos que distan de los resultados esperados, así que por esta razón la operación sigmoide fue eliminada del diccionario.

3.1.1 Importancia de la Función Exponencial

Múltiples fenómenos en la naturaleza están estrechamente relacionados con la función exponencial, fenómenos tales como la mitosis¹³ celular o el crecimiento urbano. Dado que la forma como crece la cantidad de habitantes de alguna ciudad puede ser modelado empleando la función exponencial y ya que la cantidad de

¹¹ En el contexto de este trabajo, el proceso de calibración consistió en encontrar los parámetros idóneos como: población inicial, número de generaciones y en especial operadores aritméticos, para obtener modelos matemáticos confiables.

¹² Muchos procesos naturales y curvas de aprendizaje de sistemas complejos muestran una progresión temporal desde unos niveles bajos al inicio, hasta acercarse a un clímax transcurrido un cierto tiempo; la transición se produce en una región caracterizada por una fuerte aceleración intermedia. La **función sigmoide** permite describir esta evolución.

¹³ Proceso a nivel celular que precede inmediatamente a la división celular y consiste en el reparto equitativo de material hereditario (ADN).

vehículos en una ciudad guarda una relación directamente proporcional a la cantidad de habitantes de la misma, no es de sorprender que la función exponencial fuese una de las operaciones más importantes dentro del diccionario que se empleó en el algoritmo de PG que forma parte de este trabajo.

Los experimentos que se llevaron a cabo inicialmente no incluían en el diccionario la operación exponencial. Como resultado los modelos generados por el algoritmo de PG eran modelos que no se ajustaban a los datos de entrada, es decir que la aptitud (o error cuadrático) era muy elevado, lo que significa que eran modelos poco confiables. Se realizaron muchos experimentos para finalmente concluir que, sin la operación exponencial, sería imposible generar un modelo suficientemente adecuado que sirviera para predecir el crecimiento en el parque vehicular. En la mayoría de las ciudades que se analizan en este trabajo la población de cada ciudad crece y por lo tanto el parque vehicular. Uno de los pocos casos en los que la cantidad de vehículos disminuyó, es por ejemplo la ciudad de Detroit (U.S.A.), sin embargo, esto fue debido a que la población disminuyó.

Después de agregar la operación exponencial los modelos matemáticos resultantes mostraron una mejor aptitud y se ajustaban más a los datos de entrada. Esto significa que de manera cuantitativa se puede mostrar que el crecimiento del parque vehicular crece de manera exponencial, lo cual está relacionado con el crecimiento de la población. Esta relación ya era bien conocida, sin embargo, con el trabajo realizado en esta tesis se muestra de manera fehaciente que el crecimiento del parque vehicular tiene un comportamiento exponencial, al menos en la ZMVM.

3.1.2 Protección de las Operaciones

Al momento de crear, cruzar o incluso evaluar a los individuos es necesario ejecutar las operaciones definidas en el diccionario. Potencialmente se podría dar el caso de que la evaluación de alguna de estas operaciones estuviese indeterminada como es el caso de la división ya que una división por cero no está definida y por lo tanto eso produciría un error en el algoritmo.

Por otro lado, podría darse el caso de que la evaluación de la función exponencial, tangente, seno hiperbólico o coseno hiperbólico dieran como resultado valores de tal magnitud que se desbordaran y como consecuencia se tuvieran valores de tipo **NAN**¹⁴ or **infinity**¹⁵. Es por esto que se siguió la misma metodología que

¹⁴ Proviene del acrónimo en inglés **Not a Number** (en español: **no es un número**). Este acrónimo se usa generalmente en algunos lenguajes de programación para expresar un resultado imposible de calcular, como el caso de las raíces negativas, indeterminaciones, etc., retornando *NaN* como resultado de la operación.

¹⁵ En lenguajes de programación esta etiqueta se emplea para hacer referencia a un valor tan grande o pequeño que no puede ser representado en el sistema numérico empleado en un determinado lenguaje de programación.

sugiere Koza [13]. Así que la estrategia empleada es la siguiente. En caso de que una operación arroje un resultado como los ya antes mencionados, se sustituye el resultado por el valor 999999, mismo que a pesar de ser un valor elevado en términos de magnitud, aún puede ser representado. Esto da como resultado un individuo con una aptitud pobre en el sentido de que no es un modelo adecuado, lo que implica que dicho individuo sea descartado para el procedimiento de cruza y por lo tanto sus genes no sobreviven para generaciones posteriores.

3.2 Parámetros del Algoritmo de PG

Como parte del procedimiento de regresión simbólica en PG es necesario definir 4 elementos: tamaño de la población, número de generaciones, profundidad de los individuos y diccionario de operaciones y terminales, mismos que determinarán el desempeño de los individuos y por lo tanto la exactitud del modelo obtenido. Estos parámetros se determinaron con base en la experimentación, es decir que se ejecutó el algoritmo múltiples veces con diferentes parámetros y finalmente se logró calibrar el algoritmo para obtener el modelo adecuado.

3.2.1 Población Inicial

La población inicial es uno de los parámetros que determina la rapidez o lentitud con la que el algoritmo de PG encontrará un individuo con la mejor aptitud. En un principio se experimentó con una población inicial de **100 individuos**, los modelos resultantes mostraban una baja aptitud. Se incrementó la cantidad de individuos de la población inicial hasta llegar a **1500 individuos**, sin embargo el algoritmo tardaba alrededor de 10 minutos para terminar.

Finalmente se llegó a la conclusión de que la cantidad de individuos correcta en la población inicial para mantener un equilibrio entre una buena aptitud y un desempeño ágil del algoritmo eran **1200 individuos**. De tal manera que se comprobó que 1200 individuos es una cantidad adecuada para generar modelos tanto para la ZMVM como para otras ciudades.

Otro de los parámetros que se empleó para determinar las características de un individuo fue su tamaño, es decir, el número máximo de nodos que lo pueden conformar. Y al igual que el resto de los parámetros, este se determinó mediante pruebas de desempeño, el valor que mostro ser un número adecuado para mantener una buena relación entre el desempeño y la facilidad para interpretar a los individuos fue 48 nodos, es decir, $7^2 - 1$ nodos. Este valor se deriva de las propiedades de los árboles binarios y también probó ser un valor idóneo en este trabajo.

3.2.2 Número de Generaciones

El número de generaciones es la condición de paro para este algoritmo, existen algunas variantes en las cuales la condición de paro puede ser la cantidad de nodos de los individuos o la aptitud del mejor individuo. Pero para este caso en particular el número de generaciones determina en qué momento se detiene el algoritmo. Este número representa la cantidad de veces que los mejores individuos se cruzaran para generar individuos con mejores aptitudes. Si la cantidad de generaciones es elevada el algoritmo tomará más tiempo, pero es un espacio de búsqueda mayor, ya que se dan más cruza, aunque esto no forzosamente significa que se generarán individuos con mejor aptitud. De manera muy similar a la población inicial, se comenzó experimentando con **100 generaciones**, este número produjo resultados casi de manera inmediata, pero al igual que con la población inicial los individuos resultantes no mostraban un buen desempeño. Se incrementó el número hasta llegar a **900 generaciones**, pero el análisis de los individuos mostraba que los individuos resultantes a partir de la **generación 400** tenían una aptitud considerablemente mejor que el resto de los individuos de las generaciones posteriores. Incluso se encontró que el mejor individuo generado a partir de la generación 400 (en promedio) se mantenía sin cambios hasta la generación 900. Esto indica que los individuos resultantes después de la generación 400 eran modelos muy similares o incluso idénticos a los ya generados previamente. Se dieron pocos casos en los cuales después de la generación 400 se obtenían individuos más aptos, así que finalmente se tomó la decisión de fijar este número en **500 generaciones**.

3.2.2.1 Cruce

El cruce de individuos, aunque no es un parámetro del algoritmo si es importante mencionar que se llevó a cabo mediante torneo. Existen otras formas de cruza y sobretodo de selección de individuos, sin embargo en este trabajo se empleó la cruza tradicional como se mencionó en el capítulo anterior.

3.2.2.2 Método de selección

Los individuos se eligieron mediante torneo, en el capítulo anterior se mostró por qué se eligió este método y qué beneficios conlleva el mismo.

3.2.3 Función Objetivo

A diferencia de problemas de optimización, en los cuales la función objetivo es equivalente a la función fitness, en este caso la función objetivo determina cómo es

Modelado de tráfico vehicular

que se espera que los individuos se comporten, es decir que es la función que se busca aproximar y en este caso es la función que se quiere construir a fin de obtener el modelo matemático para predecir el comportamiento del crecimiento vehicular a futuro. Por otro lado la función fitness es el ECM que emplea la función objetivo para calcular la diferencia entre esta y un modelo dado. En términos simples la función objetivo de este trabajo está determinada por los datos históricos del parque vehicular. La función objetivo es la función que nos sirve para crear individuos que se aproximen a esta, una vez que se generó un individuo con una buena aptitud, los valores que arroje la evaluación de este individuo son considerados una proyección a futuro. Dado que en este trabajo se generaron múltiples modelos la función objetivo depende de la ciudad.

- ZMVM: la función objetivo está definida por los datos históricos del parque vehicular de los años 1999 a 2014.
- L.A, Londres y Beijing.: la función objetivo está definida por los datos históricos del parque vehicular de los años 2000 a 2015.
- NYC: la función objetivo está definida por los datos históricos del parque vehicular de los años 2007 a 2015.

Para la ZMVM la función objetivo se muestra en la figura 7.

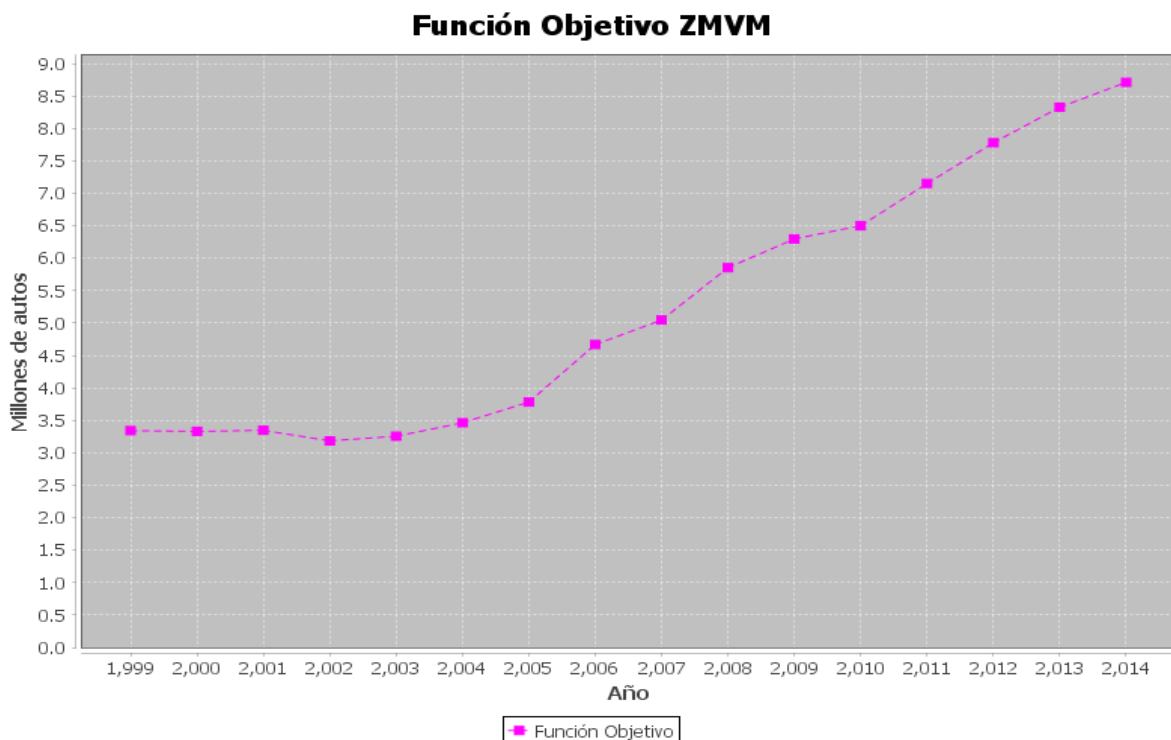


Figura 7: Función objetivo ZMVM.

En la figura 7, se muestra la función objetivo, la cual está ilustrada en color morado y muestra la cantidad de vehículos desde el año 1999 hasta el año 2014.

3.3 Datos Históricos del Parque Vehicular

Para obtener datos confiables sobre el parque vehicular fue necesario realizar una búsqueda exhaustiva en internet para encontrar fuentes confiables ya que si los datos no eran confiables los resultados de esta tesis serían incongruentes con la realidad.

3.3.1 ZMVM

En este trabajo inicialmente se pensó en considerar los datos históricos de la CDMX¹⁶ pero dadas las condiciones de la CDMX cómo la movilidad urbana o las zonas laborales se decidió que para poder generar modelos correctos que sirvan para poder predecir el comportamiento del parque vehicular y por ende el tráfico generado por el parque vehicular, fue claro que era necesario considerar una zona más amplia que tomará en cuenta las condiciones antes mencionadas. Esta decisión se tomó debido a que a pesar de que en la CDMX se tiene un censo del parque vehicular, la cantidad de vehículos que ingresan del estado de México o del estado de Hidalgo para laborar en la CDMX es considerable y contribuye al tráfico que se tiene en la misma.

Es por este motivo que se decidió considerar los datos del parque vehicular de la ZMVM. La ZMVM comprende la CDMX y 59 municipios del Estado de México y un municipio de Hidalgo.

Además de las delegaciones de la CDMX los municipios que comprenden la ZMVM se listan a continuación.

Acolman	Coyotepec	Hueyoxtla	Nextlalpan
Amecameca	Cuautitlán	Huixquilucan	Nicolás Romero
Apaxco	Chalco	Isidro Fabela	Nopaltepec
Atenco	Chiautla	Ixtapaluca	Otumba
Atizapán de Zaragoza	Chicoloapan	Jaltenco	Ozumba

¹⁶ Abreviatura para Ciudad de México.

Modelado de tráfico vehicular

Atlautla	Chiconcuac	Jilotzingo	Papalotla
Axapusco	Chimalhuacán	Juchitepec	La Paz
Ayapango	Ecatepec de Morelos	Melchor Ocampo	San Martín de las Pirámides
Coacalco de Berriozábal	Ecatzingo	Naucalpan de Juárez	Tecámac
Cocotitlán	Huehuetoca	Nezahualcóyotl	Temamatla
Temascalapa	Tlalmanalco	Tenango del Aire	Tlalnepantla de Baz
Teoloyucán	Tultepec	Teotihuacán	Tultitlán
Tepetlaoxtoc	Villa del Carbón	Tepetlixpa	Zumpango
Tepotzotlán	Cuautitlán Izcalli	Tequixquiac	Valle de Chalco Solidaridad
Texcoco	Tonanitla	Tezoyuca	

Tabla 7: Municipios que conforman la ZMVM.

Uno de los objetivos de esta tesis es que además de generar un modelo que sirva para predecir el comportamiento del parque vehicular en la ZMVM y con este modelo tener una idea del límite de vehículos en la XMVM, también se planea emplear los resultados de la evaluación de dicho modelo para generar una simulación y dicha simulación considera 3 tipos de vehículos motorizados, vehículos particulares, motocicletas y vehículos de carga transporte. Vale la pena aclarar que para el momento en el que se llevó a cabo esta tesis los datos disponibles para la ZMVM (considerando los 3 tipos de vehículos motorizados) sólo estaban disponibles en todos los municipios a partir del año 1999 y hasta el año 2016. Los datos del parque vehicular de la ZMVM se tomaron del INEGI[18] con fecha 27 de Julio del 2016.

3.3.2 Otras ciudades

De igual manera que para la ZMVM los datos que se recabaron para estas ciudades comprenden diferentes rangos.

- Beijing[19], L.A.[20] y Londres[21]: sus datos comprenden desde el año 2000 y hasta el año 2015.
- NYC[22]: sus datos comprenden desde el año 2007 y hasta el año 2015.
- Río de Janeiro[23]: sus datos comprenden del año 2001 y hasta el año 2012.

3.4 Modelo Matemático

En procedimientos como el de PG, es una práctica normal, si no es que obligatoria, iterar el proceso muchas veces para comprobar que los modelos generados son modelos confiables y que al ser empleados no arrojen datos incoherentes. Para la ZMVM se generaron 50 modelos mediante PG tomando como datos de entrada todos los datos del parque vehicular hasta el año 2014 (datos disponibles en la página del INEGI). Ya con estos 50 modelos se graficaron todos ellos y se obtuvo un modelo promedio, esto con la finalidad de mostrar que los modelos obtenidos eran confiables. La gráfica que se obtuvo se muestra en la figura 8.

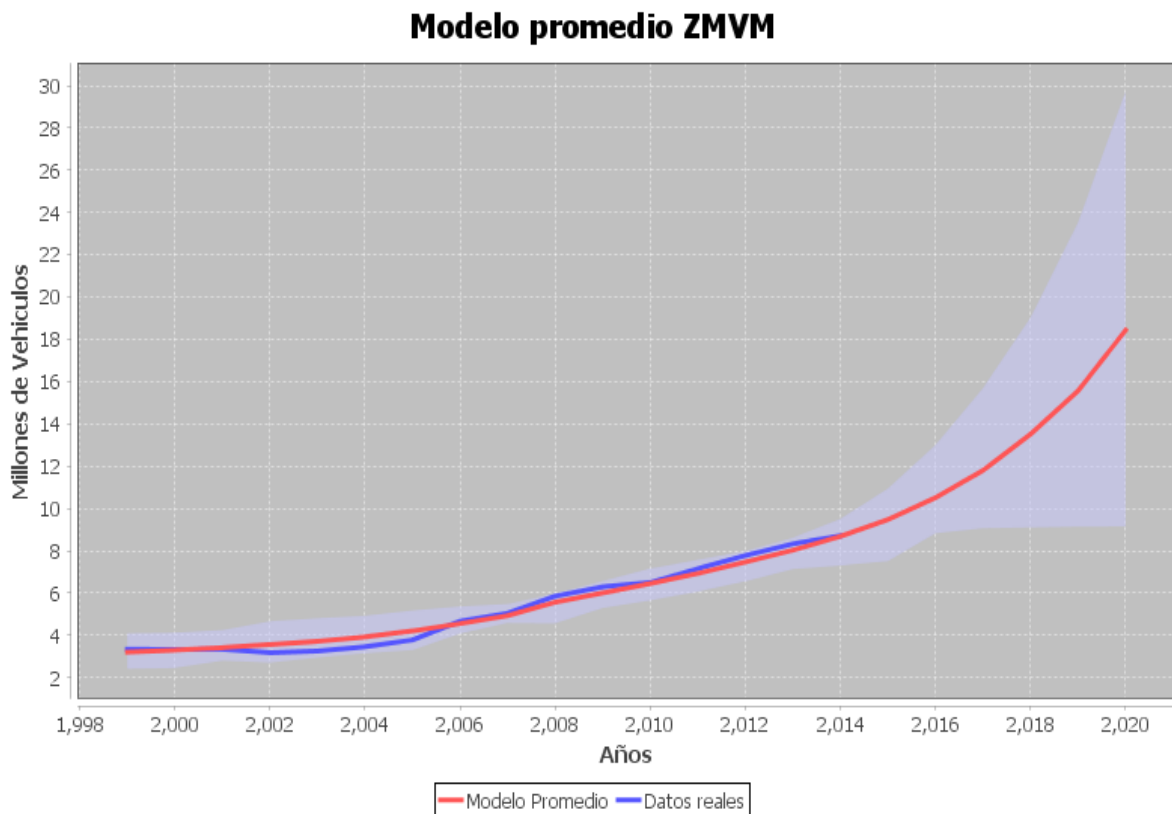


Figura 8: Grafica de los 50 modelos generado, en rojo el modelo promedio, azul, datos reales y en sombra morada el resto de los modelos.

En color rojo se muestra el promedio de los 50 modelos y en color azul los datos reales. La sombra en color morado representa el resultado de los modelos restantes. Puede notarse en la figura 8 que los modelos generados se mantienen con

una tendencia homogénea, es decir que no hay modelos generados que muestre valores exageradamente elevados o exageradamente pequeños para los años posteriores a los que se tienen datos reales. Esto nos indica que los modelos generados, son modelos confiables, es por este motivo que se decidió realizar múltiples modelos para poder promediarlos.

El mejor modelo fue el modelo 11 que tuvo un ECM de 1,066,461 vehículos y se muestra en la figura 9.

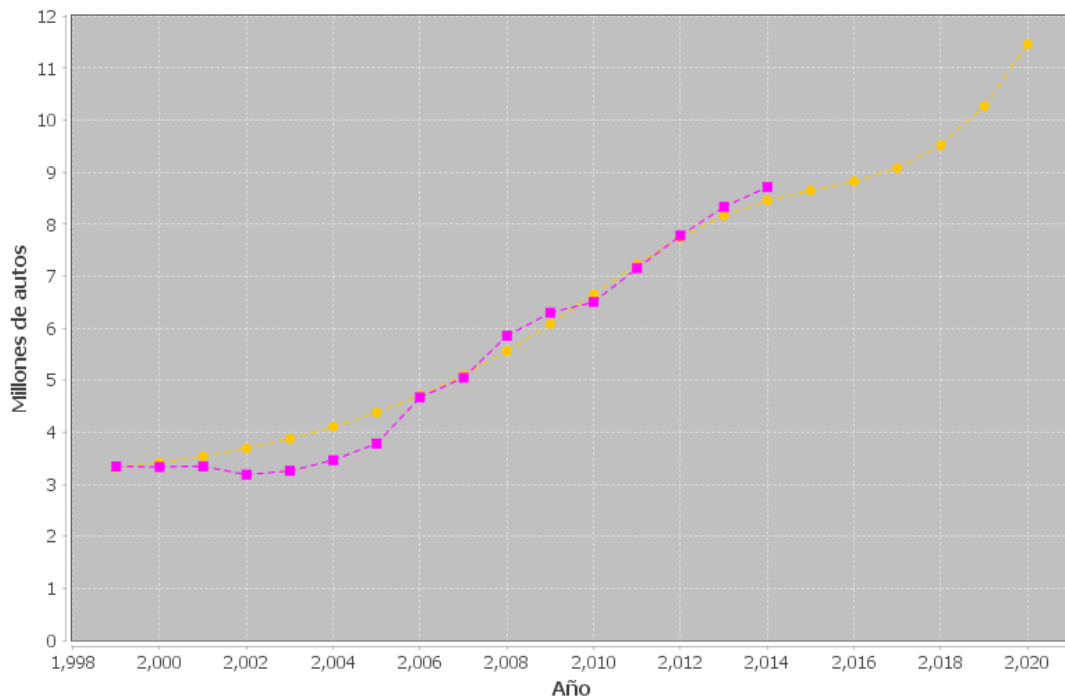


Figura 9: Mejor modelo ZMVM, morado datos reales, amarillo evaluación del modelo.

En la figura 9 en color morado se muestran los datos reales del parque vehicular y en color amarillo los valores que arrojó el mejor modelo de la ZMVM para los años 1999, 2000,, 2020, cuya expresión matemática es la siguiente.

$$F(x) = (\text{Senh}(\text{Tan}(\text{Tan}(\text{Tan}(\text{Log}(x))))) + e^{\text{Tan}(\text{Cos}(\text{Cos}(e^{e^{\text{Sin}(\text{Cos}(\text{Tan}(\text{Tan}(\text{Log}(x)))))}))}))$$

Esta expresión matemática, es el mejor individuo que genero la PG después de haberlo interpretado mediante regresión simbólica. $F(x)$ representa el valor anual del parque vehicular para el año X . Los resultados de evaluar esta expresión para cada uno de los años que comprende este trabajo se muestran en la tabla 8.

Modelado de tráfico vehicular

Año	Datos reales	Modelo
1999	3,340,492	3,323,254
2000	3,331,027	3,416,782
2001	3,345,957	3,537,603
2002	3,184,166	3,689,070
2003	3,255,373	3,875,619
2004	3,462,640	4,102,890
2005	3,782,822	4,377,727
2006	4,668,502	4,707,785
2007	5,047,400	5,100,342
2008	5,854,298	5,559,611
2009	6,297,044	6,081,886
2010	6,500,465	6,648,779
2011	7,151,484	7,221,674

Modelado de tráfico vehicular

2012	7,779,345	7,744,665
2013	8,327,810	8,163,205
2014	8,710,806	8,454,307
2015	-	8,647,941
2016	-	8,821,925
2017	-	9,077,829
2018	-	9,522,043
2019	-	10,267,551
2020	-	11,458,044

Tabla 8: Datos históricos del parque vehicular de la ZMVM y resultado del modelo de la ZMVM.

3.4.1 Validación del modelo

Una vez que se identificaron los parámetros adecuados para el correcto funcionamiento del algoritmo, y ya que se tenían modelos confiables, se procedió a generar modelos de prueba con una porción de los datos iniciales, es decir que únicamente se alimentó el algoritmo con los datos del parque vehicular de la ZMVM hasta el año 2011. Esto con la intención de calibrar el algoritmo y comprobar que el modelo generado mostrará el comportamiento esperado, contrastándolo con el resto de los datos ya conocidos.

Los modelos de prueba generados, se desempeñaron como se esperaba, aunque con una aptitud mayor que la del modelo final, esto significa que dichos

Modelado de tráfico vehicular

modelos tenían un ECM mayor al de los modelos que si empleaban la totalidad de los datos. Sin embargo, esto es completamente natural ya que, a mayor información de entrada los modelos resultantes tendrán una mejor aptitud.

La gráfica del modelo promedio con datos hasta el año 2011 se muestra en la siguiente figura.

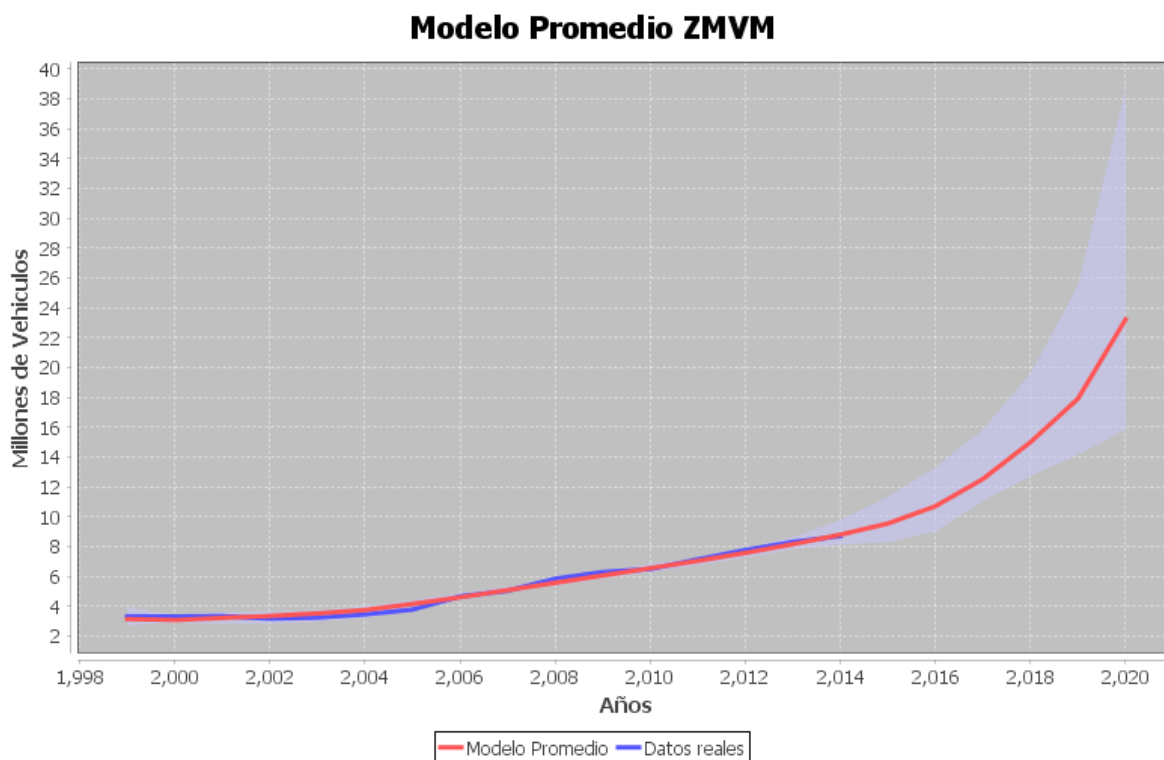


Figura 10: Gráfica de los modelos generados mediante PG considerando datos reales hasta el año 2011. En rojo el modelo promedio, azul datos reales, sombra morada el resto de los modelos.

Es notable que en la gráfica de la figura 10 ambas curvas casi siempre están sobrepuestas, incluso después del año 2011, que es el último año que se empleó para definir la función objetivo. Por otro lado, ya que se no se empleó la totalidad de los datos para definir la función objetivo la cantidad de vehículos en la proyección a futuro es mayor que en la gráfica en la cual si se hizo de todos los datos disponibles.

Las aptitudes (ECM) de los modelos generados para la validación del modelo se muestra en la tabla 9

No. Modelo	ECM
1	9,665,448
2	8,825,556
3	10,522,971
4	9,269,741
5	8,930,125
6	10,002,256
7	9,856,156

Modelado de tráfico vehicular

8	8,394,075
9	9,965,845
10	8,987,456

Tabla 9: Aptitudes de los individuos generados para la validación del modelo matemático.

En la validación el ECM del mejor modelo fue de 8,394,075 vehículos y la gráfica de este modelo se muestra en la figura 11.

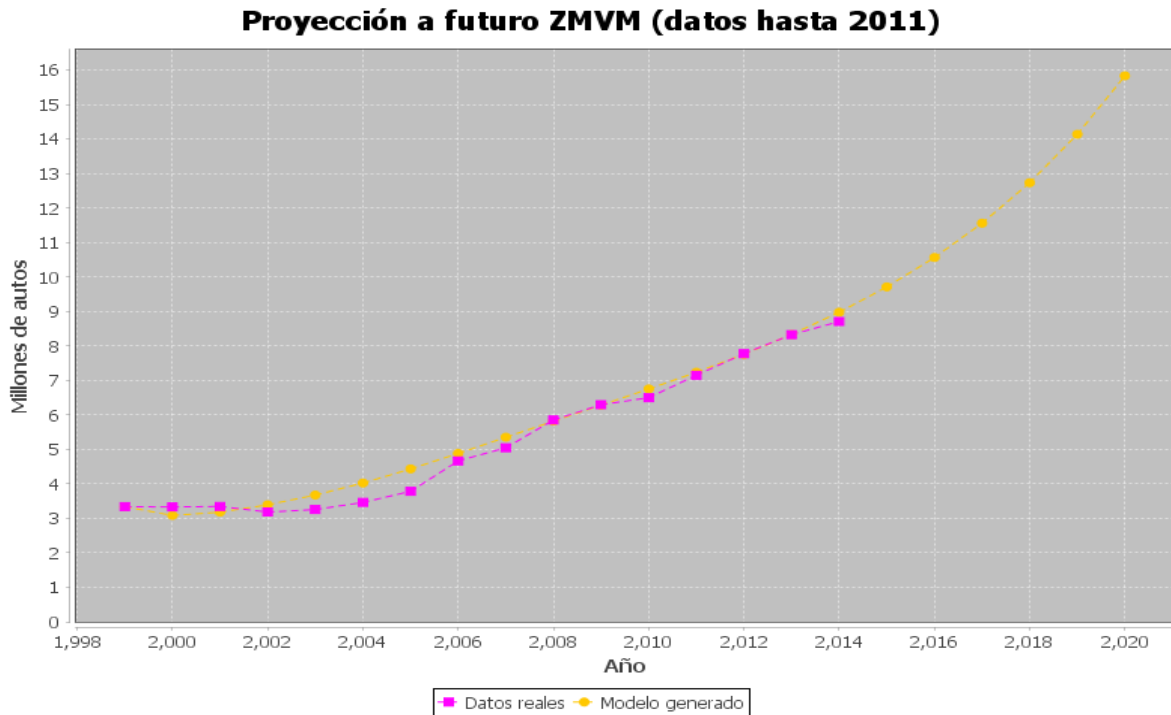


Figura 11: Mejor modelo generado PG empleando datos históricos hasta el año 2011. En color morado los datos reales y en color amarillo, los datos arrojados por el modelo matemático.

Podemos notar que el ECM del modelo que se generó en la validación (8,394,075) es mayor al del modelo que se generó empleando la totalidad de los datos reales disponibles (1,066,461), pero como ya se mencionó con anterioridad, esto es completamente normal, ya que entre más información se tenga de la función objetivo, los individuos generados tendrán una mejor aptitud. La expresión matemática correspondiente a este modelo es la siguiente.

$$F(X) = (Tan(e^{Tan(Cos(Cos(e^{Cos(Cos(Log(x))}))))})) - (Sin(Cos(Tan(Sin(Tan(e^{Cos(Cos(e^{Cos(Cos(Log(x))}))))})))))$$

Los resultados de la evaluación del modelo se muestran en la tabla 9.

Modelado de tráfico vehicular

Año	Datos reales	Modelo
1999	3,340,492	3,338,947
2000	3,331,027	3,085,895
2001	3,345,957	3,179,326
2002	3,184,166	3,393,081
2003	3,255,373	3,676,064
2004	3,462,640	4,024,284
2005	3,782,822	4,433,024
2006	4,668,502	4,884,139
2007	5,047,400	5,352,472
2008	5,854,298	5,820,287
2009	6,297,044	6,284,228
2010	6,500,465	6,752,419
2011	7,151,484	7,238,913

Modelado de tráfico vehicular

2012	<u>7,779,345</u>	7,760,146
2013	<u>8,327,810</u>	8,333,678
2014	<u>8,710,806</u>	8,978,229
2015	-	9,714,410
2016	-	10,565,951
2017	-	11,561,518
2018	-	12,737,318
2019	-	14,140,933
2020	-	15,837,157

Tabla 10: Datos históricos del parque vehicular de la ZMVM y resultado del modelo de la ZMVM empleando datos históricos hasta el año 2011. En cursiva y subrayado se muestran los datos históricos que no se emplearon en la validación, pero que si se tienen disponibles.

3.5 Detalles de la implementación

El código necesario para crear el algoritmo que implementa la programación genética y que genera los modelos descritos en esta tesis fue escrito en java bajo la versión estándar 1.8.0_101 tanto de JDK¹⁷ cómo de JRE¹⁸. Y se usó NetBeans 8.1 como IDE¹⁹.

¹⁷ Java Development Kit: entorno de desarrollo de java.

¹⁸ Java Runtime Environment: entorno de ejecución de java.

¹⁹ Integrated Development Environment: ambiente de desarrollo integrado.

3.5.1 Hardware

Los experimentos realizados para obtener los modelos descritos en esta tesis fueron llevados a cabo en una computadora con las siguientes características:

- Marca: Asus
- Modelo: GL552VW
- Procesador: Intel core I7-6700HQ 2.60 GHz 2.59Ghz
- Memoria Ram: 16 GB.
- Sistema Operativo: Windows 10 de 64 bits.

3.5.2 Arquitectura

Este proyecto fue diseñado empleando el paradigma orientado a objetos²⁰ y bajo la arquitectura MVC²¹. A continuación, en la figura 12 se muestra un diagrama de clases que enumera a detalle las clases creadas y la interacción que entre las mismas.

²⁰ Paradigma de programación, en el cual los algoritmos necesarios para obtener un resultado en particular son ejecutados por objetos que pertenecen a clases específicas. Esta forma de programar ayuda a descomponer un problema complejo en elementos de programación más pequeños y modulares.

²¹ Abreviatura para la arquitectura Modelo Vista Controlador. En esta arquitectura existen 3 capas que interactúan entre sí, el modelo, la vista y el controlador. El modelo se comunica con el controlador y viceversa, y el controlador con la vista y viceversa. No está permitida la interacción de otra forma.

Modelado de tráfico vehicular

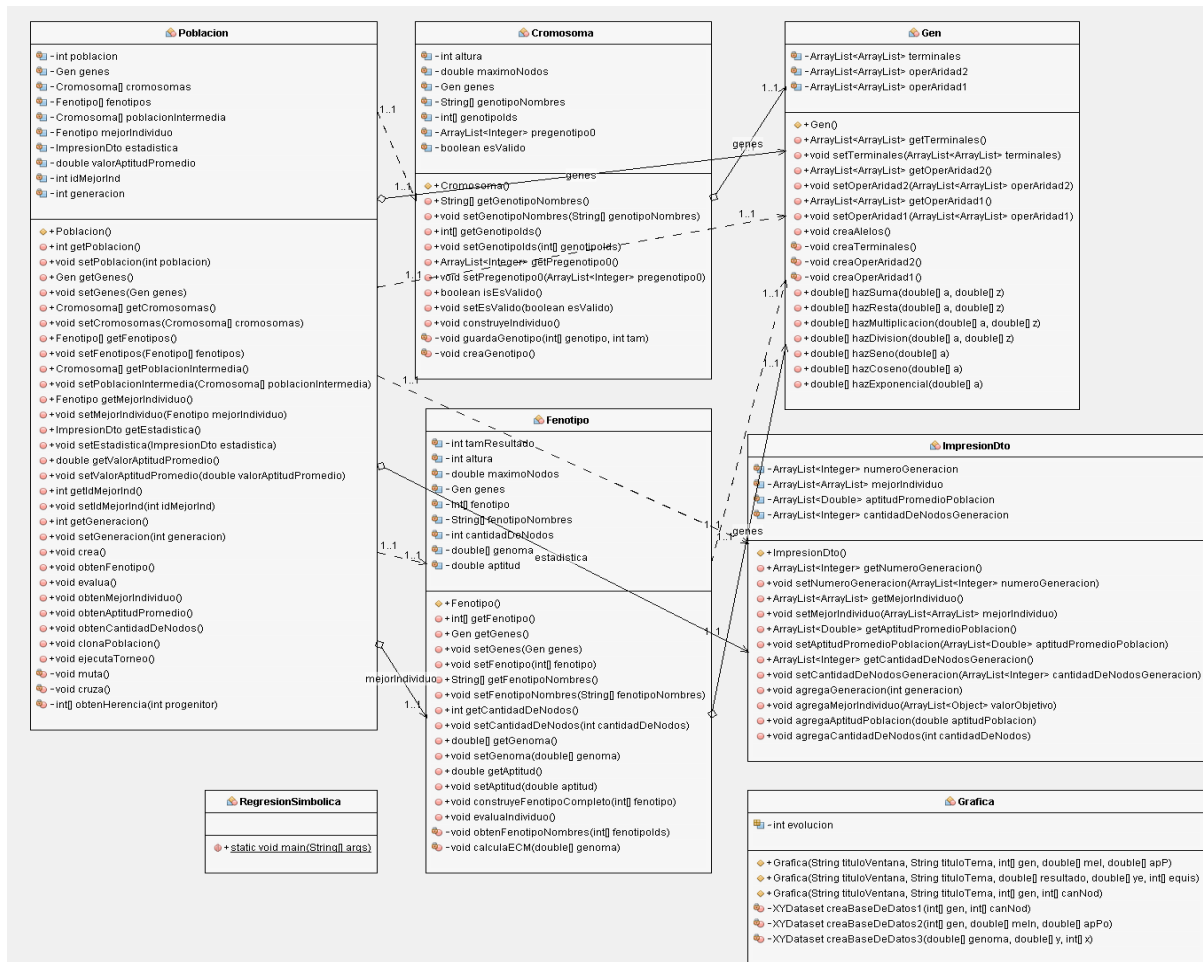


Figura 12: Diagrama de clases

- **Población:** es la clase encargada de agrupar todos los elementos necesarios para llevar a cabo la PG. En esta clase se realizan las operaciones para crear la población inicial, la cruce de individuos y selección. Es esta misma clase la que determina cuál es el individuo más apto a lo largo de las generaciones. En términos del MVC, esta clase es el controlador.
- **Cromosoma:** en esta clase se crean los individuos. Para poder construir la población inicial, la clase Población crea instancias de la clase Cromosoma, las cuales se encargan de crear los individuos. Esta clase se encarga de validar que los individuos sean sintácticamente correctos y validar su tamaño.
- **Gen:** es en esta clase donde se definen los alelos²² con los que están contruidos los individuos. Es decir que en esta clase se define el diccionario de operadores y terminales.

²² Un alelo es cada una de las formas alternativas que puede tener un mismo gen. En el contexto de este trabajo, un alelo es una operación del diccionario de operadores.

- **Fenotipo:** esta clase se encarga de obtener el fenotipo²³ asociado a un individuo y se obtiene a través de los individuos creados en la clase cromosoma. se encarga de contener la primera generación de individuos y posteriormente contendrá a los individuos de poblaciones posteriores. Contiene el ADN de cada individuo.
- **RegresionSimbolica:** es la clase principal, en esta clase se definen los parámetros con los cuales se se llevará a cabo la PG. Parámetros como, número de generaciones o función objetivo, se definen en esta clase y ella se encarga de invocar a las clases necesarias.
- **ImpresionDto:** se encarga de obtener la información de la población para que se puedan construir las gráficas.
- **Gráfica:** esta clase se encarga de organizar los resultados que se obtuvieron de la PG y los grafica. Funciona mostrando los resultados que obtiene de la clase ImpresionDato.

3.6 Resumen

En este capítulo se mostró el procedimiento completo que se llevó a cabo para poder generar el modelo matemático que describe el crecimiento del parque vehicular en la ZMVM. De igual manera se describió el procedimiento que se siguió para poder validar los modelos obtenidos. Finalmente se dieron los detalles de la implementación del código, así como las especificaciones del hardware y software que se empleó.

²³ En biología, el fenotipo es la interpretación de la cadena de ADN en términos de las características del individuo. Dado que se empleó codificación entera el genotipo y el fenotipo para este caso en particular es lo mismo. Es decir que cada alelo tiene asociado un id y dado que se empleó codificación entera, cuando se interpreta ese alelo no es necesario decodificarlo.

Capítulo 4

RESULTADOS

Es importante notar que el problema del tráfico vehicular que tiene la ciudad de México y en general la ZMVM también lo comparten otras ciudades y esto se debe en gran medida a múltiples factores como lo son:

- Población.
- Economía.
- Distribución geográfica.
- Políticas Gubernamentales.

En este capítulo analizaremos las condiciones de ciudades que padecen el mismo problema respecto al tráfico vehicular. Posteriormente se realizará una comparativa entre estas ciudades y la ZMVM.

La decisión de considerar estas ciudades fue debido a que la información del parque vehicular de estas ciudades estaba disponible y además estas ciudades son ciudades muy parecidas a la ZMVM en el sentido de que son ciudades en las cuales los habitantes desarrollan actividades laborales similares a las de la ZMVM.

Las ciudades que se tomaron en cuenta fueron:

- ZMVM.
- Londres.
- L.A.
- Beijing.
- NYC
- Río de Janeiro

4.1 Población

Es bien conocido que el comportamiento en el parque vehicular está estrechamente ligado a la cantidad de habitantes, así que en esta sección se mostrará la cantidad de habitantes desde el año 2010 y hasta el año 2014 por cada una de estas ciudades y se contrastará contra el parque vehicular.

Se eligieron estas fechas debido a que es el mismo rango de años para los cuales se tiene información de la ZMVM y la idea es hacer una comparativa. Los datos de la población se extrajeron de CITY POPULATION [24].

Modelado de tráfico vehicular

Ciudades	Habitantes 2000	Habitantes 2010	Habitantes 2014
ZMVM	18,396,677	20,116,842	20,892,724
Beijing	13,569,194	19,612,368	21,516,000
NYC	18,945,888	19,566,387	20,182,305
Londres	11,867,450	13,222,292	13,744,692
L.A.	12,365,597	12,828,966	13,340,068
Río de Janeiro (zona metropolitana)	5,851,914	6,320,446	6,498,840

Tabla 11: Población por años (2000, 2010, 2014) de ciudades analizadas.

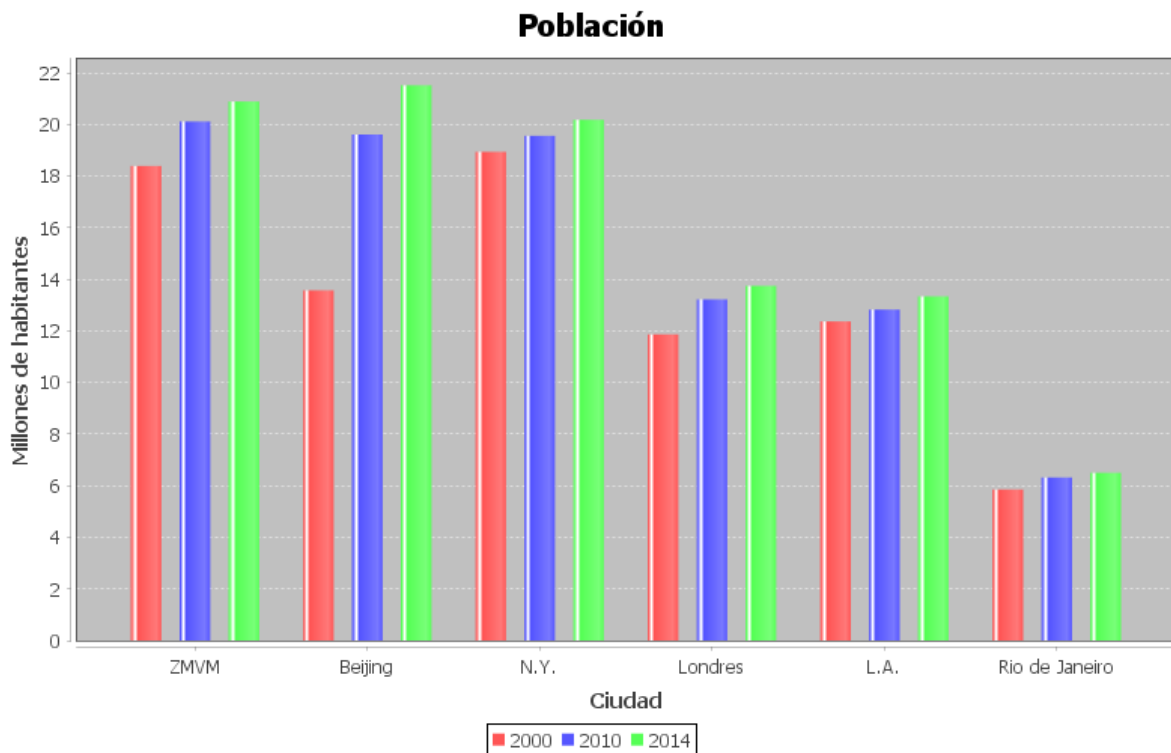


Figura 13: Comparativa de población por años, de las ciudades analizadas.

En la gráfica de la figura 13, en el eje y se muestra la cantidad de millones de habitantes de cada ciudad por años y en el eje x, las ciudades y los años (2000, 2010, 2014), estos datos se obtuvieron de la tabla 10. Se puede apreciar que todas las ciudades aumentaron su población, sin embargo también es notable que la ciudad que incrementó de manera más dramática su población fue la ciudad de Beijing. Aunque por otro lado la ZMVM es de las zonas con mayor cantidad de habitantes incluso por arriba de Nueva York.

4.2 Parque Vehicular

Ya que se mostró en la tabla 10 y figura 11 la población para cada una de estas ciudades, ahora se presentará en la tabla 11 el parque vehicular de cada ciudad analizada para mostrar una evidente relación entre el crecimiento en la población y el parque vehicular.

Ciudades	Vehículos 2000	Vehículos 2010	Vehículos 2014
ZMVM	3,331,027	6,500,465	8,710,806
Beijing	951,400	3,681,100	5,171,100
NYC	n/a	1,962,231	2,057,433
Londres	2,546,642	2,691,044	2,709,825
L.A.	6,155,565	7,067,374	7,423,446
Río de Janeiro (zona metropolitana)	n/a	2,652,961	2,835,807

Tabla 12: Población de las ciudades analizadas.

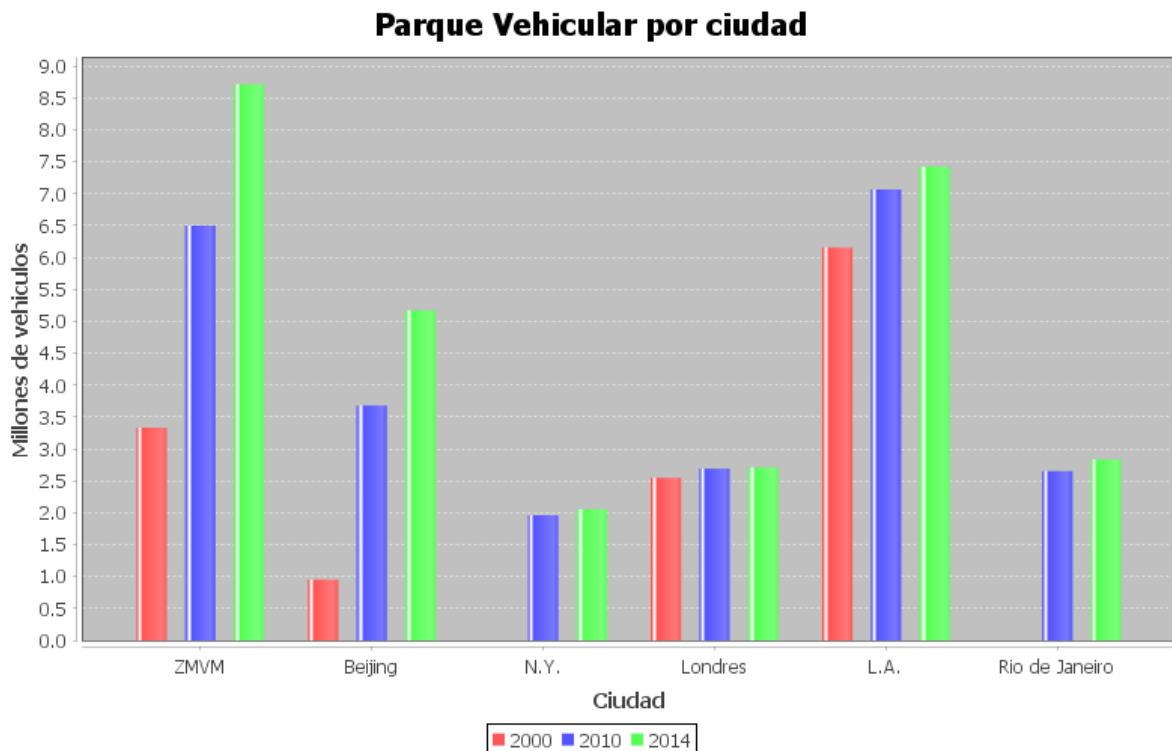


Figura 14: Comparativa de parque vehicular por años, de las ciudades analizadas.

Modelado de tráfico vehicular

En la gráfica de la figura 14, en el eje y se muestra la cantidad de millones de vehículos de cada ciudad por años y en el eje x, las ciudades y los años (2000, 2010, 2014), estos datos se obtuvieron de la tabla 11. Para las ciudades de Nueva York y Río de Janeiro no se disponía de datos del parque vehicular para el año 2000, sin embargo aun así se puede apreciar la relación que existe entre el crecimiento de la población y el crecimiento del parque vehicular.

Con ambas tablas es fácil notar que el parque vehicular crece a la par que la cantidad de habitantes por cada ciudad. Sin embargo, la población no es el único factor que influye en el crecimiento del parque vehicular.

4.2.1 Beijing

El caso de la ciudad de Beijing es uno de los más notables, ya que históricamente China y en particular la ciudad de Beijing era conocido por la cantidad de personas que deciden emplear la bicicleta como medio de transporte. Pero en la última década China ha sido de los países que han mostrado una economía en ascenso incluso con miras a ser la primera potencia mundial por arriba de E.U.A.. Muestra de este crecimiento se puede notar en 2 aspectos principales:

- Múltiples empresas internacionales han instalado sus fábricas en este país.
- Las empresas locales han decidido exportar sus productos.

Como consecuencia China se ha vuelto el primer mercado automotriz[25] en el mundo lo cual es muy bueno para su economía. Esto ha influido en gran medida en las políticas gubernamentales favoreciendo la creación de infraestructura vial. El lado negativo de este crecimiento económico es que en las principales ciudades de China (cómo Beijing), la gente ha optado por abandonar la bicicleta y adquirir un vehículo como medio de transporte, el resultado es que la cantidad de vehículos en las calles se ha incrementado considerablemente como se puede apreciar en la gráfica del parque vehicular. Finalmente esto se traduce en embotellamientos viales[26], mayores tiempos de traslados y ciudades con una mala calidad del aire.

En conclusión, para el caso de la ciudad de Beijing, la creciente economía ha sido el factor de crecimiento que ha determinado el comportamiento del parque vehicular.

4.2.2 Los Angeles

Los Angeles es otra ciudad en la cual el parque vehicular ha crecido durante muchos años y de manera constante. De la gráfica de las poblaciones se puede notar que a pesar de no ser de las ciudades más densamente pobladas (en este análisis),

el parque vehicular que posee la ciudad de L.A. es de los más numerosos superando incluso a Beijing. En este artículo [27] se menciona que la ciudad de L.A. primera en la lista de ciudades en la que las personas pasan más tiempo atascadas en el tráfico.

La historia de esta ciudad muestra que inicialmente las vialidades estaban destinadas a los tranvías, pero que con el paso del tiempo los gobernantes de esta ciudad decidieron cambiar la forma de transportarse, es decir que decidieron reemplazar las vías por las cuales transitaban los tranvías por vialidades destinadas a los taxis y no sólo eso, sino que en este artículo [28] se menciona que empresas automotrices como General Motors, influyeron para que el vehículo fuera el método de transporte favorecido.

Lo irónico en el caso de Los Angeles (y posiblemente en todos los casos), es que estudios realizados [29] muestran que el construir más vialidades destinadas para vehículos sólo generaría más congestión.

4.2.3 Londres

Londres es una de las pocas ciudades que a pesar de que su población se ha incrementado, el parque vehicular se ha mantenido con un crecimiento muy bajo y esto es debido principalmente a que las autoridades que rigen esta nación han logrado integrar políticas de gobierno que favorecen el uso de transporte público.

4.2.4 Nueva York y Río de Janeiro

Estas 2 ciudades comparten con la ZMVM la característica de que gran parte de las personas que trabajan estas ciudades se desplazan desde los suburbios o periferias de la ciudad. Por otro lado, estas ciudades también padecen el problema del tráfico vehicular.

Para la ciudad de Nueva York el factor principal que afecta el tráfico vehicular es su ubicación geográfica y la distribución de las vialidades como se menciona en [30]. A diferencia de otras ciudades, Nueva York no tiene más espacio libre en el cual pueda construir más infraestructura vial, sin embargo aun así la demanda de vehículos continúa creciendo. Una posible solución específica para esta ciudad se propone en [30], la cual consiste en imponer una especie de impuesto por las dimensiones que ocupe el vehículo en cuestión, ya que como se menciona en [30] la cantidad de espacio disponible es finita.

El problema del tráfico vehicular con la ciudad de Río de Janeiro radica básicamente en el mismo problema que poseen gran parte de las ciudades de América Latina, la mala planeación del crecimiento urbano, por un lado. Por otro lado

los recientes eventos deportivos que se han llevado a cabo en Brasil demandaron la construcción o remodelación de instalaciones deportivas, lo que trajo como consecuencia saturación en las vialidades como se menciona en [31].

4.3 Proyecciones para cada ciudad

Por cada una de estas ciudades se generó un modelo matemático para predecir cómo crecerá su parque vehicular y estos fueron los resultados.

4.3.1 Beijing

Para la ciudad de Beijing el modelo que mostró tener la mejor aptitud fue el que se muestra en la figura 15 con una aptitud (ECM) de 694,617 vehículos.

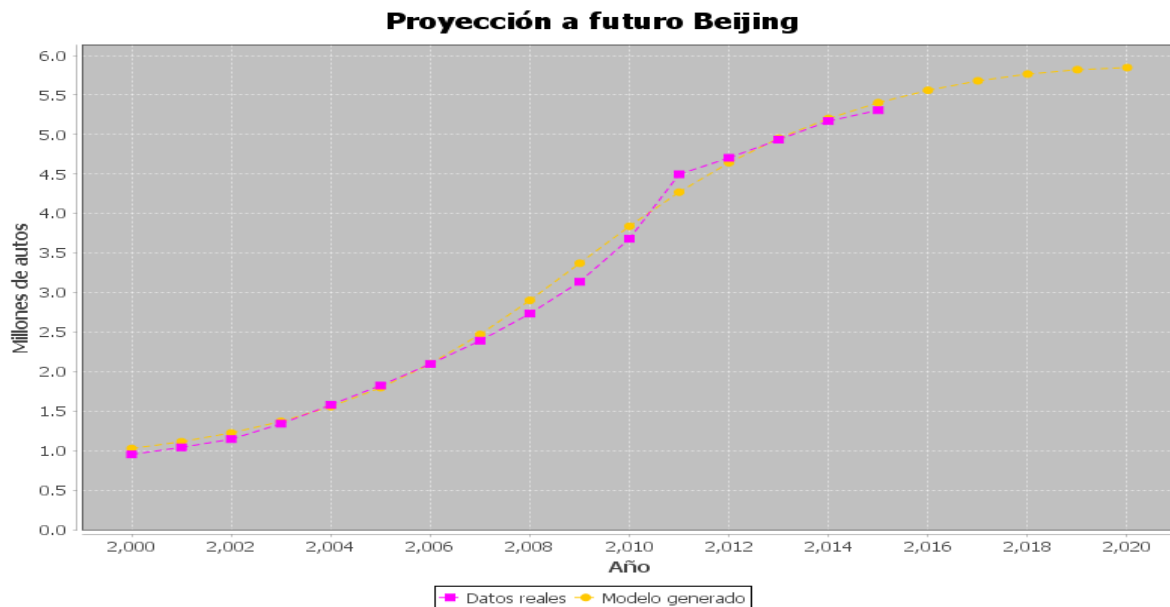


Figura 15: Modelo para la ciudad de Beijing, en el eje x los años y en el eje y el número de vehículos. En color morador los datos históricos del parque vehicular y en color amarillo, los valores arrojados por el modelo para esta ciudad.

La expresión matemática asociada a este modelo es la siguiente.

$$F(X) = \text{Senh}(\text{Sen}(\text{Sen}(\text{Senh}(\text{Tan}(\text{Log}(x)))))) + \text{Cosh}(\text{Tan}(\text{Tan}(\text{Tan}(\text{Log}(x))))$$

De igual manera que con la ZMVM para validar el buen funcionamiento del algoritmo para generar modelos, se generaron 10 modelos de la ciudad de Beijing con datos hasta el año 2010 y se generó el modelo promedio. La gráfica se muestra en la figura 15.

Modelado de tráfico vehicular

Modelo Promedio Beijing

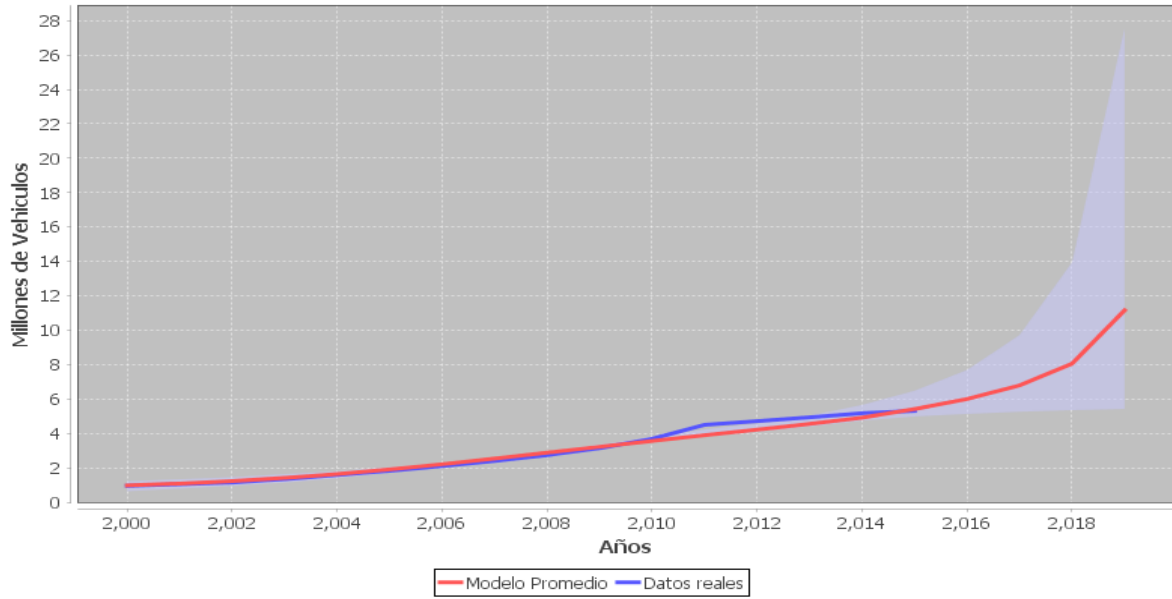


Figura 16: Gráfica del modelo promedio de la ciudad de Beijing. En color azul los datos reales del parque vehicular, en color rojo los valores promedio de los modelos generados para esta ciudad. La sombra morada denota los valores obtenidos por los modelos restantes.

4.3.2 Nueva York

Para la ciudad de Nueva York el modelo que mostró tener la mejor aptitud fue el que se muestra en la siguiente gráfica con una aptitud de 35,489 vehículos.

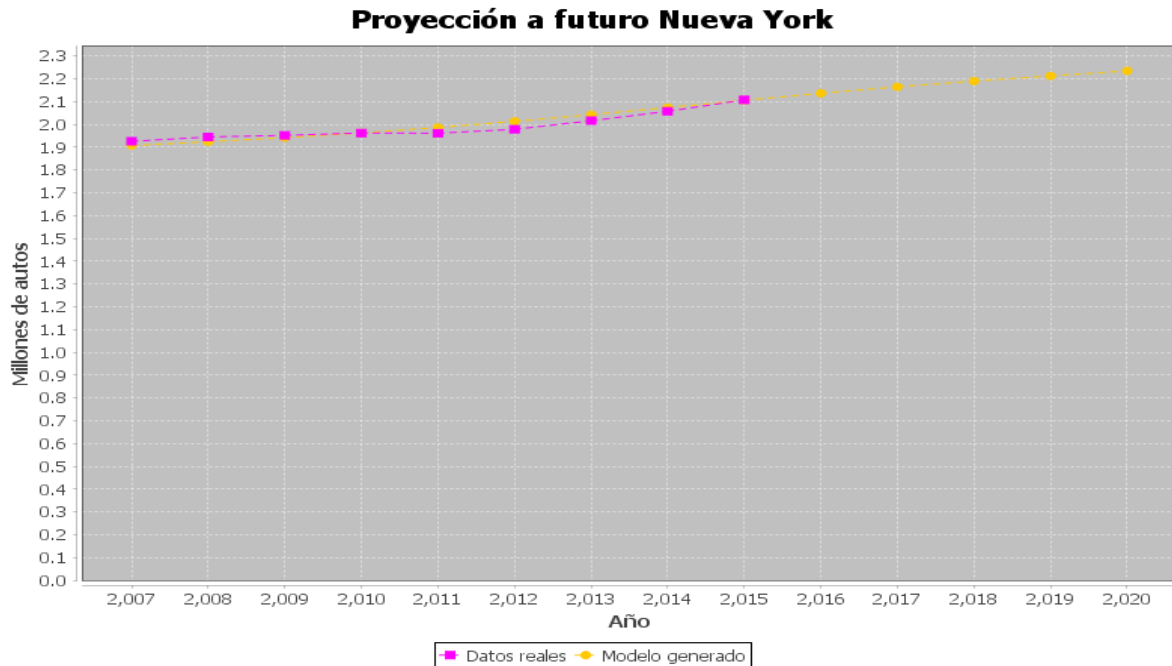


Figura 17: Mejor modelo para la ciudad de Nueva York. Se muestra la cantidad de vehículos por año, en color morado los datos históricos para el parque vehicular de esta ciudad y en amarillo, los datos arrojados por el mejor modelo generado para esta ciudad.

Modelado de tráfico vehicular

Para el modelo de la ciudad de Nueva York el modelo matemático asociado toma la siguiente expresión.

$$F(X) = \text{Tanh}(\text{Tan}(\text{Cosh}(\text{Sen}(\text{Cos}(\text{Tan}(\text{Tan}(\text{Tan}(\text{Log}(x)))))))) + \text{Tan}(\text{Tan}(\text{Log}(x)))$$

4.3.3 Londres

Para la ciudad de Londres el modelo que mostró tener la mejor aptitud fue el que se muestra en la siguiente gráfica con una aptitud de 29,946 vehículos.

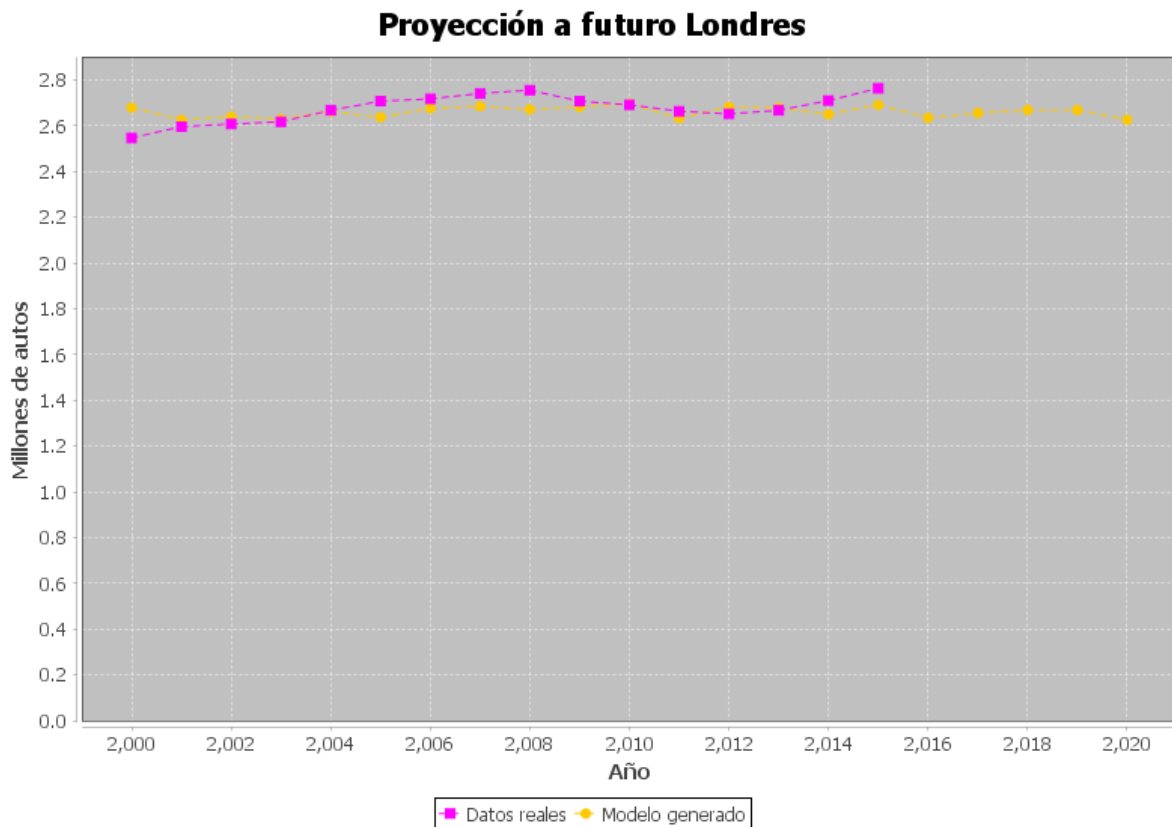


Figura 18: Mejor modelo para la ciudad de Londres. En color morado los datos históricos para el parque vehicular de esta ciudad y en amarillo, los datos arrojados por el mejor modelo generado para esta ciudad.

Para el modelo de la ciudad de Londres el modelo matemático asociado toma la siguiente expresión.

$$F(X) = e^{\text{Cos}(\text{Cos}(\text{Cos}(\text{Sen}(x*x))))} + \text{Cos}(\text{Sen}(\text{Cos}(\text{Cos}(x + 1)))) + \text{Cos}(\text{Cos}(x * x))$$

4.3.4 Los Angeles

Para la ciudad de Los Angeles el modelo que mostró tener la mejor aptitud fue el que se muestra en la siguiente gráfica con una aptitud de 218,608 vehículos.

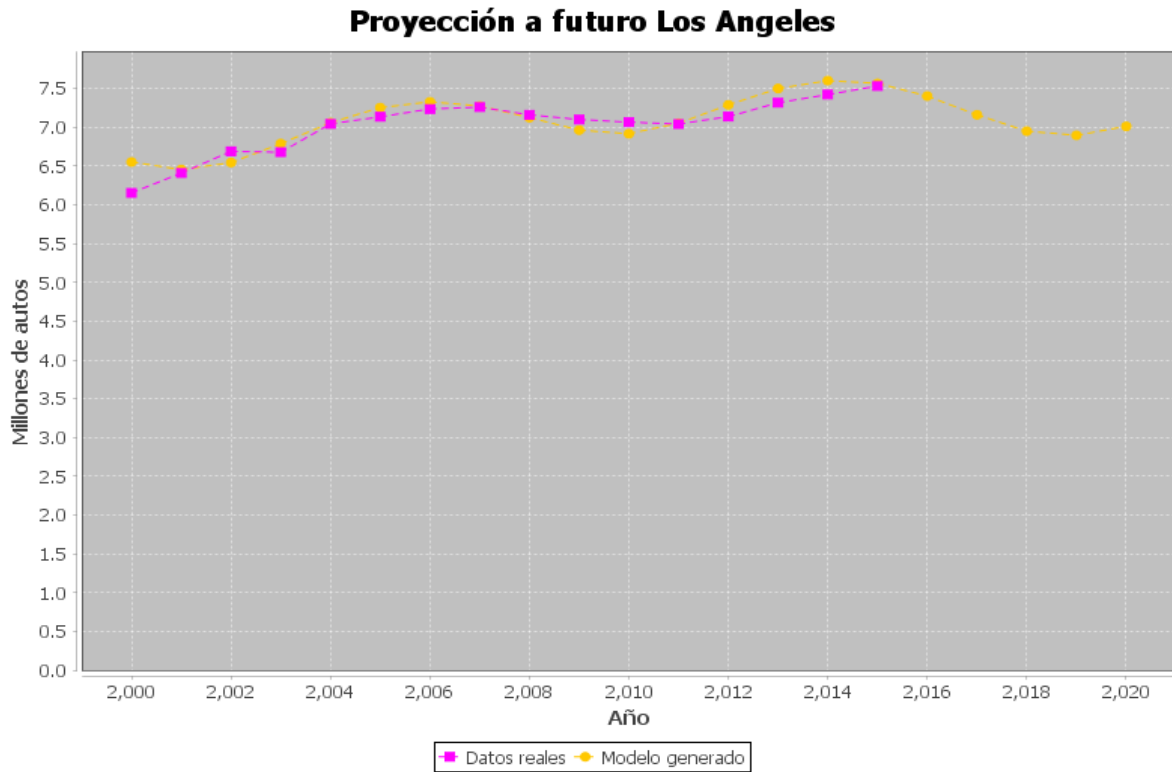


Figura 19: Mejor modelo para la ciudad de los Angeles. En color morado los datos históricos para el parque vehicular de esta ciudad y en amarillo, los datos arrojados por el mejor modelo generado para esta ciudad.

Para el modelo de la ciudad de Los Angeles el modelo matemático asociado toma la siguiente expresión.

$$F(X) = \text{Log}(\text{((Cos(Cos((Cos(Sen(Log(x)))) * x))) * (Cos(Cos((Cos(Sen(e^{(e^{(Cos(Sen(Log(x))))}))})) * (x)))) * (x)))$$

4.3.5 Río de Janeiro

Para la ciudad de Río de Janeiro el modelo que mostró tener la mejor aptitud fue el que se muestra en la figura 20 con una aptitud de 182,771 vehículos.

Modelado de tráfico vehicular

Proyección a futuro Rio de Janeiro

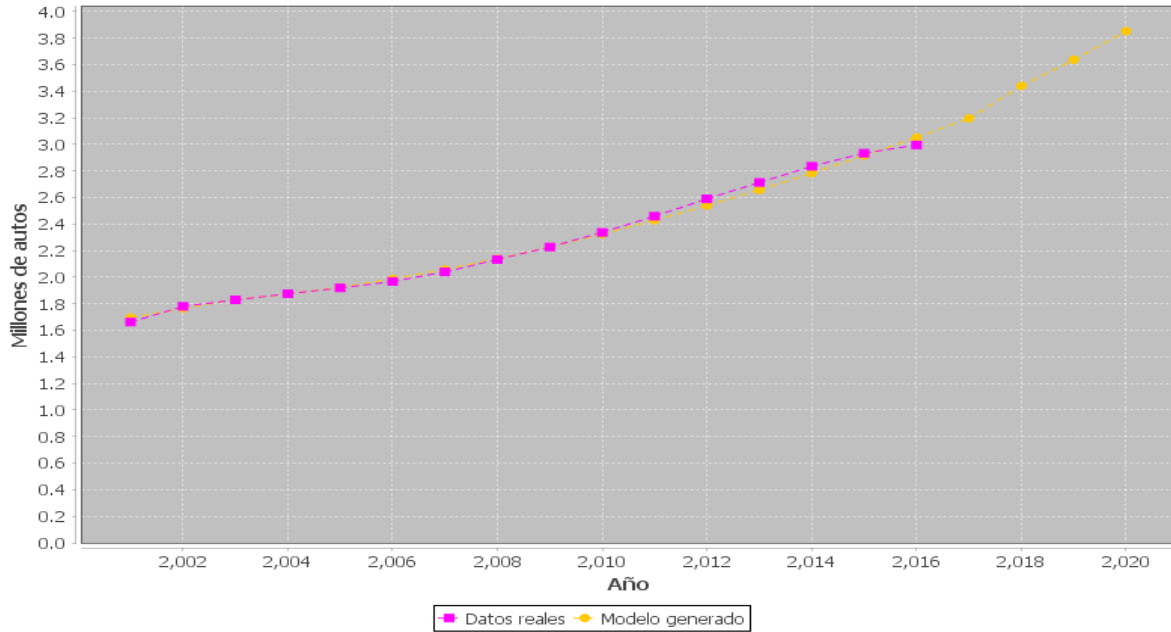


Figura 20: Mejor modelo para la ciudad de Río de Janeiro. En color morado los datos históricos para el parque vehicular de esta ciudad y en amarillo, los datos arrojados por el mejor modelo generado para esta ciudad.

También para la ciudad de Río de Janeiro se generó una gráfica promedio considerando datos hasta el año 2011, que se muestra en la figura 21.

Modelo Promedio Rio de Janeiro

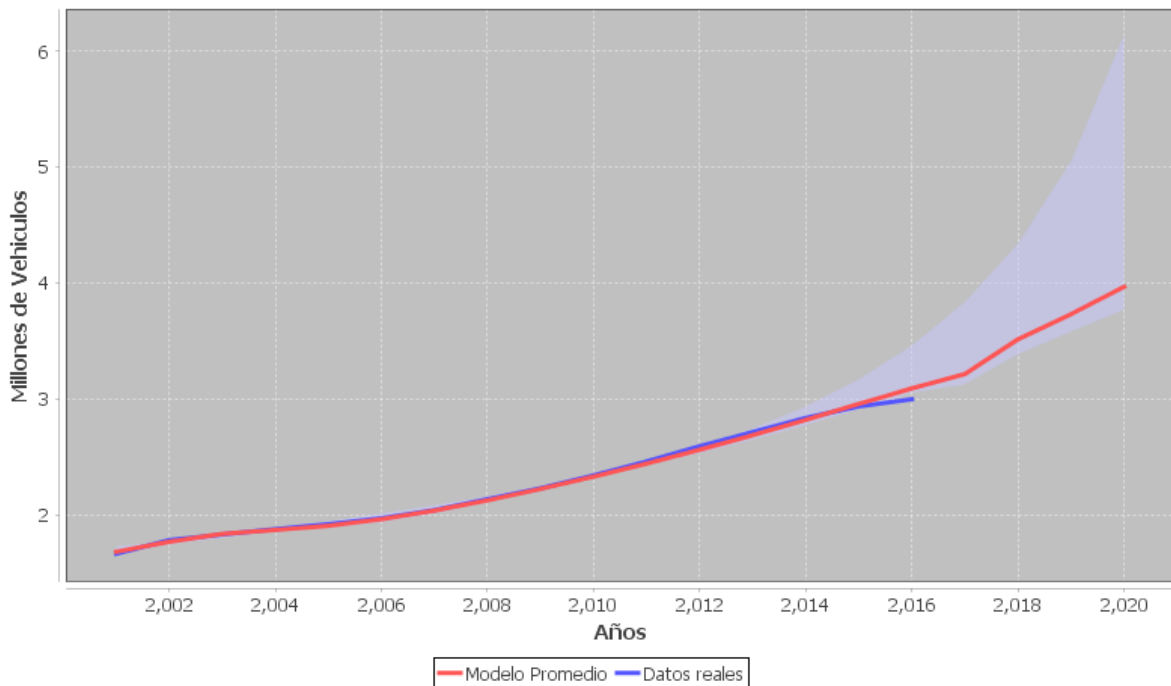


Figura 21: Grafica del modelo promedio de la ciudad de Río de Janeiro. En color azul los datos reales del parque vehicular, en color rojo los valores promedio de los modelos generados para esta ciudad. La sombra morada denota los valores obtenidos por los modelos restantes.

El modelo matemático asociado a la zona metropolitana de Río de Janeiro es el que se muestra a continuación.

$$F(X) = \frac{\text{Tan}(\text{Tan}(\text{Tan}(\text{Log}(x))))}{\text{Cos}(\text{Cos}(\text{Cos}(\text{Cos}(\text{Cos}(\text{Cos}(\text{Cos}(\text{Cos}(\text{Sen}(\text{Sen}(\text{Tan}(\text{Cos}(\text{Cos}(\text{Tan}(\text{Tan}(\text{Tan}(\text{Cos}(\text{Cos}(\text{Tan}(\text{Tan}(\text{Log}(x))))))))))))))))))$$

4.3.6 Delegación Cuauhtémoc

Para la delegación Cuauhtémoc se generó el modelo que posteriormente será empleado para la simulación en la sección 4.4. La grafica asociada a este modelo se muestra en la figura 21.

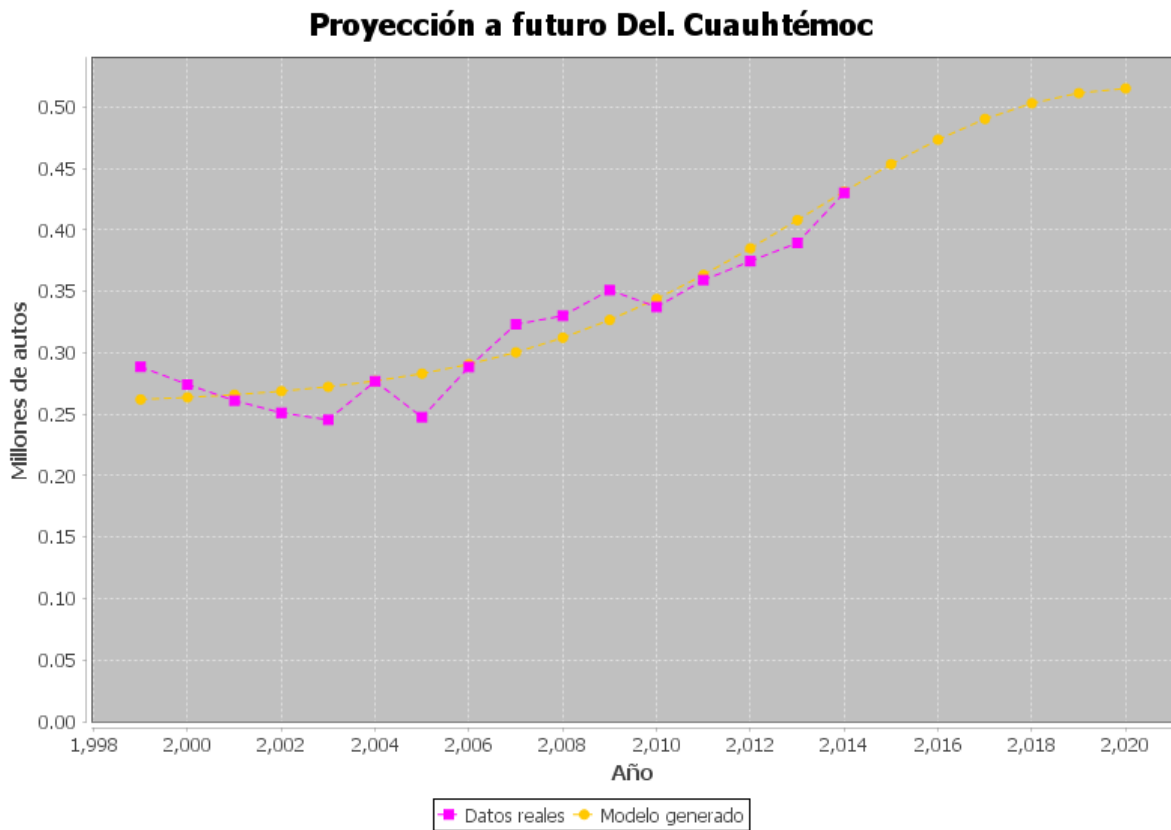


Figura 22: Mejor modelo para la delegación Cuauhtémoc. En color morado los datos históricos para el parque vehicular de esta delegación y en amarillo, los datos arrojados por el mejor modelo generado para esta delegación.

La expresión matemática asociada a este modelo es la siguiente.

$$F(X) = \frac{\text{Senh}(\text{Cos}(\text{Sen}(e^{\text{Cos}(\text{Sen}(e^{\text{Cos}(\text{Cos}(\text{Tan}(\text{Log}(x))))}))})))))}{e^{\text{Cos}(\text{Sen}(\text{Tan}(\text{Log}(x))))}}$$

4.4 Simulación

Por otra parte, para tener una idea clara de las implicaciones que tienen los datos arrojados por los modelos construidos en el capítulo 4, en particular el modelo de la delegación Cuauhtémoc, fue necesario realizar una simulación.

Esta simulación cumple con el objetivo de emplear los datos arrojados por el modelo de la delegación Cuauhtémoc como datos de entrada para hacer tangible lo que representa la proyección en el parque vehicular y como afectara esto a la ciudadanía. En esta simulación se considerarán mediciones reales de las densidades²⁴ en algunas zonas representativas de la ZMVM, como lo es la delegación Cuauhtémoc y se comparan con los datos arrojados por el modelo matemático. En lo siguiente se detallan los elementos necesarios para realizar la simulación.

4.4.1 Conceptos importantes

Dentro de la teoría del análisis del tráfico vehicular (ya sea desde el punto de vista microscópico o macroscópico) existen 4 conceptos básicos, que son:

- Velocidad.
- Tiempo de viaje.
- Densidad.
- Flujo.

4.4.1.1 Velocidad

La velocidad está definida como la razón de movimiento en distancia por unidad de tiempo y normalmente se expresa en kilómetros por hora (km/h). El manual de capacidad de carreteras (Highway Capacity Manual) determina la capacidad y los niveles de servicio en condiciones ideales. Posteriormente estas variables se adaptan en función de las discrepancias que existan entre estas condiciones y las características propias del tramo real de la carretera estudiada o proyectada. En el HCM 2000 se toma la velocidad promedio de viaje como la media de velocidad, esto es debido a que es fácil de calcular dicha velocidad observando cada vehículo dentro del tránsito o en el caso de la simulación, es posible medir dicha velocidad por cada vehículo. La velocidad es la medida estadística que mayor relevancia tiene en el análisis del tráfico vehicular.

²⁴ En este contexto, densidad se refiere a la densidad vehicular.

4.4.1.2 Tiempo de Viaje

El tiempo de viaje es el tiempo que le toma a un vehículo atravesar un segmento determinado (h), ya sea una carretera vialidad o simplemente un carril. El tiempo promedio de viaje t_a es el tiempo promedio que le tomó a una cantidad determinada de vehículos atravesar el segmento h .

4.4.1.3 Velocidad Promedio de Viaje

La velocidad promedio de viaje es otro concepto importante y se calcula dividiendo el segmento de la vialidad que se está analizando entre el tiempo promedio de viaje de los vehículos que atraviesan dicho segmento. La siguiente expresión muestra las variables a considerar.

$$v = \frac{L}{t_a}$$

Donde:

- v = Velocidad promedio de viaje (km/h),
- L = Longitud del segmento de carretera (km), y
- t_a = Tiempo promedio de viaje en el segmento (h).

4.4.1.4 Densidad

La densidad (en el contexto del tráfico vehicular) es la cantidad de vehículos que ocupa una cierta longitud dada en una carretera, vialidad o carril y normalmente se expresa en vehículos por kilómetro (veh/km). La siguiente expresión muestra las variables a considerar para el cálculo de la densidad.

$$\rho = \frac{N}{d}$$

Donde:

- ρ = Densidad (veh p/km/carril),
- N = cantidad de vehículos y
- d = longitud de la vialidad (km).

La densidad es el parámetro más importante en el contexto del análisis del tráfico vehicular porque es la medida que relaciona de manera más directa el número de vehículos en una longitud dada.

4.4.1.5 Flujo

El flujo J , es la frecuencia a la cual pasan los vehículos por un punto o sección transversal de un carril o calzada. La tasa de flujo es pues, el número de vehículos, N , que pasan durante un intervalo de tiempo específico. En otras palabras, el flujo es la cantidad de vehículos que se desplazan en el sistema (en este caso la simulación) en el tiempo. El flujo es calculado mediante la expresión siguiente.

$$J = v * \rho$$

Donde:

- J = Flujo,
- v = velocidad promedio,
- ρ = densidad.

4.5 Simulador

SUMO²⁵ fue el simulador empleado para hacer tangibles los resultados obtenido del modelo del crecimiento del parque vehicular que se generó en el capítulo 4. Se tomó esta decisión debido a que es un simulador con licencia GNU V3²⁶, lo que en pocas palabras significa que el código es libre y puede ser modificado por los usuarios, además del hecho de que su documentación y manuales están disponibles de manera gratuita.

Este simulador puede ser empleado no sólo para simular tráfico vehicular, sino para simular muchas otras formas de movilidad urbana, por ejemplo, ciclovías o transporte público.

Por otra parte, SUMO está regido por un modelo de tipo macroscópico, de manera formal podemos decir que SUMO emplea el modelo tipo car following (del que se habló en el capítulo 2). Existe gran cantidad de documentación y ejemplos de cómo hacer uso de SUMO, los cuales se pueden encontrar en [34].

En las siguientes secciones mostraré detalladamente cómo es que se utilizó esta gran herramienta en este trabajo.

²⁵ Simulatio of Urban MObility

²⁶ GNU General Public License (o simplemente sus siglas del inglés GNU GPL) es la licencia más ampliamente usada en el mundo del software y garantiza a los usuarios finales (personas, organizaciones, compañías) la libertad de usar, estudiar, compartir (copiar) y modificar el software.

4.5.1 Construcción del mapa

La forma en la cual trabaja SUMO es definiendo una red que representa el mapa del terreno que se quiere en la simulación. Para ello se emplean conceptos básicos de teoría de gráficas²⁷. Los datos (vértices y aristas) que representan el mapa a simular son almacenados en archivos de tipo .xml²⁸, estos datos posteriormente son usados por SUMO para poder definir calles, avenidas, cruces, semáforos y todo elemento que forma parte de la movilidad urbana.

Existen múltiples formas para definir una red en SUMO, una de ellas es generar los elementos manualmente, es decir, generar las aristas y vértices que representan los elementos de la simulación. En el contexto de una simulación amplia, como lo es el caso de esta tesis, este trabajo sería extenso y complicado, lo cual implica que los resultados obtenidos podrían ser incorrectos. Es por eso que se eligió descargar el mapa del segmento deseado, desde el sitio *openstreet maps*. Este sitio permite a los usuarios aportar datos reales de las zonas que habitan para construir mapas que posteriormente pueden ser utilizados con múltiples propósitos, como en este caso, el de simular el crecimiento del parque vehicular.

Una vez descargado el mapa de la simulación es necesario procesarlo y convertirlo en un archivo .xml, para que SUMO lo pueda interpretar y así dar paso a la simulación. En [35] se puede encontrar un tutorial de cómo llevar a cabo este proceso.

4.6 Simulaciones

Debido a que la cantidad de vehículos para la ZMVM se miden en millones, generar una simulación en la que interactúen millones de vehículos requerirá equipo de cómputo de alto rendimiento y para poder tener una mejor idea del impacto que el crecimiento del parque vehicular conlleva a nuestra vida diaria, se generó una simulación de la delegación Cuauhtémoc con la intención de poder medir la cantidad de contaminantes que se emite en esta delegación.

4.6.1 Contaminantes (CO₂, NO_x)

El Dióxido de Carbono (CO₂) y los Óxidos de Nitrógeno (NO_x) son de los contaminantes más nocivos emitidos por la quema de combustibles empleados en los vehículos actualmente, es por esta razón que la simulación realizada en esta tesis

²⁷ La teoría de grafos es un campo de estudio de las matemáticas y las ciencias de la computación, que estudia las propiedades de los grafos estructuras que constan de dos partes, el conjunto de vértices, nodos o puntos; y el conjunto de aristas, líneas o lados que pueden ser orientados o no.

²⁸ Siglas en inglés de *eXtensible Markup Language*, traducido como "Lenguaje de Marcado Extensible" o "Lenguaje de Marcas Extensible".

Modelado de tráfico vehicular

considero estos dos compuestos para ser medidos. Por otro lado, SUMO permite realizar una medición y análisis de dichos contaminantes. SUMO le asocia a cada vehículo un modelo de medición de contaminantes, de tal manera que se puede medir la cantidad de emisiones contaminantes que produce cada vehículo. Además, SUMO permite generar una sumatoria de los contaminantes que se emitieron por cada vialidad de la simulación, es decir, que calcula el número de vehículos que transitaron por dicha vialidad, así como los contaminantes que estos generaron.

Finalmente, mediante SUMO es posible obtener un análisis detallado múltiples factores relevantes en la simulación, pero en particular se puede determinar la cantidad de contaminantes generados por la totalidad de los vehículos empleados en la simulación a lo largo de esta.

Las herramientas suministradas por SUMO que se usaron en esta simulación, así como la documentación de cómo usarlas, se puede encontrar en [36] y se listan a continuación:

- *Duarouter*: con esta herramienta es posible generar los flujos vehiculares que SUMO emplea para la simulación.
- *Netedit*: esta herramienta se emplea en caso de que la red vehicular (mapa) construida para la simulación, necesite ser editada para agregar o quitar elementos como, semáforos, cruces, calles, avenidas, etc.
- *Instrucción –emission-output*: al ejecutar la simulación, con este comando se le pide a SUMO que genere un análisis (en formato .XML) de los contaminantes emitidos por los vehículos en la simulación.
- *plot_net_dump.py*: permite realizar una gráfica de la cantidad de emisiones generadas en la simulación.
- *xml2csv.py*: haciendo uso de esta herramienta se puede traducir el análisis generado por sumo (archivo XML) a formato .csv²⁹ para un manejo más fácil.

El análisis que se genera con SUMO permite enfocarse en muchos factores, como velocidades, tiempos, densidades y en este caso en particular emisión de contaminantes. De tal manera que a pesar de que existen otros contaminantes importantes que son emitidos por la quema de combustibles, en este trabajo nos enfocaremos en estos dos compuestos.

Los resultados de la simulación de ambos compuestos SUMO los cuantifican en Toneladas\segundo, esta unidad de medida es mejor conocida en física como Flujo Másico. Dicha unidad de medida expresa la variación de la masa en el tiempo. Matemáticamente es la diferencial de la masa con respecto al tiempo. Se trata de algo frecuente en sistemas termodinámicos, pues muchos de ellos (tuberías, toberas, turbinas, compresores, difusores...) actúan sobre un fluido que lo atraviesa.

²⁹ Abreviatura de comma-separated values, son de texto cuyos valores están separados por comas y normalmente son empleados para manejo de gran cantidad de datos organizados en columnas y filas.

4.6.2 Consideraciones de la Simulación

La cantidad de vehículos que se ingresan a esta simulación son los que obtuvieron del modelo generado para dicha delegación. Dichos valores se muestran a en la tabla 13.

Año	Vehículos
2016	473,433
2017	490,171
2018	502,922
2019	511,290
2020	515,052

Tabla 13: Cantidad de vehículos por año, empleados en la simulación.

La simulación dura un total de 1000 segundos (16.6 minutos), esto significa que para cada año y por lo tanto para cada cantidad de vehículos, la simulación toma un total de 1000 segundos.

En esta parte es importante mencionar que, en la realidad, es muy poco probable, por no decir imposible, que la cantidad total de vehículos tanto del histórico del parque vehicular de la delegación Cuauhtémoc, o la proyección del modelo estén en circulación al mismo tiempo. A pesar de esto, en la simulación si es posible y esto se hace con la intención de mostrar la cantidad de emisiones que el total de los vehículos producirá.

Es por este motivo que la simulación toma este lapso de tiempo, de otra manera se tendrían que realizar simulaciones por separado de fracciones del total de los vehículos para posteriormente realizar la suma de las cantidades de emisiones. los vehículos ingresan y salen de la simulación en diferentes puntos de la delegación Cuauhtémoc. A continuación, se muestra el mapa generado para la simulación de la delegación Cuauhtémoc.

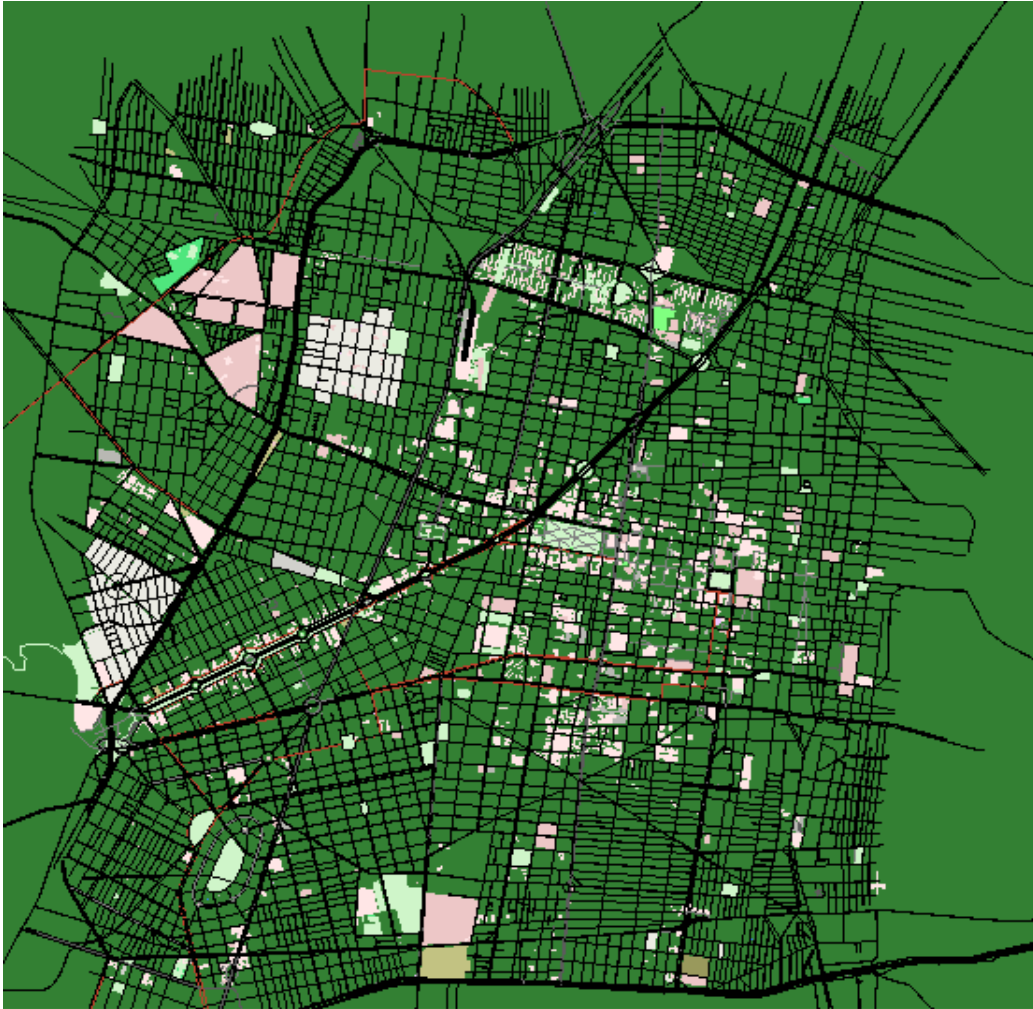


Figura 23: Mapa de la delegación Cuauhtémoc, descargado de openstreetmaps.com e interpretado con SUMO.

En la realidad se tiene una determinada cantidad de vehículos en el parque vehicular (histórico del parque vehicular), pero es evidente que no todos estos vehículos se encuentran activos, es decir, desplazándose en las vialidades. Esto se debe a que, algunos de ellos estarán estacionados o probablemente ya no estén en uso. Sin embargo, el objetivo de esta simulación es mostrar la cantidad de contaminantes que serán emitidos a la atmósfera por el total de vehículos que se proyecta para cada año, es por este motivo que la simulación si contempla la totalidad de vehículos desplazando en las vialidades de la delegación Cuauhtémoc, aunque sea en un periodo corto de tiempo.

4.6.3 Resultados de la simulación

Para cada año mostrado en la tabla 13, se llevó a cabo una simulación empleando como datos de entrada la cantidad vehículos de esta misma tabla, que además fue el resultado de evaluar el modelo de la delegación Cuauhtémoc. Una vez que se realizaron las simulaciones de todos los años, el siguiente paso fue emplear

las herramientas descritas en la sección 4.6.1. En la figura 24 se muestra las cantidades de CO₂ y NO_x emitidas por el parque vehicular para los años 2016, 2017, 2018, 2019 y 2020.

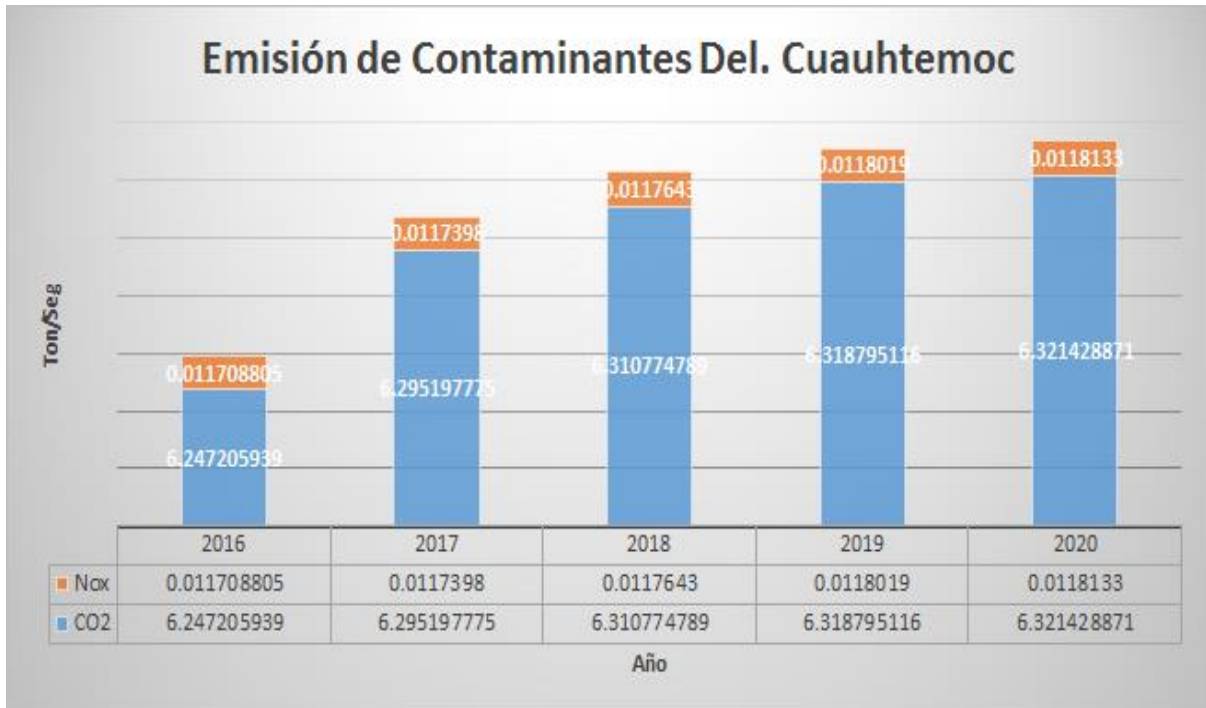


Figura 24: Proyección de las emisiones de contaminantes por año para la delegación Cuauhtémoc. En color naranja las Ton/Seg³⁰ de Nox y en azul las Ton/Seg para el CO₂.

Estos resultados muestran un alarmante aumento en la cantidad de emisiones contaminantes en la atmósfera que se traducen en una calidad de vida cada vez peor. Se ha demostrado que la calidad del aire se encuentra estrechamente relacionada con las enfermedades respiratorias. La contaminación del medio ambiente que es generado por el tráfico vehicular, aunado al problema de movilidad urbana son motivos suficientes para comenzar a tomar medidas preventivas y correctivas al respecto.

4.7 Resumen

En este capítulo se muestran tanto los resultados del modelo matemático, así como de la comparativa con las distintas ciudades consideradas en este trabajo. De igual manera se generó la simulación necesaria para hacer tangible el impacto que tendrá el crecimiento del parque vehicular.

³⁰ Flujo másico, es en física la magnitud que expresa la variación de la masa respecto al tiempo. Esta medida se emplea frecuentemente en sistemas termodinámicos (tuberías, turbinas, contaminación ambiental, etc.).

Capítulo 5

CONCLUSIONES

5.1 Ventajas de emplear PG en el contexto del parque vehicular

Como se estableció al inicio de esta tesis, una de las ventajas que se tiene al usar técnicas como la PG, es que los modelos generados no son tan restrictivos como los que se pueden generar con otras técnicas, como lo es la regresión lineal. En este trabajo se mostró que una vez que se construyó, mediante código de programación, el algoritmo asociado a la PG y una vez que fueron definidos los parámetros que la PG emplea, fue muy sencillo generar múltiples modelos para las ciudades analizadas en este trabajo.

Dichos modelos no solo sirvieron para dar una proyección del comportamiento del parque vehicular en dichas ciudades, sino que también cumplieron la tarea de comprobar que el procedimiento en general de PG funcionaba de manera correcta.

5.2 Colapso Vial

Los resultados obtenidos en este trabajo, muestran que el parque vehicular para la ZMVM continuara creciendo, de tal manera que para el año 2020 se habrán sobrepasado la cantidad de once millones de vehículos.

Si el parque vehicular continúa creciendo de la manera que lo ha hecho en años recientes, para el año 2020 se llegará al colapso vial, tal como se menciona en [32]. En este artículo el Instituto de Geografía de la UNAM llegó a los mismos resultados que se muestran en esta tesis. De acuerdo a este artículo la cantidad de vehículos en la ZMVM para el año 2020 será de once millones, empleando programación genética se llegó al resultado de que para el año 2020 se tendrán 11.45804409 millones de vehículos en la ZMVM. Estos resultados nos llevan a pensar que los tiempos de viaje, el consumo de combustibles y la cantidad de emisiones contaminantes se incrementa año con año, lo que en términos generales significa una, cada vez peor, calidad de vida.

De tal manera que además de las conclusiones antes mencionadas, este trabajo corrobora los resultados obtenidos por el Instituto de Geografía de la UNAM, lo cual indica que es necesario tomar acciones lo antes posible para frenar y tratar de disminuir el crecimiento en el parque vehicular de la ZMVM.

5.3 Contaminación Ambiental

Los resultados obtenidos por la mayoría de los modelos generados en esta tesis muestran un crecimiento en el parque vehicular, exceptuando los modelos de las ciudades de Londres y Los Angeles. Pero en particular el modelo matemático generado para la delegación Cuauhtémoc muestran que el parque vehicular en esta zona representativa de la CDMX continuara creciendo.

Los mismos resultados se vieron reflejados en la simulación de la sección 4.6.3, estos resultados indican que los índices de emisión de contaminantes continuaran aumentando con el paso de los años, lo que se traduce en una peor calidad del aire año tras año.

Las implicaciones que estos resultados tienen son muchas, aunque las más evidentes y preocupantes tienen que ver con la calidad de vida de los habitantes de esta delegación. De no encontrar una manera para disminuir el parque vehicular en la ZMVM y en particular en la delegación Cuauhtémoc, se concluye con base en este trabajo que la calidad de vida de los habitantes de esta delegación será peor con el paso de los años.

5.4 Inseguridad Vial

Es claro que el crecimiento del parque vehicular en una ciudad está estrechamente relacionado con el tráfico vehicular en la misma. En fechas recientes se ha dado un gran incremento en el índice de asaltos a automovilistas, esto se puede verificar al revisar los noticieros o incluso en las redes sociales. Estos crímenes han sido llevados a cabo en vialidades de la ZMVM que reportan un índice elevado de tráfico vehicular, lo que significa que el tráfico vehicular tiene una relación directa con el número de asaltos a automovilistas.

De continuar así, podríamos suponer que la cantidad de asaltos a automovilistas se comportan de manera muy similar al crecimiento del parque vehicular. Este es un punto muy importante que nuestras autoridades deben tomar en cuenta, ya que uno de los principales factores que decide la calidad de vida es la inseguridad.

De continuar incrementando el parque vehicular como se muestra en este trabajo el resultado inminente será un incremento en la inseguridad de los automovilistas de la ZMVM.

5.5 Propuestas

Dado que la fecha en la cual se llegará al colapso vial[32] no está muy lejana (año 2020), en este trabajo se proponen algunas soluciones que podrían ayudar a atenuar el problema del crecimiento del parque vehicular.

5.5.1 Uso de la Bicicleta

La principal propuesta que se hace en este trabajo es incentivar el uso de transportes alternativos, como es el uso de la bicicleta. Este medio de transporte ha probado ser una excelente alternativa por sus múltiples beneficios tanto en el aspecto de disminuir el tráfico vehicular, como en la salud de los usuarios. En ciudades como Alemania, Estocolmo y Dinamarca el uso de la bicicleta va en aumento y como resultado se tiene la disminución en el tráfico vehicular. En [37] se menciona que para la ciudad de México la bicicleta es el medio de transporte más rápido en trayectos cortos. Además, en [38] se mencionan algunos beneficios para la salud, como los son:

- Reducción en los niveles de colesterol en la sangre.
- Ayuda a mejorar la coordinación motriz.
- Reduce los niveles de estrés y mejora el estado de ánimo.

5.5.2 Compartir Vehículos

En numerosos países de Europa, otra solución que se ha encontrado para el problema del tráfico, ha sido el que los automovilistas compartan los vehículos existentes. Existen numerosas aplicaciones como: *blablacar* o *compartoche*, que ayudan a los usuarios a publicar las rutas que seguirán y así es posible compartir los gastos con otros usuarios que compartan la misma ruta. En la U.N.A.M. se están desarrollando interesantes propuestas como PumaRide[39]. Por otra parte, el compartir el vehículo tiene múltiples ventajas, no solo económicas, si no también ambientales y sociales, en [40] se mencionan algunas de ellas, además de otras modalidades de cómo se puede compartir el medio de transporte.

La dinámica de compartir el vehículo en sus múltiples formas es relativamente una idea reciente sin embargo es una de las ideas que, apoyándose en elementos tecnológicos como las aplicaciones móviles, ha tenido un fuerte impulso. Esta es una de las alternativas para la disminución en el crecimiento vehicular más viables hasta el momento.

5.5.3 Gravámenes a la compra de vehículos

Otra opción es aplicar un impuesto extra a la compra de vehículos motorizados, de tal manera que lo recabado con este impuesto se pueda emplear para solucionar el problema existente. Aunque se podría pensar que la solución es crear más vialidades, se ha mostrado [29] que el crear nuevas vialidades sólo empeora el problema del tráfico vehicular. Por otro lado, la solución está en emplear los recursos recaudados en crear más ciclovías o extender los programas ya existentes a más zonas de la ZMVM. Por otro lado invertir en transporte público es una prioridad. En trabajos futuros se podría dar una estimación de lo que se recaudaría con estos gravámenes y lo que se podría hacer con estos, en caso de que sean bien empleados.

5.5.4 Distribución de horarios laborales y escolares por zonas

Una propuesta interesante para las autoridades sería la de asignar horarios laborales y escolares por zonas, de tal manera que el fenómeno de la “hora pico” se pudiera disminuir. La propuesta en general sería asignar horarios de entrada y de salida con un lapso de 30 o 60 minutos de diferencia para las distintas demarcaciones de la ZMVM. Así las vialidades podrían disminuir su flujo vehicular ya que se distribuiría de una mejor manera los vehículos que en estas se desplazan.

5.6 Trabajo a futuro

En este trabajo se muestran el proceso que se llevó a cabo para obtener los resultados que se analizan a detalle sin embargo en todo el proceso del desarrollo de esta tesis se identificaron puntos que tienen que ser mejorados, pero debido al alcance de la tesis ya no se pudieron abordar con detalle y en esta sección se enumeran para tomarlos como partida en un trabajo futuro.

5.6.1 Enfoque multifactorial

A lo largo de esta tesis fue evidente que en el crecimiento del parque vehicular influyen muchos factores, como lo es, el mismo parque vehicular de años anteriores, el precio de la gasolina, las políticas públicas, la creciente aparición de nuevos métodos de transporte, etc. Debido al alcance de esta tesis estos factores no fueron considerados, sin embargo, eso no significa que no pueden ser considerados empleando el enfoque que la PG ofrece, aunque si fue evidente (en particular en la sección 3.4.1) que entre mayor sea la cantidad de datos y de factores que intervienen en la creación del modelo, este será más confiable y reflejara de manera más

fidedigna el fenómeno que se esté modelando. Se deja para trabajo a futuro emplear un enfoque multifactorial empleando PG para el contexto del crecimiento del parque vehicular.

5.6.2 Programación en paralelo

La programación en paralelo es una herramienta que en últimas fechas ha tenido un gran auge, sobretodo debido a que ha demostrado ser una herramienta que reduce en gran medida el tiempo empleado por un algoritmo computacional.

El programa desarrollado en esta tesis no emplea programación en paralelo, es decir que cada rutina se lleva a cabo de manera secuencial. Dadas las características con las que se lleva a cabo el proceso de la PG, lo hacen un candidato ideal para realizar una implementación de dicho procedimiento empleando programación en paralelo.

5.6.3 Estimación y análisis del índice de asaltos a automovilistas en vialidad conflictivas

Seria de mucha utilidad para las autoridades y para los habitantes de la ZMVM tener una estimación de cómo se incrementan los asaltos a automovilistas, en que horarios y sobretodo en cuales vialidades. Ya que con esta estimación las autoridades podrían implementar acciones que ayuden a disminuir el tránsito en estas vialidades, ya sea agilizando el tránsito o supervisando dichas vialidades para impedir que se den estos asaltos.

Referencias:

- 1.- M. Lighthill, G. Whitman, pro. R. Soc. Lon. A229 (1955) 317.
- 2.- Koorey, G F. Mitchell, T J. Fisk, C R. TRAFFIC GROWTH PREDICTION (2002).
- 3.- W. Afzal, R. Torkar Suitability of genetic programming for software reliability growth modeling (2008).
- 4.- M. Schmidt, H. Lipson Distilling Free-Form Natural Laws from Experimental Data (2009).
- 5.- H. J. Payne. FREEFLO: A microscopic simulation model of free way traffic. Transportation Research Record 722 (Urban System Operations, National Academy of Sciences, Washington D. C. , 1979).
- 6.- R. Kühne. Proc. Int. Symp. on Transportation and Traffic Theory (MIT press Cambridge, 1987).
- 7.- H.J. Payne. FREEFLO: A macroscopic simulation model of freeway traffic. Transportation Research Record 722 (Urban System Operations, National Academy of Sciences, Washington D.C. ,1979); H. J. Payne, in Proc. Conf. on Research directions in computer control of urban traffic systems, 1979, edited by W. S. Levine, E. Lieberman, and J. J. Fearnside (American Society of Civil Engineers, New York, 1979), p. 251.
- 8.- D. Helbing. Improved Fluid Dynamic Model for Vehicular traffic. Physical Review E, 51:3164, 1995
- 9.- K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic", J. Physique. I 2, 2221, 1992.
- 10.- O. Biham, A. Middletown, and D. Levine, Phys. Rev. A 46, R6124 (1992).
- 11.- D. C. Gazis, R. Herman y R. Rothery, Nonlinear follow the leader models of traffic Flow . Operations Research., 9:545, 1961.
- 12.- Cramer, Michael Lynn (1985), "A representation for the Adaptive Generation of Simple Sequential Programs" in *Proceedings of an International Conference on Genetic Algorithms and the Applications*, Grefenstette, John J. (ed.), Carnegie Mellon University.
- 13.- Koza, J.R. (1992), *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press.
- 13.- Koza, J.R. (1992), *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press.
- 14.- Samuel, A. (1963) Some Studies in Machine Learning Using the Game of Checkers. In *Computers and Thought* (Feigenbaum, E. and Feldman, J., editors). McGraw-Hill.
- 15.- Mitchell, M. (1996) *An Introduction to Genetic Algorithms*. The MIT Press.
- 16.- Diday, E., The symbolic approach in clustering and related methods of Data Analysis. H.H. Bock (ed.), *Classification and Related Methods of Data Analysis*. Amsterdam: North-Holland, 1987.
- 17.- Diday, E., Introduction à l'approche symbolique en analyse des données. Première Journées Symbolique-Numérique. Université Paris IX Dauphine, Diciembre 1987.
- 18.- http://www.inegi.org.mx/est/lista_cubos/consulta.aspx?p=adm&c=8

Modelado de tráfico vehicular

- 19.- <http://www.stats.gov.cn/english/Statisticaldata/AnnualData>
- 20.- <http://www.laalmnac.com/transport/tr02.htm>
- 21.- <http://data.london.gov.uk/dataset/licensed-vehicles-numbers-borough/resource/45c47aba-682d-4be4-b62a-42215203c2ad#>
- 22.- <https://dmv.ny.gov/about-dmv/archives-statistical-summaries>
- 23.- http://www.observatoriodasmetropoles.net/download/auto_motos2013.pdf
- 24.- <http://www.citypopulation.de>
- 25.- <http://www.iprofesional.com/notas/122664-El-parque-automotor-de-China-alcanz-las-100-millones-de-unidades>
- 26.- <http://www.forbes.com/sites/jimgorzelay/2015/10/15/the-worlds-worst-traffic-jams-ever/#1260a2026c6c>
- 27.- <http://www.latimes.com/local/lanow/la-me-ln-la-worst-traffic-20160314-story.html>
- 28.- <http://mentalfloss.com/article/66462/why-la-traffic-so-awful>
- 29.- <https://www.fastcompany.com/1756746/building-more-roads-only-causes-more-traffic>
- 30.- <http://www.streetsblog.org/2015/09/17/message-from-brewers-congestion-hearing-its-the-road-pricing-stupid/>
- 31.- <http://riotimesonline.com/brazil-news/rio-real-estate/rio-has-third-worst-traffic-of-146-cities-in-the-world/>
- 32.- <http://www.elfinanciero.com.mx/nacional/en-2020-colapsarian-vialidades-habria-el-doble-de-autos-en-df.html>
- 33.-
<http://www.ptolomeo.unam.mx:8080/xmlui/bitstream/handle/132.248.52.100/417/A4.pdf?sequence=4>
- 34.- http://www.dlr.de/ts/en/desktopdefault.aspx/tabid-9883/16931_read-41000/
- 35.- <http://sumo.dlr.de/wiki/Networks/Import/OpenStreetMap>
- 36.- <http://sumo.dlr.de/wiki/Tutorials>
- 37.- <http://diarioenbici.com/la-bicicleta-el-transporte-mas-rapido-de-la-ciudad-de-mexico/>
- 38.- <https://www.sura.com/blogs/autos/ventajas-bicicleta-medio-transporte.aspx>
- 39.- https://www.facebook.com/pg/PumaRide/about/?ref=page_internal
- 40.- <http://eleconomista.com.mx/finanzas-personales/2014/08/28/beneficios-compartir-auto>