

---

**Desarrollo y validación de una metodología para la evaluación del desempeño ambiental del parque automotor liviano de Bogotá región a partir de ciclos típicos de conducción desarrollados por métodos estadísticos y de aprendizaje automático**

---



**Daniel Robayo Rueda**

**UNIVERSIDAD DE LA SABANA  
FACULTAD DE INGENIERÍA  
2023**

---

**Desarrollo y validación de una metodología para la evaluación del desempeño ambiental del parque automotor liviano de Bogotá región a partir de ciclos típicos de conducción desarrollados por métodos estadísticos y de aprendizaje automático**

---

Tesis presentada en cumplimiento parcial de los requisitos para obtener el grado en:

**Magister en Diseño y Gestión de Procesos**

Facultad de Ingeniería

Maestría en Diseño y Gestión de Procesos – Énfasis en Procesos Químicos

Modalidad de investigación

Presenta:

Daniel Robayo Rueda

Asesor:

Ph.D. Msc Cesar Luis Barraza Botet

Co asesor:

Ph.D. Msc Miguel Angel Uribe Laverde



FACULTAD DE INGENIERÍA

2023



---

**Desarrollo y validación de una metodología para la evaluación del desempeño ambiental del parque automotor liviano de Bogotá región a partir de ciclos típicos de conducción desarrollados a través de métodos estadísticos y de aprendizaje automático**

---

---

Ph.D. Msc Cesar Luis Barraza Botet  
Director de tesis

---

Ph.D. Msc Miguel Angel Uribe Laverde  
Co director de tesis

---

Ph.D. Msc. Fabiana Fernanda Franceschi García  
Profesor asistente e investigadora, Universidad de La Salle  
Jurado 1

---

Ph.D. Msc. Zamir Mera  
Profesor asistente, Universidad Técnica del Norte, Ecuador  
Jurado 2

---

Msc. Manuel Alfredo Figueredo Medina  
Director de la maestría en Gerencia de Ingeniería, Universidad de La Sabana  
Jurado 3

## **Dedicatoria**

*Soli Deo Gloria*

*A Dios, ya que si no fuese por su infinita gracia y misericordia nada podría haber hecho,*

*A mi madre Blanca Robayo, quien siempre me ha apoyado*

*A mi padre y mi abuela quien están desde el cielo en la presencia de mi Señor*

*A mis profesores por su continuo apoyo*

*A mi familia, y mis hermanos en la fe.*

*Así que, sea que coman o beban o cualquier otra cosa que hagan, háganlo todo para la gloria de  
Dios.*

*1 Corintios 10:31 NTV*

## **Agradecimientos**

En primer lugar quiero agradecer a Dios ya que día tras día me recuerda que está conmigo a pesar de las dificultades, quien por su gracia y amor me permite servirlo, incluso desde el ámbito profesional.

A mi mamá Blanca Flor Robayo, quien me ha enseñado a amar a Dios y a confiar en Él, siempre apoyándome en cumplir los sueños de Dios en mi vida, por su apoyo diario e incondicional, quien a pesar de mis muchas afecciones de salud durante los últimos meses ha estado allí apoyándome.

A mi papá José Otoniel Robayo, quien ya se encuentra en la presencia de Dios, y me enseñó a buscar día tras día más a Dios, y que Dios no es un cuento de hadas sino alguien real y que me ama, a quien puedo servir desde mi ámbito profesional, quien fue el que me metió toda esta locura de aprender cómo funcionan las cosas, el amor por los vehículos, y el trabajo bien hecho, de quien recuerdo sus enseñanzas y una de sus últimas palabras, “No quites tu mirada de Jesús”.

A mi abuela Aura Matilde Rueda de Cortes, quien está disfrutando de la presencia de su amado Dios en la eternidad. quien me enseñó también a amar a Dios, a hacer bien las cosas, y el amar a los demás en servicio, quien se sentaba conmigo cuando era pequeño a estudiar y a enseñarme a dibujar, leer, escribir y sumar, de quien recuerdo su abrazo de amor, un abrazo de mi Señor Jesús.

A mi primo Juan Robayo, quien a pesar de la distancia ha estado presente brindándome apoyo y algunos regaños.

A mi tío Ricardo Robayo, quien me ha apoyado desde que era muy pequeño, quien me regalo mi primer computador, y ha sido para mí un ejemplo de amor a Dios y a su familia.

Especialmente a mi profesor, maestro y tutor en el ámbito profesional el profesor Cesar Luis Barraza Botet, muchas gracias, profesor por creer en que puedo desarrollar y aplicar mis conocimientos en servicio a los demás, a dar siempre más de mí y a trabajar duro, muchas gracias también por su ejemplo, su paciencia y su dedicación.

A mi cotutor de tesis el profesor Miguel Angel Uribe Laverde, muchas gracias por enseñarme más acerca de este mundo de los algoritmos, y la inteligencia artificial, así como a repetir los trabajos cuando es necesario para aprender más.

Un agradecimiento especial al profesor Julian Mauricio Echeverry Mejía, por todos sus consejos y asesoramiento en la construcción del dispositivo IVDR, conocimiento en el uso del dinamómetro de chasis y conceptos acerca de los ciclos de conducción, por su paciencia de quien fuese mi director de programa mientras estudiaba ingeniería mecánica, así como todo el apoyo y comprensión, muchas gracias.

A todos los profesores, quienes me apoyaron y enseñaron para que este proyecto saliera adelante.

A los técnicos de laboratorio quienes estuvieron allí para el desarrollo del proyecto.

A mis profesores de música y líderes de jóvenes durante 10 años, Yamid Martinez, y Diana Diaz, quienes son mis tutores desde que llegue a la academia de música, y me enseñaron disciplina y trabajo duro, así como amor y servicio a Dios, gracias por ayudarme en esta etapa y por creer en mí.

A mis líderes de jóvenes actuales Willy y Carito, quienes siempre oran por mí y a quienes aprecio mucho.

A mi querida hermana mayor (en la fe), Lorena Gualtero y a su esposo Julian, mi mejor amiga, por apoyarme en oración y aguantar mis chistes malos siempre.

A mi querida mejor amiga y hermana en la fe Alejandra Vargas, por aguantarme siempre que la he querido molestar, y también mi mal carácter en ocasiones, y por alentarme a iniciar esta etapa que ya culmina, y quien ha sido un ejemplo de dedicación durante estos años.

A mis mejores amigos Santiago, Bryan y Esteban, así como a mis amigas Paula y Leidy por su apoyo y oración.

A la congregación de la iglesia Wesleyana Suba quienes se nunca me han soltado de sus oraciones.

Muchas gracias.

## Índice

Dedicatoria .....	5
Agradecimientos .....	6
Lista de figuras.....	11
Lista de tablas .....	14
Lista de Algoritmos.....	16
Lista de ecuaciones.....	17
Lista de abreviaturas .....	18
Resumen .....	20
Abstract .....	22
Resumen gráfico.....	24
<b>1. Planteamiento del problema y justificación .....</b>	<b>25</b>
<b>2. Objetivos.....</b>	<b>29</b>
<b>2.1. Objetivo general.....</b>	<b>29</b>
<b>2.2. . Objetivos específicos .....</b>	<b>29</b>
<b>3. Marco conceptual.....</b>	<b>30</b>
<b>3.1. Generalidades de los ciclos típicos de conducción .....</b>	<b>30</b>
<b>3.1.1. Principales Características de los ciclos típicos de conducción.....</b>	<b>30</b>
<b>3.1.2. Desarrollo de un ciclo típico de conducción .....</b>	<b>32</b>
<b>3.1.3. Ciclos típicos de conducción estándar.....</b>	<b>33</b>
<b>3.1.4. Comparación de los ciclos típicos de conducción contruidos y los ciclos típicos de conducción estándar .....</b>	<b>35</b>
<b>3.2. Adquisición de datos en tiempo real de automóviles .....</b>	<b>38</b>
<b>3.2.1. Protocolos de comunicación.....</b>	<b>38</b>
<b>3.2.2. Desarrollo de un IVDR (In-vehicle data recorder) .....</b>	<b>42</b>
<b>3.2.3. Análisis previo de los datos adquiridos por un IVDR.....</b>	<b>45</b>
<b>3.3. Generalidades sobre los métodos de aprendizaje automático y estadísticos para la construcción de los ciclos típicos de conducción .....</b>	<b>45</b>
<b>3.3.1. Métodos generales de aprendizaje automático y estadísticos para la construcción de ciclos típicos de conducción .....</b>	<b>46</b>
<b>3.4. Medición de emisiones en vehículos del parque automotor liviano a gasolina .....</b>	<b>52</b>
<b>3.4.1. Generalidades sobre la medición de emisiones hechas sobre un dinamómetro de chasis</b>	<b>53</b>
<b>3.4.2. Medición de emisiones con base en la aplicación de un ciclo típico de conducción ....</b>	<b>54</b>
<b>3.4.3. Métodos para el análisis de las emisiones registradas .....</b>	<b>55</b>

<b>4. Metodología</b> .....	57
<b>4.1. Desarrollo metodológico del primer objetivo específico</b> .....	57
4.1.1. Diseño, construcción e instalación de los instrumentos de medición IVDR.....	58
4.1.2. Diseño de experimentos en ruta.....	66
4.1.3. Toma de datos en campo.....	67
<b>4.2. Metodología para el desarrollo del segundo objetivo específico</b> .....	70
4.2.1. Selección de técnicas para la construcción de los ciclos típicos de conducción característicos de la zona (Bogotá – Región).....	71
4.2.2. Preprocesamiento de datos el dispositivo IVDR.....	81
4.2.3. Evaluación comparativa de los métodos de análisis de Aprendizaje automático.....	91
4.2.4. Definición del ciclo típico de conducción para Bogotá - Región.....	92
4.2.5. Comparación del CTC encontrado con los ciclos estándar internacionales.....	94
<b>4.3. Desarrollo metodológico del tercer objetivo específico</b> .....	94
4.3.1. Selección de vehículos experimentales para el desarrollo de pruebas dinámicas.....	95
4.3.2. Diseño de experimento en dinamómetro de chasis.....	95
4.3.3. Desarrollo de pruebas dinámicas bajo el CTC Bogotá y CTCs estándar.....	95
4.3.4. Medición de diferentes contaminantes con el analizador de gases.....	97
<b>4.4. Desarrollo metodológico del objetivo cuarto específico</b> .....	98
4.4.1. Cálculo de los factores de emisión del CTC Bogotá.....	98
4.4.2. Comparación de los factores de emisión del CTC Bogotá respecto a los CTCs estándar. <sup>99</sup>	
<b>5. Resultados</b> .....	99
<b>5.1. Desarrollo del dispositivo IVDR</b> .....	99
5.1.1. Ensamblaje mecánico del dispositivo IVDR.....	99
5.1.2. Esquemático de conexiones del dispositivo IVDR.....	102
5.1.3. Validación del dispositivo IVDR.....	103
5.1.4. Resultados de los datos obtenidos por el módulo GPS.....	105
<b>5.2. Desarrollo de algoritmos para la construcción de CTCs (Evaluación comparativa de los   métodos de aprendizaje automático)</b> .....	106
5.2.1. Análisis general de las metaheurísticas utilizadas.....	106
5.2.2. Análisis estadístico de las metaheurísticas utilizadas.....	111
5.2.3. Ciclos típicos desarrollados para cada base de datos.....	116
5.2.4. Comparación de los ciclos construidos con los ciclos internacionales.....	118
5.2.5. Mejoramiento de los CTCs desarrollados.....	119
<b>5.3. Construcción del CTC para la zona Bogotá Región – CTC Bogotá</b> .....	125

5.3.1.	Validación de la metaheurística ILS-RANDOM para la base de datos Bogotá-región.	
	125	
5.3.2.	Construcción, comparación y selección del CTC Bogotá .....	128
5.4.	Pruebas dinámicas preliminares del CTC Bogotá y CTC estándar. ....	136
5.5.	Demostración de determinación de factores de emisión.....	138
6.	Conclusiones.....	139
7.	Trabajo futuro.....	141
	Referencias .....	141
	Anexos.....	147
A.	Desarrollo del software del dispositivo IVDR .....	147
A.1.	<i>Desarrollo del programa principal</i> .....	147
A.2	<i>Interrupción RX UART</i> .....	149
A.3.	<i>Interrupción RX CAN</i> .....	150
A.4.	<i>Interrupción PIT</i> .....	152
B.	Distribución de velocidades para todas las metaheurísticas sin refinamiento de parámetros. ....	153
C.	Resultados de la evaluación comparativa de algoritmos con el refinamiento de parámetros	
	156	
D.	Distribución de velocidades para todas las bases de datos de Bogotá región.....	164

## Lista de figuras

Figura 3.1. Representación general de un ciclo típico de conducción. (Giakoumis, 2016) .....	30
Figura 3.2. Metodología general para la construcción de un ciclo típico de conducción. (Giakoumis, 2016) .....	32
Figura 3.3. Estructura general para la trama de datos de un mensaje CAN estándar. (Echeverry Mejía, 2018; International Organization for Standardization (ISO), 2006) .....	39
Figura 3.4. Diagrama modular del dispositivo IVDR (Echeverry Mejía, 2018) .....	44
Figura 3.5. Esquema de un dinamómetro de chasis conectado a un analizador de gases para la medición de emisiones. (Giakoumis, 2016).....	54
Figura 3.6. Esquema de emisiones generadas por un vehículo (Ministerio de Ambiente y Desarrollo Sostenible, 2017) .....	56
Figura 4.1. Esquema de la metodología para el desarrollo del primer objetivo específico .....	58
Figura 4.2. Microcontrolador SKEAZ128MLK4 – Tarjeta de desarrollo NXP FRDM KEAZ128 .....	59
Figura 4.3. Datalogger SparkFun OpenLog DEV -137 12.....	60
Figura 4.4. Módulo GPS NEO 6M UBLOX .....	60
Figura 4.5. Convertidores DC-DC BUCK XL4015 y BUCK LM2596.....	61
Figura 4.6. Máquina de estados del dispositivo IVDR.....	65
Figura 4.7. Vehículos experimentales para la toma de datos. De izquierda a derecha, Nissan Versa 2013 y Ford Fiesta 2017 .....	68
Figura 4.8. Esquema de la metodología para desarrollar el segundo objetivo específico .....	70
Figura 4.9. Ejemplo de operación Crossover .....	77
Figura 4.10. Metodología para el Desarrollo del preprocesamiento de datos .....	81
Figura 4.11. Comparación entre datos de OBDII en formato binario (izquierda) y datos de OBDII en formato ASCII (derecha) .....	82
Figura 4.12. Metodología para el cálculo del consumo de combustible .....	86
Figura 4.13. Proceso metodológico para el desarrollo del tercer objetivo específico .....	94
Figura 4.14. Izquierda: Vehículo experimental durante una prueba dinámica. Derecha: Dinamómetro de chasis Mustang Dyno MD-110SE.....	97
Figura 4.15. Características de los sensores de NO <sub>x</sub> del analizador de gases. ....	98
Figura 4.16. Características de los sensores de CO <sub>2</sub> , CO, THC, O <sub>2</sub> y del analizador de gases .....	98
Figura 4.17. Conexión del analizador de gases al vehículo experimental.....	98
Figura 5.1. Ensamblaje mecánico del dispositivo IVDR y sus componentes. ....	100
Figura 5.2. Ensamblaje mecánico del dispositivo IVDR.....	101
Figura 5.3. Dispositivo IVDR final.....	101
Figura 5.4. Esquemático de conexiones del dispositivo IVDR.....	102
Figura 5.5. Recorridos realizados durante el proceso de recolección de datos .....	105
Figura 5.6. Distribución de los microciclos acorde a la velocidad media para la metaheurística ILS-RANDOM. ....	111
Figura 5.7. Box-plot del error medio para todas las metaheurísticas en todas las bases de datos .....	112
Figura 5.8. Box-plot del tiempo de ejecución para todas las metaheurísticas en todas las bases de datos .....	113
Figura 5.9. Ciclo típico de conducción construido para El Paso .....	116
Figura 5.10. Ciclo típico de conducción para San Antonio .....	117
Figura 5.11. Ciclo típico de conducción para Houston/Galveston .....	117
Figura 5.12. Comparación del error medio de los parámetros para cada base de datos. ....	119
Figura 5.13. CTC construido para El Paso con el refinamiento de parámetros.....	121

Figura 5.14. CTC construido para San Antonio con el refinamiento de parámetros .....	122
Figura 5.15. CTC construido para Houston/Galveston con el refinamiento de parámetros .....	122
Figura 5.16. Comparación del error medio de los parámetros refinados, respecto al error medio de los CTCs construidos.....	124
Figura 5.17. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística ILS -RANDOM. ....	127
Figura 5.18. Ciclo seleccionado por la metodología MWD-CP. ....	128
Figura 5.19. CTC desarrollado para la zona Bogotá-región con base al error medio general y sin considerar el tipo de vehículo. ....	129
Figura 5.20. CTC desarrollado para la zona Bogotá-región con base al error medio general y considerando el tipo de vehículo.....	130
Figura 5.21. CTC desarrollado para la zona Bogotá-región con base la metodología combinada y sin considerar el tipo de vehículo. ....	131
Figura 5.22. CTC desarrollado para la zona Bogotá-región con base la metodología combinada y considerando el tipo de vehículo.....	132
Figura 5.23. Ciclo seleccionado – Ciclo Bogotá .....	136
Figura 5.24. Seguimiento del vehículo experimental al CTC construido para la zona Bogotá región. ....	137
Figura 5.25. Seguimiento del vehículo experimental al CTC FTP-75.....	137
Figura 5.26. Seguimiento del vehículo experimental al CTC WLTC-clase 1.....	138
Figura A. 1. Diagrama de flujo de la rutina del programa principal	149
Figura A. 2. Diagrama de flujo de la rutina RX UART0.....	150
Figura A. 3. Diagrama de flujo de la rutina RX CAN.....	152
Figura A. 4. Diagrama de flujo de la rutina PIT0.....	153
Figura A. 5 Distribución de los microciclos acorde a la velocidad media para la metaheurística GRASP-RANDOM .....	153
Figura A. 6. Distribución de los microciclos acorde a la velocidad media para la metaheurística GRASP-FORGY.....	154
Figura A. 7. Distribución de los microciclos acorde a la velocidad media para la metaheurística ILS-FORGY.....	155
Figura A. 8. Distribución de los microciclos acorde a la velocidad media para la metaheurística GKA. ....	155
Figura A. 9. Distribución del error medio para todas las bases de datos con refinamiento de parámetros .....	157
Figura A. 10. Distribución del error medio para todas las bases de datos y cada algoritmo con refinamiento de parámetros.....	158
Figura A. 11. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GRASP -RANDOM con refinamiento de parámetros.....	159
Figura A. 12. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GRASP -FORGY con refinamiento de parámetros .....	160
Figura A. 13. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística ILS-RANDOM con refinamiento de parámetros .....	161
Figura A. 14. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística ILS-FORGY con refinamiento de parámetros.....	162
Figura A. 15. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GKA con refinamiento de parámetros.....	163
Figura A. 16. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GRASP -FORGY .....	164

Figura A. 17. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GRASP -FORGY .....	164
Figura A. 18. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística ILS -FORGY .....	165
Figura A. 19. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GKA .....	165

## Lista de tablas

Tabla 1.1. Comparación de los límites establecidos para los contaminantes de aire. (Elaboración propia)	26
Tabla 3.1. Categorización de los ciclos típicos de conducción acorde al tipo de ruta (Elaboración propia)	32
Tabla 3.2. Características principales de los ciclos típicos de conducción estándar. (Elaboración propia)	34
Tabla 3.3. Comparación de los ciclos típicos de conducción algunos de CTCs construidos (Elaboración propia)	36
Tabla 3.4. Definición de cada una de las partes de la trama de datos del protocolo CAN. (Adaptado Echeverry Mejía, 2018)	40
Tabla 3.5. PIDs comúnmente utilizados (Echeverry Mejía, 2018)	41
Tabla 3.6. Descripción de un mensaje de OBDII (Elaboración propia)	42
Tabla 3.7. Características generales de un IVDR (Adaptado de Echeverry Mejía, 2018)	43
Tabla 3.8. Requisitos de medición de gases para un analizador de gases acorde a la norma NTC 4983 (ICONTEC, 2012)	55
Tabla 4.1. Costos de cada uno de los componentes del IVDR	62
Tabla 4.2. Ficha técnica general de los vehículos experimentales para la toma de datos	68
Tabla 4.3. PIDs solicitados – PIDs de cada vehículo experimental	69
Tabla 4.4. Parámetros seleccionados para el análisis del ciclo típico de conducción	79
Tabla 4.5. Revisión de la literatura acerca de la definición de los modos de conducción	80
Tabla 4.6. Descripción de las fórmulas para el cálculo de los parámetros de OBDII	83
Tabla 4.7. Organización de un mensaje GPS	90
Tabla 5.1. Validación de tiempo del dispositivo IVDR	104
Tabla 5.2. Descripción de la base de datos	106
Tabla 5.3. Parámetros calculados para las bases de datos seleccionadas	107
Tabla 5.4. <i>Resultados generales de las metaheurísticas utilizadas</i>	107
Tabla 5.5. Distribución de microciclos en cada uno de los clusters para cada base de datos	109
Tabla 5.6. Velocidad promedio en cada clúster para cada base de datos	109
Tabla 5.7. Validación de supuestos para el promedio del error medio para cada base de datos	114
Tabla 5.8. Validación de supuestos para el tiempo de ejecución para cada base de datos	114
Tabla 5.9. Resultados de las pruebas no paramétricas para el promedio del error medio y para tiempo de ejecución de los CTCs para cada base de datos	115
Tabla 5.10. Comparación de los parámetros de cada ciclo típico de conducción construido	118
Tabla 5.11. Comparación de los errores medios para cada base de datos	118
Tabla 5.12. Parámetros seleccionados para el mejoramiento de los CTCs	120
Tabla 5.13. Comparación de parámetros de los CTCs desarrollados con el refinamiento de parámetros	123
Tabla 5.14. Comparación de los errores medios para cada base de datos para los CTC con refinamiento de parámetros	124
Tabla 5.15. Resultados generales de las metaheurísticas utilizadas para zona Bogotá región	126
Tabla 5.16. Distribución de porcentajes en cada microciclo para cada metaheurística en la base de datos Bogotá región	126
Tabla 5.17. Velocidad media de los microciclos asignados en cada clúster para la base de datos Bogotá región	126
Tabla 5.18. <i>Pesos relativos de cada uno de los parámetros característicos en cada metodología desarrollada. (CV – con el parámetro tipo de vehículo, SV - sin el parámetro tipo de vehículo)</i>	133
Tabla 5.19. Comparación de los parámetros de los CTCs construidos para la zona Bogotá región	134

Tabla 5.20. Comparación del error medio y las diferencias ponderadas para todos los ciclos construidos para la zona Bogotá región.....	135
Tabla 5.21. Comparación de los factores de emisión para el vehículo de pruebas dinámica Nissan Versa 2013.....	139
Tabla A. 1. Resultados generales para cada algoritmo con el refinamiento de parámetros.	156
Tabla A. 2. Distribución de porcentajes de microciclos para todas las bases de datos. ....	156
Tabla A. 3. Promedio de la velocidad media en cada cluster para cada base de datos. ....	156

## Lista de Algoritmos

Algoritmo 4.1 Análisis de Componentes Principales (PCA).....	72
Algoritmo 4.2 Algoritmo K-means .....	73
Algoritmo 4.3 Metaheurística GRASP combinada con k-means .....	74
Algoritmo 4.4 Metaheurística ILS combinada con el algoritmo k-means .....	75
Algoritmo 4.5 Genetic k-means .....	78

## Lista de ecuaciones

Ecuación 3.1 .....	46
Ecuación 3.2 .....	47
Ecuación 3.3 .....	49
Ecuación 3.4 .....	51
Ecuación 3.5 .....	53
Ecuación 3.6 .....	57
Ecuación 4.1 .....	80
Ecuación 4.2 .....	84
Ecuación 4.3 .....	84
Ecuación 4.4 .....	85
Ecuación 4.5 .....	86
Ecuación 4.6 .....	88
Ecuación 4.7 .....	89
Ecuación 4.8 .....	89
Ecuación 4.9 .....	89
Ecuación 4.10 .....	89
Ecuación 4.11 .....	89
Ecuación 4.12 .....	93

## Lista de abreviaturas

ABREVIATURA	SIGNIFICADO
<i>ACC</i>	Ajuste de combustible a corto plazo
ACK	Acknowledgement
<i>AE</i>	Error medio
$A_f$	Área frontal del vehículo
$AF_{ideal}$	Relación aire combustible ideal
ANOVA	Analysis of Variance
ASCII	American Standard code for Information Interchange
CAN	Controller Area Network
$C_d$	Coefficiente de arrastre
CDCA	Common Driving Cycle ARTEMIS
$C_{fr}$	Coefficiente de rodadura
CRC	Cyclic Redundancy Check
CTC	Ciclo típico de conducción
DLC	Data length code
DTC	Diagnostic trouble codes
<i>E</i>	Emisión producida
ECU	Engine Control Unit
<i>ED</i>	Desplazamiento del motor
EOF	End of frame
<i>Fat</i>	Factor de actividad en el tiempo $t$
<i>FC</i>	Consumo de combustible
$FE_i$	Factor de emisión del contaminante $i$
$FE_i$	Light-emitting diode
FTP-72	Federal Test Procedure - 1972
FTP-75	Federal Test Procedure - 1975
GKA	Genetic K-means Algorithm
GPGGA	Global Positioning System Fix Data
GPS	Global Positioning System
GRASP	Greedy Randomized Search Procedure
IARC	Agencia Internacional para la Investigación en Cancer
IAT	Intake Air Temperature sensor
ID	Identifier
IDE	Identifier extension
ILS	Iterated Local Search
IRQ	Interruption request
IVDR	In-vehicle Data Recorder

MAF	Mass Air flow sensor
MAP	Mass Air Pressure sensor
<i>ML</i>	Carga del motor
<i>MM</i>	Masa molar del aire
$m_v$	Masa del vehículo
MWD - CP	Minimum Weighted Differences of Characteristics Parameters
<i>n</i>	Cantidad de parámetros
NEDC	New European Driving Cycle
NREL	National Renewable Energies Laboratory of United States
OBD2	On Board Diagnostic version 2.0
OMS	Organización Mundial de la Salud
PCA	Principal Component Analysis
PEM	Portable Emission System
$\hat{P}_i$	Parametro promedio de todos los microciclos
$P_i$	Parametro <i>i</i> del CTC constuido
PID	Parameter Identifier
PIT	Periodic Interrupt Timer
PM	Material particulado
<i>PPA</i>	Posición del pedal del acelerador
<i>R</i>	Constante universal de los gases
RPM	Revoluciones por minuto
RTR	Remote Transmission Request
RWA	Relative Weight Analysis
SAE	Society of Automotive Engineers
SOF	Start of frame
TPM	Transition Probability Matrix
TPS	Throttle Position sensor
UART	Universal Asynchronous Receiver – Transmitter
(UC)/LA 92	California Unified Cycle
<i>VE</i>	Eficiencia volumétrica
$V_{max}$	Voltaje máximo medido
$V_{min}$	Voltaje mínimo medido
<i>VSO</i>	Voltaje del sensor de oxígeno
$V_t$	Voltaje medido por el sensor de oxigeno
$W_i$	Peso relativo del parámetro característico de cada parámetro
WLTP	World harmonized light-duty test procedure
$\lambda$	Relación entre la relación aire combustible real y la relación aire combustible ideal
$\rho_a$	Densidad del aire

## Resumen

La contaminación del aire ha producido en los últimos años graves enfermedades en los seres humanos, tales como: muertes prematuras, enfermedades cardiovasculares y cáncer de pulmón (World Health Organization, 2018). En el año 2013, la OMS (Organización Mundial de la Salud) junto con la IARC (*World Health Organization International Agency for Research on cancer*), establecieron que la contaminación del aire es cancerígena para los seres humanos. Bogotá es una de las ciudades donde se presenta contaminación del aire, la cual desde el año 1998 excedió los niveles recomendados para el material particulado menor a 10 micrómetros (PM<sub>10</sub>) (Observatorio ambiental de Bogotá, 2022), controlando este indicador desde el año 2012 en promedio anual, pero mostrando altos índices para algunos sectores de la ciudad (Observatorio Ambiental de Bogotá, 2022) en comparación a los límites establecidos por la resolución 2254 del 2017 (Ministerio de ambiente y desarrollo sostenible, 2018).

Si bien, se tienen medidas regulatorias sobre los niveles contaminantes para Bogotá como la resolución 2254, estos límites establecidos son mayores a los recomendados por la OMS (World Health Organization, 2021a). Para el año 2020, el parque automotor liviano de la ciudad de Bogotá representó un 72.4% del parque automotor de la ciudad, los cuales se encuentran entre la primera o segunda fuente móvil de generación de contaminantes (Secretaría de Ambiente de Bogotá, 2022). Debido a la cantidad de contaminantes producidos en la ciudad de Bogotá, y el porcentaje de vehículos livianos que transitan, se plantea desarrollar una metodología para estimar los factores de emisión de este sector del parque automotor para contaminantes como: dióxido de carbono (CO<sub>2</sub>), monóxido de carbono (CO), óxidos de nitrógeno (NO<sub>x</sub>), e hidrocarburos (THC).

Para esto, se propone la construcción de ciclos típicos de conducción de rutas variables para vehículos livianos de combustión interna los cuales utilicen como combustible gasolina para la zona Bogotá región, mediante la recolección de datos a través de un dispositivo IVDR (In-vehicle data recorder, por sus siglas en inglés), el cual fue construido y validado. Los datos obtenidos fueron preprocesados, y fueron la base para la construcción de ciclos típicos de conducción utilizando métodos estadísticos y de aprendizaje automático tales como el método de diferencias ponderadas (Minimum weighted differences – MWD, por sus siglas en inglés), una combinación

entre el algoritmo de cadenas de Márkov, el algoritmo K-means, las metaheurísticas ILS (*Iterated Local Search*, por sus siglas en inglés), GRASP (*Greedy Randomized Search Procedure*, por sus siglas en inglés) y GKA (*Genetic K-means Algorithm*, por sus siglas en inglés), junto con un Análisis de Componentes Principales (PCA, por sus siglas en inglés), y una combinación entre el método estadístico y la combinación de los métodos de aprendizaje automático.

Estos ciclos típicos de conducción construidos evidenciaron la representatividad general de los vehículos en la zona Bogotá región y el consumo de combustible; de igual manera, permitieron la validación en el dinamómetro de chasis, y la medición de los factores de emisión, los resultados obtenidos comparados con los límites establecidos por la Secretaría distrital de ambiente y el Ministerio de ambiente y desarrollo sostenible, demuestran la flexibilidad respecto al control de contaminantes y el mejoramiento de tecnologías. Así como, la posible implementación de la metodología presentada en futuros contextos regulatorios para el control de emisiones, mediante pruebas dinámicas.

**Palabras clave:** Analizador de gases, Ciclo típico de conducción, dinamómetro de chasis, factores de emisión IVDR, k-means, metaheurística.

## Abstract

Air pollution in recent years has led to serious illnesses in humans, such as premature deaths, cardiovascular diseases, and lung cancer (World Health Organization, 2018). In 2013, the WHO (World Health Organization), along with the IARC (International Agency for Research on Cancer), established that air pollution is carcinogenic to humans. Bogotá is one of the cities with larger levels of emissions, which has exceeded recommended levels for particulate matter smaller than 10 micrometers (PM10) since 1998 (Observatorio ambiental de Bogotá, 2022). While this indicator has been under control on an annual average since 2012, there are still high levels in certain sectors of the city (Observatorio Ambiental de Bogotá, 2022) compared to the limits set by Resolution 2254 of 2017 (Ministerio de ambiente y desarrollo sostenible, 2018).

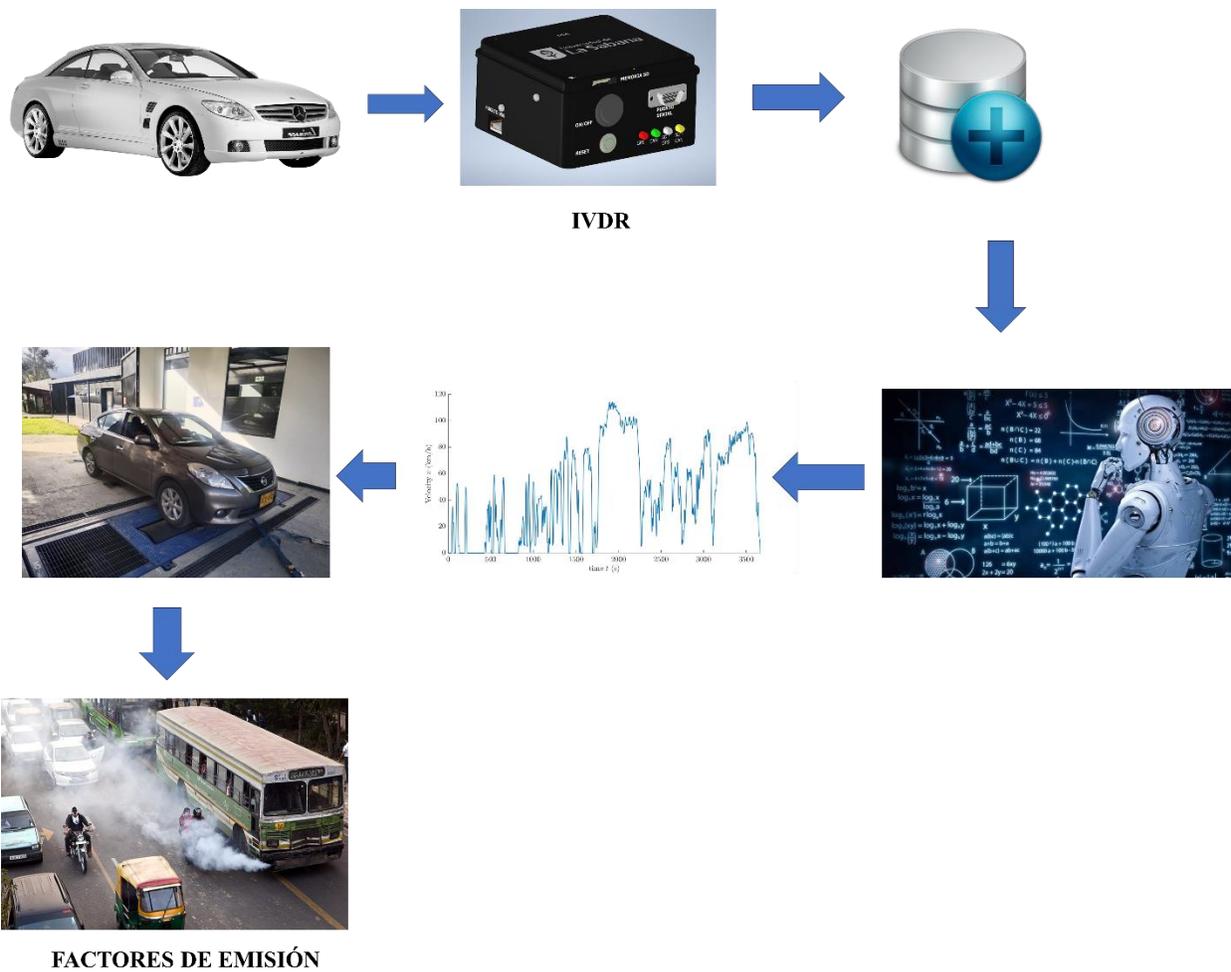
Although there are regulatory measures in place regarding pollutant levels, such as Resolution 2254, these established limits are higher than those recommended by the WHO (World Health Organization, 2021a). In the year 2020, light duty vehicles accounted for 72.4% of the vehicle fleet in Bogotá, which are among the first or second mobile resources of pollutant generation (Secretaría de Ambiente de Bogotá, 2022). Due to the amount of pollutants produced in Bogotá and the percentage of light duty vehicles on the roads, there is a need to understand the emission factors of this vehicle segment for pollutants such as carbon dioxide (CO<sub>2</sub>), carbon monoxide (CO), nitrogen oxides (NO<sub>x</sub>), and hydrocarbons (THC).

To address this, the construction of typical driving cycles involving varied routes for internal combustion engine light-duty vehicles fueled with gasoline in the Bogotá region is proposed. This involves collecting data through an In-Vehicle Data Recorder (IVDR) device, which has been constructed and validated. The obtained data underwent preprocessing and served as the foundation for constructing typical driving cycles using statistical and machine learning methods such as the Minimum Weighted Differences (MWD) method, a combination of the Markov Chain algorithm, K-means algorithm, metaheuristics including Iterated Local Search (ILS), Greedy Randomized Search Procedure (GRASP), and Genetic K-means Algorithm (GKA). Additionally, Principal Component Analysis (PCA) was used along with a combination of statistical methods and machine learning techniques.

These constructed typical driving cycles demonstrated the overall representativeness of vehicles in the Bogotá region and their fuel consumption. They also facilitated validation on a chassis dynamometer and measurement of emission factors. Results obtained, when compared with limits set by Secretaría distrital de ambiente and Ministerio de ambiente desarrollo sostenible, highlight the flexibility in terms of pollution control and technology enhancement. The methodology proposed could be implemented in future vehicle emissions regulatory contexts.

**Keywords:** Chassis dynamometer, Emissions factors, Gas analyzer, k-means, Metaheuristic, Typical Driving Cycle

## Resumen gráfico



## 1. Planteamiento del problema y justificación

Para el año 2016, la Organización Mundial de la Salud (OMS) evidenció que la contaminación del aire en áreas rurales como urbanas fue la causa de 4.2 millones de muertes prematuras por año. (World Health Organization, 2018). Así mismo, la contaminación ambiental se encuentra relacionada con enfermedades respiratorias, cardiovasculares y con el desarrollo de cáncer de pulmón, generando el 6% de las muertes ocasionadas por la contaminación del aire (World Health Organization, 2018). Una evaluación realizada por la Agencia para la investigación del cáncer de la OMS (IARC) en el año 2013 estableció que la contaminación del aire es cancerígena para los humanos, en mayor relación con el material particulado  $PM_{2.5}$  (material particulado con un diámetro de 2.5 micrómetros o menos) (International Agency for Research on Cancer, 2013).

Por esta razón, en el año 2021, la OMS actualizó los límites para los diferentes tipos de contaminantes producidos por diferentes fuentes, como: fuentes de combustión estacionaria, fuentes industriales, fuentes intermitentes y fuentes móviles, entre otras. Estos límites se establecieron para contaminantes tales como: Material particulado ( $PM_{2.5}$  –  $PM_{10}$ ), Ozono ( $O_3$ ), dióxido de nitrógeno ( $NO_2$ ), dióxido de azufre ( $SO_2$ ), y monóxido de carbono (CO) (World Health Organization, 2021b). Debido a los efectos producidos en la salud humana por la contaminación del aire, es necesario desarrollar métodos que permitan conocer la cantidad de contaminantes emitidos, con el fin de evitar estas consecuencias.

En el año 2018 el Ministerio de ambiente y desarrollo sostenible presentó límites máximos permisibles de contaminantes en el aire ( $PM_{2.5}$ ,  $PM_{10}$ ,  $SO_2$ ,  $O_3$  y CO) (Ministerio de ambiente y desarrollo sostenible, 2018). En contraste, para el año 2021 la Organización mundial de la Salud (OMS) estableció los lineamientos para los límites de contaminantes en el aire (World Health Organization, 2021b), donde al comparar estos niveles se evidencia que los niveles máximos permitidos propuestos por cada organización difieren entre sí. Específicamente, para contaminantes como  $PM_{2.5}$ ,  $PM_{10}$ ,  $SO_2$ , los límites establecidos por el Ministerio de ambiente superan los lineamientos establecidos, donde los límites para estos contaminantes son establecidos por la OMS dentro del objetivo provisional 2, en camino a llegar a progresivamente a los lineamientos establecidos.

Tabla 1.1. Comparación de los límites establecidos para los contaminantes de aire. (Elaboración propia)

Contaminante ( $\mu\text{g}/\text{m}^3$ )	Nivel máximo permitido por el Ministerio de Ambiente	Periodo de exposición	Lineamientos establecidos por la OMS	Periodo de exposición
PM <sub>10</sub>	50	Anual	15	Anual
	100	24 horas	45	24 horas
PM <sub>2.5</sub>	25	Anual	5	Anual
	50	24 horas	15	24 horas
SO <sub>2</sub>	50	24 horas	40	24 horas
	100	1 hora	-	1 hora
NO <sub>2</sub>	60	Anual	10	Anual
	200	1 hora	200	1 hora
O <sub>3</sub>	100	8 horas	100	8 horas
CO	5000	8 horas	10000	Anual
	35000	1 hora	35000	1 hora

La ciudad de Bogotá, con más de siete millones de habitantes (DANE, 2019), el parque automotor estaba conformado por 2456104 vehículos de los cuales el 72.4% eran vehículos livianos (automóviles, camperos y camionetas) para el año 2022. Para el mismo año, el parque automotor generó 1788,9 toneladas de PM<sub>10</sub>, 1610,7 toneladas de PM<sub>2.5</sub>, así como 43097 toneladas de óxidos de nitrógeno (NO<sub>x</sub>), 292677,1 toneladas de monóxido de carbono (CO), 50566,3 toneladas de compuestos orgánicos volátiles (VOC), y 824,7 toneladas de óxidos de azufre (SO<sub>x</sub>). Estos valores se pueden comparar con las emisiones generadas por fuentes fijas industriales en la ciudad de Bogotá, en el mismo año, donde se generaron 363 toneladas de PM<sub>10</sub>, 159 toneladas de PM<sub>2.5</sub>, 3290 toneladas de óxidos de nitrógeno (NO<sub>x</sub>), 690 toneladas de dióxido de azufre (SO<sub>2</sub>), 1130 toneladas de monóxido de carbono (CO) y 184 toneladas de compuestos volátiles orgánicos (VOC), acorde a la metodología *Bottom-Up* (Secretaría de Ambiente de Bogotá, 2022). En comparación, las fuentes móviles generan una mayor cantidad de contaminantes relacionados con el material particulado, óxidos de carbono y monóxido de carbono, por lo cual es de gran importancia emplear métodos para la reducción de estos contaminantes generados por fuentes móviles.

Particularmente, el parque automotor liviano fue el responsable de producir 15% de PM<sub>10</sub>, 15% de PM<sub>2.5</sub>, 32,1% de NO<sub>x</sub>, 30,1 de CO, 70% de VOC y 59,5% de SO<sub>x</sub> de las emisiones generadas por el parque automotor de la ciudad, donde los porcentajes restantes corresponden a vehículos de transporte, vehículos pesados y motocicletas. Debido a la cantidad de contaminantes

generados por el parque automotor liviano, es necesario desarrollar una metodología que permita conocer los factores de emisión de los vehículos que pertenecen a este sector (Secretaría de Ambiente de Bogotá, 2022). Existen diferentes metodologías para conocer los factores de emisión de dichos vehículos.

- Indirecta, mediante datos proporcionados por el fabricante, los cuales en muchos casos no son muy confiables.
- Directa, mediante la instrumentación de una gran cantidad de vehículos con analizadores de gases portables, lo cual hace el método bastante costoso.
- Directa, mediante la construcción de un ciclo típico de conducción, con el objetivo de realizar pruebas experimentales en un dinamómetro de chasis con los diferentes vehículos pertenecientes al parque automotor liviano, para analizar las emisiones producidas mediante un analizador de gases.

La primera opción presenta datos que no son confiables, debido a la variabilidad en los contextos y métodos utilizados para el análisis realizado por los fabricantes, y la segunda opción requiere la instrumentación de una gran cantidad de vehículos en ruta. Por esta razón, la tercera opción corresponde a la opción más confiable y una menor inversión comparativamente.

Los ciclos típicos de conducción representan el comportamiento de los vehículos en una región específica a través del tiempo mediante una gráfica la cual representa velocidad en función del tiempo (Giakoumis, 2016), teniendo en consideración diferentes factores tales como: velocidad, aceleración, consumo de combustible, emisiones, entre otros. Estos han sido utilizados desde hace varios años para realizar las certificaciones de emisiones que un vehículo puede generar., La mayoría de los vehículos son certificados al ser puestos a prueba con ciclos de conducción tales como el NEDC (New European driving cycle), el WLTP (World harmonized light-duty test procedure), el CDCA (Common driving cycle ARTEMIS), entre otros (Giakoumis, 2016). Sin embargo, estos ciclos de conducción estandarizados no necesariamente tienen una relación directa con la realidad (Yang et al., 2019), razón por la cual pueden sobrestimar o subestimar la cantidad de emisiones que se producen en un lugar determinado.

En la ciudad de Bogotá, el tráfico que se presenta tiene un comportamiento complejo que no es posible analizar mediante un ciclo de conducción estandarizado, específicamente en la zona

norte de la ciudad Bogotá y los municipios aledaños que presentan diferentes tipos de comportamientos en el tráfico a lo largo del día y de la semana. Por este motivo, es necesario construir un ciclo típico de conducción específico para la zona de Bogotá región, con el fin de conocer la cantidad de emisiones producidas que se acerquen más a la realidad. Realizando mediciones de CO<sub>2</sub>, CO, NO<sub>x</sub>, O<sub>2</sub>, e hidrocarburos no quemados mediante experimentos en el dinamómetro de chasis, es posible determinar los factores de emisión para estos contaminantes.

Un ciclo de conducción puede ser construido directamente sobre las mediciones realizadas a diferentes tipos de vehículos. Actualmente, es posible construir un ciclo de conducción con base en mediciones realizadas a diferentes tipos de vehículos en rutas establecidas, realizando agrupamiento de datos para generar un ciclo representativo (Huertas et al., 2018; Yang et al., 2019). En este caso, se utilizan métodos de aprendizaje automático, donde los métodos más comunes utilizados son: el método de cadenas de Markov, el método k-means, el método de árbol de decisiones, el método de componentes principales, entre otros (Nyberg et al., 2014; Peng et al., 2020; Yuhui et al., 2019).

En este proyecto, se desarrollan dos de estos métodos, el MWD-CP (*Minimum Weighted Differences of Principal Components*, por sus siglas en inglés) como método de análisis de datos estadístico. Además, se utilizan métodos de aprendizaje automático bajo el enfoque de microciclos, mediante la combinación del algoritmo de aprendizaje no supervisado k-means combinado con la metaheurística ILS-RANDOM (*Iterated Local Search*, por sus siglas en inglés), el algoritmo de cadenas de Markov y una reducción de dimensionalidad mediante un Análisis de Componentes Principales (PCA) (Yuhui et al., 2019), así como la combinación entre el método MWD-CP y el método de aprendizaje automático, para la selección o construcción del CTC (*Typical Driving Cycle*, por sus siglas en inglés o Ciclo típico de conducción). Comparando los resultados, se selecciona el CTC que tenga un menor promedio de error medio respecto a la base de datos de ciclos recolectada, para vehículos del parque automotor liviano en la ciudad de la zona Bogotá región. Luego, se comparará con diferentes ciclos de conducción estandarizados a nivel internacional, con el fin de analizar los factores de emisión generados.

Por lo anteriormente expuesto, la pregunta de investigación planteada es la siguiente: **¿Cómo se puede evaluar de manera estandarizada y eficiente el desempeño ambiental del parque automotor liviano de la zona de Bogotá región**

## **2. Objetivos**

### **2.1. Objetivo general**

Diseñar y validar una metodología para evaluación del desempeño ambiental del parque automotor liviano de la zona de Bogotá región a partir del desarrollo de ciclos típicos de conducción utilizando métodos estadísticos y de aprendizaje automático.

### **2.2. . Objetivos específicos**

- Desarrollar un sistema de medición y adquisición de datos de parámetros de conducción para vehículos ligeros a partir de IVDRs (In - vehicle data recorder).
- Construir un ciclo típico de conducción para vehículos ligeros en la zona Bogotá región mediante métodos estadísticos y de aprendizaje automático.
- Validar la metodología para la evaluación del desempeño medio ambiental de vehículos del parque automotor liviano de Bogotá región a partir de la medición parámetros medio ambientales (eficiencia energética, emisiones GEI y contaminantes producidos), mediante pruebas dinámicas realizadas en el dinamómetro de chasis.
- Determinar factores de emisión generales del parque automotor liviano de Bogotá región, para ser comparados con los factores de emisión registrados en el inventario de emisiones actualmente vigente.

### 3. Marco conceptual

#### 3.1. Generalidades de los ciclos típicos de conducción

Los ciclos típicos de conducción o CTC (*typical driving cycle*, por sus siglas en inglés) empezaron a ser usados a partir de la década de 1960, con el objetivo de cumplir los estándares de emisiones para automóviles, utilizando en su mayoría ciclos estándar. Un ciclo típico de conducción es un ciclo de evaluación transitorio (debido a su variación en el tiempo), el cual consiste en una secuencia de puntos donde la velocidad del auto es definida, y el auto sometido a estudio sigue dicha velocidad, discretizado mediante pasos de tiempo en su mayoría segundos (Giakoumis, 2016).

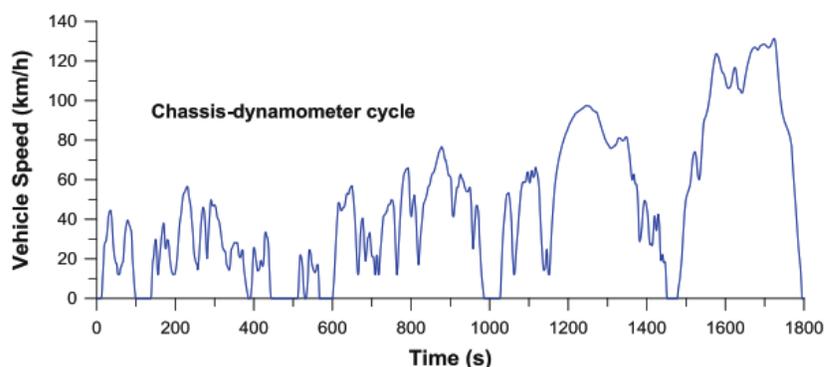


Figura 3.1. Representación general de un ciclo típico de conducción. (Giakoumis, 2016)

En la Figura 3.1 se presenta un ciclo de conducción general, donde se muestra la velocidad en función del tiempo, en muchos casos al igual que la velocidad, se presenta la pendiente de la carretera en función del tiempo; en su mayoría los ciclos de conducción tienen una duración aproximada de 30 minutos, aunque se han desarrollado ciclos típicos de conducción de mayor duración. (Giakoumis, 2016)

##### 3.1.1. Principales Características de los ciclos típicos de conducción

Los ciclos típicos de conducción en su mayoría son utilizados para la evaluación de emisiones, y consumo de combustible (Echeverry Mejía, 2018; Ho et al., 2014; Yang et al., 2019; Yuan et al., 2021) durante el proceso de certificación de los vehículos, así mismo los ciclos de conducción son usados para la comparación de vehículos sometidos al mismo procedimiento, y en algunos casos para el diseño de motores con base en los límites establecidos por los estándares de emisiones,

para inspección y mantenimiento de vehículos y para el modelamiento e investigación de la generación de emisiones y su disminución.

En muchos casos los ciclos de conducción no representan de manera correcta la realidad de los vehículos en un lugar específico, es por esto, que esta característica es muy importante para los CTCs que tienen como objetivo realizar análisis de generación de emisiones.

Por lo anteriormente mencionado, un ciclo de conducción empleado para aprobación, o en este caso para la medición de factores de emisión, debe tener las siguientes características (Giakoumis, 2016).

- Representar de manera correcta el comportamiento del vehículo en la realidad, para lo cual debe considerar diferentes condiciones del vehículo como: encendido en frío, encendido en caliente, marcha en ralentí (revoluciones por minuto a las cuales se encuentra el motor mientras el vehículo se encuentra en reposo, 800 rpm - 1000 rpm), representación de las carreteras del lugar de análisis (carreteras rurales y/o urbanas).
- Debe ser práctico, no debe ser demasiado largo o complicado, con el objetivo de poder ser reproducido de manera sencilla y costo-efectiva, y con resultados reproducibles.
- Debe poder ser completamente controlado.

Las características anteriormente mencionadas son muy importantes ya que los ciclos típicos de conducción son utilizados para la realización de certificaciones, y el análisis de los contaminantes producidos por los automóviles (Giakoumis, 2016). Los ciclos típicos de conducción también son utilizados en el diseño de vehículos, ya que estos permiten conocer el desempeño que tendrá el vehículo en carretera y realizar las respectivas modificaciones; los principales sistemas en los cuales se utilizan los CTCs en el área de diseño, son: Diseño del tren de potencia, diseño de la computadora motor, y diseño del tren de la computadora motor. Es importante mencionar que, los CTCs son utilizados en gran medida dentro la industria automotriz (como se ha visto hasta el momento), en el área energética y legislativa.

Los CTCs se pueden categorizar de la siguiente manera, acorde del tipo de ruta la cual estén recorriendo.

Tabla 3.1. *Categorización de los ciclos típicos de conducción acorde al tipo de ruta (Elaboración propia)*

Categorización de los CTC en base al tipo de ruta	
Tipo de ruta	Características del CTC
Ruta fija	Estos CTCs son construidos para una ruta fija, es decir se realizan diferentes mediciones para la misma ruta siempre entre un punto A y un punto B, así mismo sus aplicaciones son limitadas debido a que se caracterizan por estar determinados por una ruta fija. Estos ciclos pueden ser representados por la distancia recorrida. (Echeverry Mejía, 2018)
Ruta variables o rutas mixtas	A diferencia de los CTCs para rutas fijas, estos ciclos se basan en mediciones realizadas en diferentes rutas, sin ningún orden establecido, claramente este ciclo no puede ser representado por la distancia recorrida, ya que la distancia recorrida no es la misma en todos los trayectos, sin embargo se puede caracterizar por diferentes parámetros, como por ejemplo, velocidad máxima alcanzada, velocidad promedio, consumo de combustible etc. Estos CTC tienen bastantes aplicaciones dentro de la industria automotriz, así como en el área energética, ejemplos de estos CTCs, son algunos de los ciclos estándar. Este proyecto se enmarca en este tipo de CTCs. (Echeverry Mejía, 2018; Peng et al., 2020)

### 3.1.2. Desarrollo de un ciclo típico de conducción

El desarrollo de un ciclo típico de conducción en la actualidad sigue una metodología establecida, acorde a lo presentado en la literatura, donde se establece una de las aplicaciones de *Big data* debido a la cantidad de información que debe ser procesada, y también una aplicación de los métodos de aprendizaje automático. La metodología general expuesta en la literatura (Giakoumis, 2016; Wu et al., 2020; Yuhui et al., 2019), para el desarrollo de un ciclo típico de conducción es la siguiente:

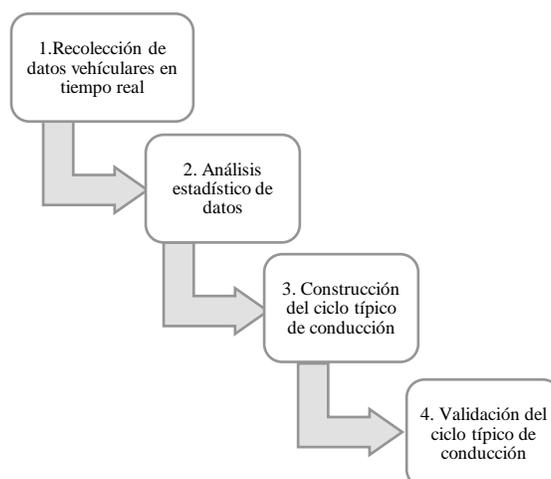


Figura 3.2. Metodología general para la construcción de un ciclo típico de conducción. (Giakoumis, 2016)

En la Figura 3.2 se presentan cuatro pasos para el desarrollo de un ciclo típico de conducción, claramente en algunos casos se presentan pequeñas diferencias en el orden de los pasos, o en los métodos utilizados para el desarrollo del ciclo típico de conducción, pero en general la metodología actual es la anteriormente presentada.

### **3.1.3. Ciclos típicos de conducción estándar**

Desde la década de 1950 se han desarrollado diferentes tipos de CTCs, en sus comienzos se realizaron CTCs basados en modos (California 7-mode), en los cuales se establecían velocidades para ciertos rangos de tiempo, mas no tenían un comportamiento directamente relacionado con la realidad; desde la década de 1970 se empezaron a desarrollar CTCs transitorios, como lo fue el CTC FTP-72 (Federal Test Proccesure – 1972, por sus siglas en inglés), conocido actualmente FTP-75 (Federal Test Proccedure - 1975, por sus siglas en inglés) debido a algunas variaciones realizadas, este CTC fue desarrollado en Norteamérica.

Alrededor del mundo se han construido diferentes CTCs estándar, para diferentes aplicaciones, con diferentes tipos de vehículos incluyendo, vehículos ligeros y/o pesados, vehículos de combustión interna con motores de gasolina o diésel, vehículos híbridos y vehículos eléctricos. Este proyecto se establecerá en el marco del análisis de vehículos ligeros de combustión interna con motores de gasolina.

A continuación, se presentan diferentes CTCs estándar elaborados en diferentes partes del Mundo, para diferentes aplicaciones.

Tabla 3.2. Características principales de los ciclos típicos de conducción estándar. (Elaboración propia)

Ciclos típicos de conducción estándar					
CTC	Ciudad/País	Fecha	Tipo de ruta	Método de construcción	Características
California 7 - mode	California - Estados Unidos	1960	Ruta fija	Basado en modos, combinación de frecuencias de conducción, de aceleración y desaceleración.	Duración 137 segundos, velocidad máxima 80 km/h, velocidad promedio 41.8 km/h. Utilizado como ciclo para la certificación de vehículos entre los años 1966 - 1972.(Clarkson & Middleton, 1962; Giakoumis, 2016; Hass & Brubacher, 1962)
FTP - 72	California - Estados Unidos	1972	Ruta variable	Mediante métodos estadísticos se selecciona el ciclo que representa la mayor cantidad de muestras de datos tomadas	Duración 1372 segundos, dividido en dos partes, encendido en frío y periodo de estabilización del motor, velocidad máxima 91,2 km/h, distancia recorrida 12 km. (Giakoumis, 2016)
FTP - 75	California - Estados Unidos	1975	Ruta variable	Mediante métodos estadísticos se selecciona el ciclo que representa la mayor cantidad de muestras de datos tomadas	Variación del ciclo FTP-72, desarrollado en el año 1975, primer ciclo en ser utilizado para la certificación de vehículos ligeros incluyendo vehículos a base de Diesel. Este ciclo tiene una tercera fase con encendido en caliente, después de una parada de 10 minutos, se recolectan las emisiones durante todo el ciclo a diferencia del FTP - 72 y el NEDC, proveyendo de una mayor exactitud en la medición de emisiones. Duración 2477 segundos. Debido al uso equipos desactualizados de la época de los 60s, este ciclo no tiene concordancia con la realidad. (Division & Agency, 1993; Lin & Niemeier, 2002)
NEDC	Unión Europea	1970	Ruta fija	Basado en modos, combinación de frecuencias de conducción, de aceleración y desaceleración.	Duración del ciclo 1180 segundos, velocidad máxima 120 km/h, distancia recorrida 10.6 km. Este ciclo fue utilizado desde el año 1970, teniendo en consideración los ciclos ECE y EUDC, para la certificación de vehículos respecto a emisiones. Este ciclo no representa una relación directa con la realidad debido a sus bajas

					aceleraciones, y velocidades estacionarias. (Giakoumis, 2016)
WLTC	Internacional	2017	Ruta variable	Mediante métodos estadísticos y a partir de una gran base de datos recolectados en diferentes países alrededor del mundo, se genera y valida el CTC.	Ciclo típico de conducción desarrollado dentro del WLTP (Worldwide Harmonized light duty test procedure), con el objetivo de tener un CTC aplicable de manera internacional. Duración del ciclo 1800 segundos en una de sus tres versiones, velocidad máxima 131.2 km/h, este ciclo se encuentra dividido en tres clases, las cuales son aplicables a diferentes tipos de vehículos. CTC utilizados para regulación de vehículos en Europa desde 2017. (Pavlovic et al., 2018; Tutuianu et al., 2015).
CDCA	Unión Europea	-	Ruta variable	Mediante métodos estadísticos se selecciona el ciclo que representa la mayor cantidad de muestras de datos tomadas	El CDCA (Common driving cycle ARTEMIS), si bien no es un CTC legislado es comúnmente utilizado en países de la unión europea para la medición de contaminantes producidos por vehículos. Este ciclo corresponde a un ciclo transitorio, el cual tiene una duración de 3143 segundos, en su versión completa, y una velocidad máxima de 150.4 km/h. (Giakoumis, 2016)

Como se presenta en la Tabla 3.2, los CTCs estándar tienen características muy similares entre sí, estos ciclos son utilizados para la certificación de vehículos o medición de contaminantes. Los CTCs presentados no corresponden a todos los ciclos estándar, sino a los más conocidos mundialmente, existen diferentes ciclos específicos, realizados en países como India, Japón, China, entre otros.

#### **3.1.4. Comparación de los ciclos típicos de conducción construidos y los ciclos típicos de conducción estándar**

Para la zona Bogotá región no se cuenta con un CTC para vehículos livianos a gasolina con rutas variables. Existen diferentes tipos de ciclos de conducción estándar utilizados para diferentes propósitos. Pero, estos ciclos si bien son realizados mediante métodos estadísticos muy fiables, no representan de manera correcta la realidad de los vehículos en una región específica, al tener

diferentes tipos de carreteras, comportamientos de tráfico complejos y diferentes aspectos que varían dependiendo del lugar donde se aplica el CTC. Un CTC estándar puede estar bastante alejado de representar la realidad de un lugar diferente al que se construyó, por esta razón en diferentes lugares del mundo se han desarrollado CTCs específicos (Fotouhi & Montazeri-Gh, 2013; Huertas et al., 2018; Yang et al., 2019; Yuan et al., 2021) .

Estos ciclos si bien no son estándar, representan de una mejor manera la realidad de los vehículos a analizar en una zona determinada. Una comparación entre los ciclos de conducción estándar y los construidos a partir de diferentes métodos, se presenta en la Tabla 3.3.

Tabla 3.3. *Comparación de los ciclos típicos de conducción algunos de CTCs construidos (Elaboración propia)*

<b>Comparación entre los CTCs estándar y los CTCs construidos</b>					
<b>CTC</b>	<b>Tipo</b>	<b>Ciudad/ País - Fecha</b>	<b>Tipo de ruta</b>	<b>Método de construcción</b>	<b>Características</b>
California 7 - mode	Estándar	California - Estados Unidos - 1960	Ruta fija	Método de modos	Utilizado para la certificación de vehículos
FTP - 72	Estándar	California - Estados Unidos - 1972	Ruta variable	Métodos estadísticos	Utilizado para la certificación de vehículos
FTP - 75	Estándar	California - Estados Unidos - 1975	Ruta variable	Métodos estadísticos	Utilizado para la certificación de vehículos
NEDC	Estándar	Unión Europea - 1970	Ruta fija	Método de modos	Utilizado para la certificación de vehículos
WLTC	Estándar	Internacional - 2017	Ruta variable	Métodos estadísticos	Utilizado para la certificación de vehículos
CADC	Estándar	Unión Europea	Ruta variable	Métodos estadísticos	Utilizado para la medición de contaminantes
Nanjing, China	Construido	Nanjing. China - 2019	Ruta variable	Métodos estadísticos y de aprendizaje automático, específicamente método de análisis de componentes principales y algoritmo k-means	Utilizado para la caracterización de las emisiones de la ciudad, basándose en las implicaciones de la velocidad y aceleración del ciclo. 18 parámetros característicos seleccionados. (Yang et al., 2019)

Flota DNA	Construido	Estados Unidos -2021	Ruta Variable	Métodos estadísticos y de aprendizaje automático, específicamente método árbol de decisiones	CTC construido a partir de mediciones realizadas a la flota de vehículos pesados DNA en Estados Unidos, se realizan 5 CTCs para cinco tipos de vehículos basados en diferentes parámetros de análisis, a los cuales se les da un peso específico acorde al método de árbol de decisiones. Finalmente se genera un ciclo representativo y cuatro ciclos generales, para la promoción del diseño de un tren de potencia. (Zhang et al., 2021)
Toluca México	Construido	Toluca - México - 2017	Ruta Fija	Métodos estadísticos - método de diferencias ponderadas.	Diseño del tren de potencia. (Huertas, 2017)

Como se presenta en la Tabla 3.3; **Error! No se encuentra el origen de la referencia.**, se realiza una comparación entre los ciclos estándar y los construidos por investigadores, o bien ciclos locales. Estos ciclos representan de una mejor manera el comportamiento de una zona específica, basándose en parámetros del vehículo y de carretera los cuales pueden llegar a tener una mayor incidencia en el ciclo, así mismo se emplean métodos estadísticos avanzados y métodos de aprendizaje automático, para obtener un CTC coherente con el comportamiento de los vehículos en el lugar de estudio.

Claramente existe una mayor cantidad de ciclos construidos por investigadores (Chauhan et al., 2020; Fotouhi & Montazeri-Gh, 2013; Qiu et al., 2018; Yuan et al., 2021), sin embargo, los ciclos presentados corresponden a una revisión de la literatura de los años más recientes. Estos ciclos construidos, ayudan a los investigadores a tener una mayor aproximación a la realidad, respecto a

la medición de emisiones, diseño de vehículos, y las aplicaciones de los ciclos anteriormente mencionados.

### **3.2. Adquisición de datos en tiempo real de automóviles**

Como se presentó en la Figura 3.2, el primer paso para la construcción de un CTC es la recolección de información, esto claramente después de haber seleccionado el tipo de ruta del CTC y las técnicas de construcción del CTC. La recolección de información del vehículo debe ser realizada en tiempo real, es decir se debe tomar la información en el mismo instante que el vehículo realiza el trayecto. Para esto se debe tener una comunicación directa con el vehículo, mediante la computadora motor (ECU – *Engine Control Unit*, por sus siglas en ingles). De la literatura analizada los investigadores utilizan diferentes dispositivos para la medición de los parámetros del vehículo en carretera tales como, PEMs (*Portable Emissions Measurment System*, por sus siglas en ingles) el cual recolecta información de emisiones y parámetros del vehículo en tiempo real (Yang et al., 2019), dispositivos GPS, en otros casos se realiza comunicación directa con la ECU mediante el protocolo de comunicación CAN (*Controller Area Network*), y su sistema de diagnóstico OBDII (*On Board Diagnostic version 2.0*, por sus siglas en ingles), y en algunos casos se utilizan ambos tanto el dispositivo GPS, como la comunicación mediante protocolo CAN (Kharrazi et al., 2019; Zhang et al., 2021),

En el marco del proyecto de investigación, se realizará la recolección de datos mediante la comunicación con la ECU, a través un dispositivo el cual recolecta los datos necesarios en este caso un IVDR (Echeverry-Mejía et al., 2022).

#### **3.2.1. Protocolos de comunicación**

Para la adquisición de datos en tiempo real de los vehículos, es necesario solicitar y leer estos datos, mediante comunicación a través del protocolo CAN y el sistema de diagnóstico OBDII.

**Protocolo CAN (*Controller Area Network*).** El protocolo CAN, es el protocolo de comunicación actualmente utilizado para la comunicación de datos entre las computadoras de un vehículo. Este protocolo fue diseñado por la compañía Robert Bosch GmbH en la década de los 80's, para el año 1993 este protocolo se estableció como el protocolo estándar (International Organization for Standardization (ISO), 2006). Este protocolo de comunicación permite una

transferencia de datos de hasta 1 Mbps, con tramas de 8 bits de información, y 11 bits para las respectivas direcciones (IDs). Físicamente el protocolo CAN, consta de dos cables los cuales son denominados CAN HIGH y CAN LOW, estos cables manejan la misma señal, pero en el caso del CAN LOW de manera invertida, manejado diferentes rangos de voltajes entre sí, esto se hace con el objetivo de eliminar interferencias en la señal que se recibe.

El protocolo CAN presenta algunas ventajas frente a otros protocolos de comunicación, entre las cuales se encuentran:

- Reduce la cantidad de cables para la comunicación, ya que permite la comunicación mediante dos cables.
- Recibe información de 8 bytes por cada trama de datos, tiene la capacidad de que en cada nodo (computadora de alguno de los sistemas del vehículo), los mensajes enviados son recibidos por todos los nodos los cuales tienen la capacidad de aceptar o no dicho mensaje mediante máscaras.
- Posee métodos sofisticados de detección de errores y retransmisión de datos erróneos, tales como: CRC (Cyclic redundancy check, por sus siglas en ingles), Frame check y Acknowledge error. (Echeverry Mejía, 2018; International Organization for Standardization (ISO), 2006).

Los datos que se envían mediante el protocolo CAN tienen una estructura específica, la cual se presenta a continuación.

<b>SOF</b> 1 Bit	<b>ID</b> 11 Bits	<b>RTR</b> 1 Bit	<b>IDE</b> 1 Bit	<b>DLC</b> 4 Bits	<b>Campo de datos</b> 8 Bytes	<b>CRC</b> 15 Bits	<b>ACK</b> 3 Bits	<b>EOF</b> 7 Bits
---------------------	----------------------	---------------------	---------------------	----------------------	----------------------------------	-----------------------	----------------------	----------------------

*Figura 3.3.* Estructura general para la trama de datos de un mensaje CAN estándar. (Echeverry Mejía, 2018; International Organization for Standardization (ISO), 2006)

Cada una de las partes tiene una importancia y significado (por esto la asignación de la cantidad de bits de cada parte del mensaje) importante dentro de un mensaje de CAN, la definición de cada una de estas partes se presenta en la Tabla 3.4 (Echeverry Mejía, 2018; International Organization for Standardization (ISO), 2006).

Tabla 3.4. Definición de cada una de las partes de la trama de datos del protocolo CAN. (Adaptado Echeverry Mejía, 2018)

Parte del mensaje	Definición
SOF (Start of frame)	Define el comienzo de la trama de mensajes
ID (Identifier)	Esta parte de la trama de datos corresponde al identificador o bien el nombre del mensaje, así mismo establece el arbitraje del mensaje y la prioridad de este
RTR (Remote Transmission Request)	Permite diferenciar la trama únicamente de datos, de una trama de solicitud remota
IDE (Identifier extension)	Permite la diferenciación entre tramas de datos extendidas (29 Bits) y estándar (11 Bits)
DLC (Data length code)	Corresponde a la longitud del campo de datos enviados o recibidos, en si a los bits que se están utilizando en el mensaje enviado o recibido
Campo de datos	Corresponde al campo de datos de la trama correspondiente a un mensaje enviado o recibido, contiene la información requerida, o solicitada.
CRC (Cyclic Redundancy Check)	Sistema de redundancia cíclica, en este caso se asignan 15 bits, para que, a través de un polinomio generado, este pueda coincidir con el número de recepción en el nodo de llegada
ACK (Acknowledgement)	Esta parte de la trama de datos permite conocer si el mensaje llegó exitosamente a los nodos. Cada nodo al recibir el mensaje envía un bit, esta parte de la trama, en caso de no recibirse el nodo transmisor lo envía de nuevo.
EOF (End of frame)	Indica el final de la trama de datos.

En la Tabla 3.4, y la Figura 3.3 se presentan la estructura y la definición de cada una de las partes de la trama de datos del protocolo CAN, cada una de las computadoras de un vehículo puede enviar mensajes entre ellas, y cada una tiene un ID específico el cual permite que cada computadora acepte los mensajes que requiere para su operación. Dentro del protocolo de comunicación CAN circulan diferentes tipos de mensajes, para el desarrollo de este proyecto se analizan dos de ellos.

- *Mensaje de diagnóstico OBDII*. Este tipo de mensajes maneja una estandarización, mediante la cual se pueden solicitar datos y recibir datos de manera estándar para todos los vehículos.
- *Mensajes CAN*. Estos mensajes no son estándar y varían acorde a los IDs que se asignen a cada una de las computadoras de un vehículo, esto depende de cada casa automotriz, así que para conocer estos mensajes es necesario tener la información directamente de la compañía que programó el sistema CAN del vehículo.

**Sistema de diagnóstico OBD II (On Board Diagnostic)**. El sistema de diagnóstico OBDII, es un sistema estandarizado para la mayoría de los vehículos desde el año 2001, y desarrollado por la sociedad de ingenieros automotrices descrito en la norma SAE J1979 y SAE J1962 (Society of Automotive Engineers, 2007, 2016), este sistema de diagnóstico se ejecuta sobre el protocolo de comunicación CAN.

El sistema de diagnóstico tiene 10 modos de operación, entre los cuales se encuentra la adquisición de datos del vehículo en tiempo real (modo 1), la lectura de DTCs (*Diagnostic trouble codes*, por sus siglas en inglés) (modo 4), y el borrado de DTCs (modo 3). Acorde al enfoque del proyecto el modo utilizado corresponde al modo 1, adquisición de datos en tiempo real del vehículo.

Para esto el sistema de diagnóstico OBDII en el modo uno de operación contiene diferentes PIDs (*Parameters identifiers*, por sus siglas en inglés), en este caso estos identificadores se emplean como direcciones de memoria que almacenan parámetros del vehículo en tiempo real, en la Tabla 3.5 se presentan algunos de los PIDs más utilizados los cuales se encuentran en sistema Hexadecimal.

Tabla 3.5. *PIDs comúnmente utilizados* (Echeverry Mejía, 2018)

PIDs Hex.	Bytes	Descripción	min.	máx.	res.	Unid.	Fórmula
0x04	1	Carga calculada del motor	0	100	1	%	A/2.55
0x05	1	Temperatura refrigerante del motor	-40	215	1	°C	A-40
0x0B	1	Presión múltiple de admisión	0	255	1	KPa	A
0x0C	2	RPM del motor	0	16383.75	0.25	RPM	(256A+B)/4
0x0D	1	Velocidad lineal del vehículo	0	255	1	Km/h	A
0x11	1	Posición de la válvula mariposa	0	100	1	%	A/2.55
0x49	1	Posición del pedal acelerador	0	100	1	%	A/2.55

Ya que el sistema de diagnóstico OBDII se ejecuta en el protocolo de comunicación CAN, este tiene un ID específico para la solicitud de información, y un ID específico para la recepción de la información solicitada, estos IDs corresponden a 7DF (2015) y 7E8 (2024) en hexadecimal respectivamente. Teniendo en consideración solamente el ID, el DLC y el campo de datos, en la Tabla 3.6 se presenta un ejemplo de un mensaje de OBDII, en el caso de solicitar las revoluciones por minuto del motor (PID 0x0C).

Tabla 3.6. *Descripción de un mensaje de OBDII (Elaboración propia).*

Descripción	ID Hex	DLC	Campo de datos							
Mensaje de solicitud OBDII	7DF	08	02	01	0x0C	00	00	00	00	00
Mensaje recibido OBDII	7E8	08	04	41	0x0C	09	DF	00	00	00
			<b>Bit 1</b>	<b>Bit 2</b>	<b>Bit 3</b>	<b>Bit 4</b>	<b>Bit 5</b>	<b>Bit 6</b>	<b>Bit 7</b>	<b>Bit 8</b>
			<b>DLC OBD II</b>	<b>MODO</b>	<b>PID</b>	<b>Valor A</b>	<b>Valor B</b>			

En el caso presentado en la Tabla 3.6 se presenta los PIDs de petición o escritura, y de recepción o lectura en el primer caso el bit 1, representa el DLC de OBDII, el cual, para el mensaje de petición, solamente es de dos, ya que solo se utilizan dos bits, el bit 2 representa el modo, el bit 3 el PID. Para el caso del mensaje de lectura, se tiene la misma estructura de datos con la variación del bit 1, el cual tiene un valor de cuatro, ya que se utilizan cuatro bits, y los bits 4 y 5, los cuales representan los valores A y B respectivamente, para los cuales acorde a la ecuación de la tabla 4, corresponderían a 631.75 RPM.

### 3.2.2. Desarrollo de un IVDR (In-vehicle data recorder)

Un IVDR (In- Vehicle Data Recorder), es un dispositivo el cual conectado al puerto SAE J1962, puede obtener mediante el sistema diagnóstico OBDII información en tiempo real de un vehículo. Este dispositivo debe ser diseñado de manera tal que lea y almacene la información requerida. Los dispositivos IVDR han sido desarrollados e implementados en diferentes aspectos acorde a la literatura (Echeverry-Mejía et al., 2022; Gitelman et al., 2018; Toledo et al., 2008), especialmente para la recolección de datos. Existen diferentes tipos de IVDRs en el mercado, pero su costo es demasiado alto, por esta razón se propone el diseño y construcción de un IVDR para las mediciones necesarias.

**Diseño de un IVDR.** Acorde a lo expuesto en trabajos previos (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022), un IVDR debe cumplir con las siguientes características:

Tabla 3.7. *Características generales de un IVDR (Adaptado de Echeverry Mejía, 2018)*

<b>Característica/parámetro de diseño</b>	<b>Descripción</b>
CAN/OBDII	Debe tener comunicación con el protocolo CAN y su sistema de diagnóstico OBDII
GPS	Debe contar con un GPS integrado
Perdida de información	Debe ser capaz de leer la mayor cantidad de información, teniendo la mínima pérdida de datos
Validación con la industria	Al menos uno de los componentes (microcontrolador), debe estar validado en la industria automotriz
Grabar memoria	Debe tener la capacidad de almacenar la información recolectada en una memoria SD
Grabar en la nube (opcional)	Debe tener la capacidad de guardar los registros en la nube (para el objetivo del proyecto esta característica no es necesaria)
Portable	Debe ser portable, y lo más pequeño posible
Programable y con bajo costo de software	Debe poder ser programable para que tenga la capacidad de añadir funciones, y tener un lenguaje de Código libre
Costo	Debe ser un dispositivo que se pueda replicar fácilmente, con un bajo costo.

En la Tabla 3.7 se presentan las características necesarias para la construcción de un IVDR, acorde a esto se debe plantear el diseño del hardware y el software del dispositivo.

*Diseño de Hardware.* Teniendo en consideración la información presentada en la Tabla 3.6, es necesario realizar un diseño del hardware en sí, con el objetivo de que cumpla con todos los parámetros de diseño, y el dispositivo pueda cumplir con la función establecida. Para esto, es necesario tener el conocimiento de los componentes necesarios para el desarrollo, y el respectivo esquemático de su construcción.

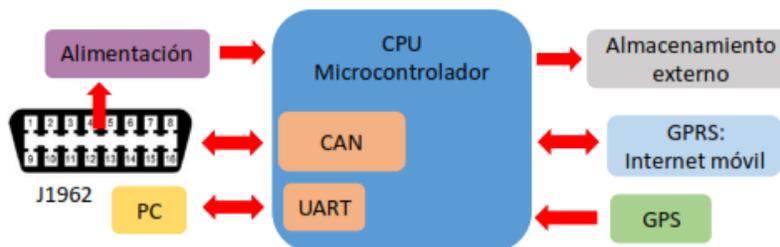


Figura 3.4. Diagrama modular del dispositivo IVDR (Echeverry Mejía, 2018)

En la Figura 3.4 se presenta el esquema modular de un IVDR, en este caso el IVDR construido por (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022) en el IVDR se deben encontrar los siguientes componentes de Hardware.

- CPU – MICROCONTROLADOR: Consiste en la unidad programable, la cual permite la comunicación entre el vehículo y el PC, debe contar con CAN y UART (*Universal Asynchronous Receiver – Transmitter*, por sus siglas en inglés) (Protocolo de comunicación serial con el PC)
- Módulo de almacenamiento externo: Permite guardar la información recolectada en una memoria externa
- Modulo GPRS (opcional): Provee de señal de internet al dispositivo IVDR. Este componente no es necesario para el proyecto de investigación, ya que no se requiere almacenar datos en internet.
- Modulo GPS: Este módulo permite la adquisición de posicionamiento global en tiempo real
- Fuente(s) de alimentación: Suple de energía eléctrica al dispositivo tomándola del alternador del vehículo
- Conector serial: Conector el cual permite transferir los datos del vehículo al dispositivo

*Diseño de software.* El diseño de software de IVDR debe estar en marcado dentro de la programación del microcontrolador a utilizar, y para esto acorde a los IVDRs desarrollados, y específicamente el IVDR desarrollado por (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022), la programación del IVDR debe estar basada en interrupciones, lo cual determina que el programa está controlado por eventos externos o las mismas interrupciones. Estas interrupciones se establecen cada cierto tiempo, cuando se solicita información, o se corren rutinas para el correcto

funcionamiento del sistema. Las rutinas de interrupción son las siguientes (Echeverry Mejía, 2018):

- Interrupción de recepción por CAN
- Interrupciones PIT (*Periodic Interrupt Timer*, por sus siglas en inglés)
- Interrupciones seriales UART
- Interrupción IRQ (External interrupt request)
- Rutina del programa principal
  - Rutinas para la verificación de errores
  - Rutinas del control del módulo GPRS
  - Rutina de finalización segura
  - Rutinas simples para la interfaz de usuario.

Este tipo de programación fue validada, pero cabe aclarar que el dispositivo IVDR está expuesto a actualizaciones y mejoras tanto de software como de hardware.

### **3.2.3. Análisis previo de los datos adquiridos por un IVDR**

A partir de los datos adquiridos por el IVDR de los vehículos analizados es necesario realizar un preprocesamiento de datos, en primer lugar una limpieza de datos, ya que es posible debido a la comunicación con el puerto serial que se pierda cierta información durante algunos periodos de tiempo, o se presenten datos atípicos debido a interferencias, por lo cual se deberán reemplazar dichos datos faltantes, mediante diferentes algoritmos de posprocesamiento. Así mismo, se debe determinar si el periodo en el que dichos datos faltantes es mayor varios segundos, esos datos deben ser excluidos de la muestra, entre otros procesos de limpieza de datos (Duran & Earleywine, 2012; Nguyen et al., 2020) .

### **3.3. Generalidades sobre los métodos de aprendizaje automático y estadísticos para la construcción de los ciclos típicos de conducción**

Una de las tecnologías que ha tenido gran importancia en los últimos años, ha sido la inteligencia artificial, así como las ramas que la componen como lo son, el aprendizaje automático, el aprendizaje profundo, el procesamiento del lenguaje natural, entre otros. La construcción de un CTC es una aplicación clara aplicación del aprendizaje automático, ya que mediante los métodos

propuestos por esta área de la inteligencia artificial es posible determinar ciclos típicos de conducción los cuales representen de una mejor manera el comportamiento de los vehículos en la realidad para zonas específicas, (Yang et al., 2019; Yuan et al., 2021). }

### 3.3.1. Métodos generales de aprendizaje automático y estadísticos para la construcción de ciclos típicos de conducción

Para la construcción de un CTC es posible utilizar diferentes métodos, como se presentó en la sección 3.1.3 en el caso los ciclos estándar utilizando métodos modales o métodos estadísticos, como el uso de la distribución de probabilidad chi cuadrado, y en los ciclos más recientes construidos por investigadores, se utilizan métodos más avanzados como los métodos de aprendizaje automático, con el fin de realizar una mejor manejo de los datos y construir un CTC que se acople más a la realidad. A continuación, se presentan algunos métodos estadísticos y métodos de aprendizaje automático para la construcción de CTCs.

**Cadenas de Márkov.** El método de cadenas de Márkov es un método estocástico aplicado al aprendizaje automático, este método fue elaborado por Andréi Márkov, con base en el análisis de la ley de los grandes números propuesta por Jacob Bernoulli, en este caso mediante este método se establece que la ley de los grandes números no solo es aplicable para eventos de carácter independiente, sino también para eventos de carácter dependiente, es decir el estado actual de un evento contiene la información necesaria para conocer la probabilidad del siguiente estado.

Este método estocástico es utilizado en los métodos de aprendizaje de automático, y también es utilizado como enfoque para la construcción de CTCs con base en los datos recolectados de las rutas realizadas por los vehículos. Como se mencionó anteriormente, las cadenas de Márkov establecen que la probabilidad del estado actual de un sistema depende del estado anterior, o la probabilidad del estado siguiente de un sistema, depende del estado actual. Como se muestra en la ecuación *Ecuación 3.1* (Nyberg et al., 2014) .

$$P(X_{n+1} = x_{n+1} | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n) \quad \text{Ecuación 3.1}$$

Para la construcción de los CTCs y acorde a la literatura, el estado actual es definido como  $x_n = (v_n, a_n)$ , donde el estado actual contiene la información de velocidad y aceleración del vehículo, por lo consiguiente la información necesaria para conocer la probabilidad del siguiente estado ( $x_{n+1}$ ) tanto de aceleración ( $a_{n+1}$ ) como velocidad ( $v_{n+1}$ ), las probabilidades de cada uno de los estados se guardan y analizan en una matriz de transición de probabilidad, la cual contiene las probabilidades de cada uno de los estados de transición y se describe mediante la ecuación Ecuación 3.2 (Nyberg et al., 2014).

$$p_{ij} = P(X_{n+1} = j | X_n = i) \quad \text{Ecuación 3.2}$$

Donde  $i$  corresponde al estado actual,  $j$  al estado siguiente, y  $p_{ij}$  corresponde a la probabilidad del siguiente estado con base en el estado actual. Las probabilidades  $p_{ij}$  se determinan mediante el análisis de las transiciones realizadas entre los diferentes estados de todos los ciclos recolectados, estas probabilidades son almacenadas en una matriz de transición de probabilidad o TPM (*Transition probability matrix*, por sus siglas en inglés), realizando una correcta discretización de la frecuencia de análisis tanto de la velocidad como de la aceleración 1 km/h y 0.2 m/s<sup>2</sup> respectivamente acorde a la literatura.

La construcción de la TPM (Onnegren, 2013), se determina mediante el análisis de la transición de cada uno de los estados de los diferentes ciclos recolectados, para que una vez incluidas todas las transiciones de todos los ciclos se conozca la cantidad de veces que ocurrió una transición de estado y así conocer las respectivas probabilidades.

A partir de esto se selecciona un estado inicial (velocidad y aceleración igual a cero acorde a la literatura), en el cual se unen los siguientes estados de velocidad y aceleración basado en la TPM, hasta que en un tiempo prudencial se llegue a un estado de velocidad cero. Cada uno de los valores del vector de velocidad es analizado, con el objetivo de conocer si cumple con los requisitos de CTC requerido, de no ser así se repite el ciclo hasta que se cumpla con los requisitos, con base en variables representativas (parámetros característicos en el método MWD) establecidas las cuales tengan una incidencia específica en los ciclos medidos, estas variables acorde a la literatura analizada se determinan a través de diferentes métodos (Nyberg et al., 2014; Onnegren, 2013).

Ya que el método de las cadenas de Márkov establece que la probabilidad es estable cuando se tiene una cantidad de estados la cual tiende a infinito, es de mencionar que mediante un análisis de valores y vectores propios se puede calcular el vector de probabilidades estable para cada uno de los estados del CTC (Shi et al., 2016).

***K-means y el método de Análisis de Componentes Principales.*** El método K-means es un método de aprendizaje automático no supervisado, el cual está basado en la segmentación de los ciclos recolectados mediante microciclos para después con base en diferentes parámetros realizar la unión de los microciclos que más se asemejen a la muestra total.

Un microciclo, se define un grupo de datos entre dos paradas consecutivas, es decir un grupo de datos entre los cuales la velocidad inicia en 0 m/s y finaliza en 0 m/s incluyendo su periodo de ralentí inicial, teniendo así un periodo de ralentí en cada microciclo (Desineedi et al., 2020), pero esta no es la única definición de un microciclo, en la literatura se establecen diferentes definiciones para este concepto (Nouri & Morency, 2017; Roy & Morency, 2020), sin embargo la definición anteriormente presentada es la utilizada en el presente trabajo.

Una vez los ciclos son segmentados en microciclos, mediante las diferentes definiciones, tales como, parada entre dos puntos, distancias de 250 m, o modos de conducción, se procede a realizar la categorización en grupos o clusters (nombre comúnmente utilizado para la segmentación de información en grupos en los métodos de aprendizaje automático), con base en las características definidas las cuales tengan en común cada uno de los microciclos segmentados. Las características con las cuales el algoritmo k-means clasifica los microciclos, se han establecido a partir de diferentes enfoques, pero con el objetivo de tener una caracterización, la cual considere la mayoría de los parámetros necesarios para la clasificación se utiliza el método de análisis de componentes principales o PCA (*Principal component análisis*, por sus siglas en inglés). El cual permite tener una gran cantidad de parámetros que caracterizan los microciclos, los cuales estén relacionados entre sí, reduciéndolos a parámetros que no se relacionan entre sí, pero que reflejan en un alto porcentaje la variabilidad de todos parámetros seleccionados. En la literatura (Peng et al., 2020; Yuhui et al., 2019), diferentes parámetros característicos los cuales son reducidos dimensionalmente en tres componentes principales, calculando la contribución de cada uno de estos componentes, donde establece el número máximo de componentes, para los cuales la suma porcentual de sus contribuciones sea mayor a 98%, es decir la cantidad de componentes principales

los cuales en su contribución conjunta representen más del 98% de la variabilidad de los parámetros seleccionados, esto implica directamente una reducción de dimensionalidad.

Teniendo como base los parámetros encontrados a partir del método PCA, se procede a clasificar en grupos cada uno de los microciclos, mediante el método k-means. Como se explicó este es un método de aprendizaje automático no supervisado, cuyo nombre consiste en K, el cual hace referencia a la cantidad de clústeres o grupos en los cuales se segmenta la información en este caso los diferentes microciclos, means hace referencia a que se toma un punto de referencia de cada clúster (centroide) con el objetivo de calcular la distancia euclidiana entre los puntos cercanos al punto de referencia, es decir con base en los parámetros se calcula la distancia a la cual se encuentra cada microciclo del centroide, la distancia mínima entre el centroide y los microciclos, establecerá a que clúster hará parte cada microciclo. La Ecuación 3.3 hace referencia a la función de costo a minimizar en el algoritmo k-means, la cual corresponde al promedio de las distancias de cada microciclo a su respectivo centroide.

$$J(c^{(1)}, \dots, c^{(m)}, \mu_1, \dots, \mu_K) = \frac{1}{m} \sum_{i=1}^m \|x^{(i)} - \mu_{c^{(i)}}\|^2 \quad \text{Ecuación 3.3}$$

Donde  $x^{(i)}$ , corresponde a la norma del microciclo  $i$ , y  $\mu_{c^{(i)}}$  corresponde a la norma del centroide al cual el microciclo fue asignado, con  $m$  cantidad de microciclos y  $K$  cantidad de clústeres. A partir del PCA es posible conocer la dimensión del espacio de análisis. La cantidad de clústeres en los cuales se segmentará la información es necesario conocerlo, en la literatura se aplica la ecuación de Silhouette (Yuhui et al., 2019), aunque en otros estudios se considera la cantidad de clústeres revisados en la literatura los cuales varían entre 2 a 7 (Roy & Morency, 2020), de igual manera otros métodos como el método el codo se establecen para conocer la cantidad de clústeres necesaria donde se establece convergencia en la solución (Desineedi et al., 2020).

Posterior a la clasificación de los microciclos en clústeres, y con base en el centroide de referencia de cada clúster, se seleccionan diez microciclos candidatos desde el más cercano al centro del clúster (microciclo de referencia), hasta el más lejano, si un clúster tiene menos de diez microciclos todos estos microciclos son propuestos como candidatos, para que de manera aleatoria se genere un ciclo de conducción general candidato a ser el CTC, el cual es analizado con base en diferentes parámetros de evaluación seleccionados de acuerdo con el objetivo de investigación, o

a una técnica de evaluación la cual también tiene en consideración los parámetros de interés, acorde a la metodología utilizada por Yuhui et al. (Yuhui et al., 2019).

En la literatura se evidencia el uso de cadenas de Márkov junto con el método k-means, para la selección de cada uno de los microciclos en la construcción del ciclo candidato, para esto es necesario construir la TPM, de los estados de transición de los microciclos para generar el nuevo ciclo de conducción (Desinedi et al., 2020; Peng et al., 2020). Un método común y parecido al método MWC-CP, para determinar el ciclo que representa de mejor manera los ciclos recolectados, es el método del error medio, el cual analiza el error medio del ciclo construido mediante el método k-means, referente a parámetros característicos o de evaluación, el ciclo construido que tenga el menor error medio respecto a todos los parámetros será el CTC seleccionado (Desinedi et al., 2020; Roy & Morency, 2020; Yang et al., 2019).

Una de las desventajas del uso del algoritmo k-means es que sus resultados, si bien corresponden a un mínimo local de la Ecuación 3.3, no es uno de los mejores valores que se puede obtener. Para solucionar esto, es posible implementar metaheurísticas las cuales utilizan el algoritmo de búsqueda local en este caso el algoritmo k-means, para encontrar valores menores valores de la ecuación de función de costo, entre estas metaheurísticas se encuentran las siguientes:

***Greedy Randomized Search Procedure (GRASP)*** (Cano et al., 2002).

Esta metaheurística consiste en desarrollar varias veces el algoritmo k-means de manera tal, que la solución entregada sea aquella con el menor valor de la Ecuación 3.3. Esto se realiza una cierta cantidad de veces hasta encontrar convergencia, es decir que el mejor resultado obtenido no cambie durante una cierta cantidad de iteraciones o hasta un límite máximo de iteraciones, esto con el fin de optimizar el tiempo de cómputo para la búsqueda de la solución

***Iterated Local Search (ILS)*** (Berthold et al., 2003)

La Metaheurística ILS, consiste en desarrollar la búsqueda local inicial, mediante el algoritmo k-means, para posteriormente “perturbar” la solución encontrada. En este caso, la perturbación de la solución consiste en seleccionar uno de los centroides calculados por el algoritmo k-means de manera aleatoria para reemplazarlo con un nuevo centroide acorde al método de inicialización utilizada. Posteriormente, se repite este ciclo hasta encontrar la convergencia del menor valor

encontrado para la función de costo del algoritmo k-means o hasta que una cierta cantidad de iteraciones se hayan completado

***Algoritmo genético k-means (GKA, Genetic k-means, por sus siglas en inglés)*** (Yuan et al., 2021)

Una de las alternativas para el mejoramiento del valor de la función de costo del algoritmo k-means es la implementación de un algoritmo genético, teniendo así una variación del algoritmo k-means o GKA (*Genetic k-means Algorithm*, por sus siglas en inglés). Este algoritmo es explicado en detalle por Doenhoff et al. (Doenhoff et al., 1999) y aplicado para la construcción de ciclos típicos de conducción Yuan et al. (Yuan et al., 2021).

***Método de mínimas diferencias ponderadas de parámetros característicos (Minimim Weighted Differences of Characteristic Parameters, MWD-CP)***

El método de mínimas diferencias ponderadas de parámetros característicos (MWD -CP, por sus siglas en inglés), ha sido desarrollado y validado en diferentes trabajos relacionados con la construcción de CTCs (Echeverry Mejía, 2018; Huertas et al., 2018), de igual manera, este método ha sido comparado con algunos de los métodos presentados anteriormente específicamente el enfoque de microciclos y el método de cadenas de Markov (Huertas et al., 2019).

El método MWD-CP, principalmente consiste en seleccionar un ciclo de conducción de la entre todos los ciclos medidos, el cual tenga el menor valor acorde a la Ecuación 3.4(Echeverry Mejía, 2018; Huertas et al., 2018), la cual representa el método MWD-CP.

$$TDC = \min_j \left( \sum_i^n W_i \text{abs}(P_{i,j} - \bar{P}_i) \right) \quad \text{Ecuación 3.4}$$

Donde:

$P_{i,j}$ : Parámetro característico  $i$  del ciclo  $j$

$\bar{P}_i$ : Promedio del parámetro  $i$  de todos los ciclos recolectados

$W_i$ : Peso relativo del parámetro característico de cada parámetro  $i$

$n$  : Número total de parámetros característicos

Es posible analizar que el método estadístico MWD-CP, depende directamente de la cantidad de recorridos realizados y en los cuales se recolectó información, así como de los parámetros característicos seleccionados para la selección del CTC. Se destacan dos términos dentro de la ecuación, en primer lugar se analizan los parámetros seleccionados, estos parámetros como ya se mencionó y como se verá en la sección Metodología deben ser seleccionados con base a la influencia que tienen directamente sobre la(s) variable(s) de estudio respecto a las cuales se busca construir el CTC.

El segundo parámetro resaltado corresponde a los pesos relativos asignados a cada parámetro, para lo cual existen diferentes métodos de asignación de estos pesos, sin embargo uno de los métodos que ha sido validado y utilizado en trabajos previos (Echeverry Mejía, 2018; Huertas et al., 2018) para este análisis corresponde al método de pesos relativos (Relative Weight Analysis – RWA, por sus siglas en inglés), donde se asignan pesos a cada parámetro de manera tal que, la suma de estos sea 100%, acorde a su aporte al  $R^2$ , teniendo en consideración las correlaciones entre los parámetros seleccionados, con base a la predicción de la variable a la cual se quiere que el CTC sea representativo (Johnson, 2000; Lundby & Johnson, 2006). Teniendo en consideración que este método de asignación de pesos a los parámetros ha sido utilizado en proyectos con características similares, este será utilizado en el contexto de este proyecto.

Es importante mencionar que al realizar la selección del CTC acorde a la metodología MWD-CP, se debe tener en consideración toda la base de datos recolectados de los ciclos obtenidos durante los recorridos realizados, incluyendo los ciclos atípicos buscando tener una mayor cantidad de recorridos, de manera tal que, estos ciclos no sean seleccionados como el CTC. Siendo así, que el CTC seleccionado por esta metodología, será aquel que tenga en sumatoria acorde al aporte de cada parámetro el valor mínimo de la Ecuación 3.4, aquel que represente de manera general a toda la muestra de ciclos recolectados.

### **3.4. Medición de emisiones en vehículos del parque automotor liviano a gasolina**

La medición de las emisiones generadas por los vehículos livianos de combustión interna de gasolina es posible realizarla de dos formas, principalmente mediante analizadores portables de emisiones o PEMs por sus siglas en inglés (*Portable emissions measurement*), el cual permite realizar el análisis de los gases de combustión de un vehículo en tiempo real mientras el vehículo

se encuentra en carretera, o a través de la construcción de un CTC, como se vio en las secciones anteriores, para simular en un dinamómetro de chasis el recorrido en un vehículo y mediante un analizador de gases realizar la medición de las emisiones generadas. Esto con el fin de conocer los factores de emisiones del parque automotor para el cual se realizó el CTC.

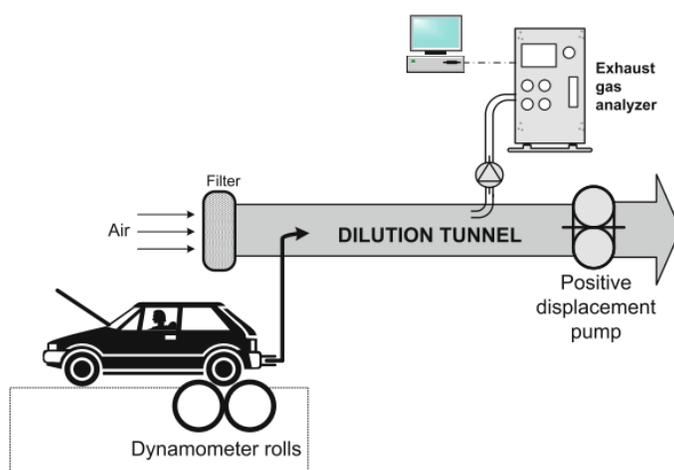
### 3.4.1. Generalidades sobre la medición de emisiones hechas sobre un dinamómetro de chasis

Un dinamómetro de chasis es un instrumento el cual permite simular condiciones de carretera respecto a diferentes parámetros estableciendo diferentes cargas al motor, esto mediante diferentes tipos de pruebas entre ellas la aplicación de un CTC construido para la zona de análisis. En sí, un dinamómetro de chasis consiste en un rodillo el cual conectado a un eje varia su velocidad angular, simulando la “velocidad de la carretera”, sobre los cuales se posicionan las llantas traseras o delanteras del vehículo a analizar, el vehículo es encendido y el dinamómetro de chasis calcula la velocidad a la cual el vehículo se debe encontrar, de igual manera el dinamómetro de chasis debe simular cada una de las fuerzas existentes en carretera tales como, las fuerzas debido a una pendiente, las fuerzas aerodinámicas, la resistencia a la rodadura, y la inercia del vehículo, siendo la fuerza total de tracción la sumatoria de las fuerzas anteriormente mencionadas, como se presenta en la Ecuación 3.5 (Giakoumis, 2016).

$$F_{tr} = C_{fr}m_v g + m_v g \sin(\theta) + \frac{1}{2} \rho_a C_d A_f V^2 + m_v \frac{dV}{dt} \quad \text{Ecuación 3.5}$$

En la Ecuación 3.5, el primer término corresponde a la fuerza de resistencia por rodadura, donde  $C_{fr}$  corresponde al coeficiente de rodadura, y  $m_v$  corresponde a la masa del vehículo, el segundo termino corresponde a la fuerza ocasionada por la inclinación de la carretera, el tercer termino corresponde a la fuerza aerodinámica o fuerza de arrastre, donde  $\rho_a$  corresponde a la densidad del aire,  $C_d$  al coeficiente de arrastre, y  $A_f$  al área frontal del vehículo, finalmente el ultimo termino corresponde a la fuerza de inercia del vehículo. Las fuerzas anteriormente mencionadas deben ser simuladas por el dinamómetro de chasis con el objetivo de representar el comportamiento en carretera y simular las cargas reales al vehículo.

Las pruebas desarrolladas en un dinamómetro de chasis tienen diversas aplicaciones así como los CTCs, pero específicamente una de ellas es la medición emisiones producidas por un vehículo, cuando este se expone a cargas reales, claramente el vehículo permanece estable, mientras se realiza la prueba, el dinamómetro de chasis simula todas las fuerzas, y sigue el CTC con ciertas tolerancias establecidas, mientras que a la par un analizador de gases conectado al exhosto del vehículo realiza la medición de los gases de combustión, para conocer la cantidad de emisiones producidas por el vehículo. En la Figura 3.5, se presenta el esquema de un dinamómetro de chasis conectado aun analizador gases para la medición de emisiones.



*Figura 3.5.* Esquema de un dinamómetro de chasis conectado a un analizador de gases para la medición de emisiones. (Giakoumis, 2016)

En la Figura 3.5, se puede evidenciar el uso de un túnel de dilución y una bomba de desplazamiento positivo, estos elementos son opcionales, ya que actualmente se cuenta con analizadores de gases los cuales son directamente conectados al exhosto del vehículo.

### **3.4.2. Medición de emisiones con base en la aplicación de un ciclo típico de conducción**

Cuando se cuenta con un CTC de la zona la cual se requiere analizar, y el vehículo realiza el seguimiento del CTC en el dinamómetro de chasis, se procede a realizar la medición de los gases de combustión, para los cuales comúnmente se realizan mediciones son  $\text{CO}_2$ ,  $\text{CO}$ ,  $\text{NO}_x$ ,  $\text{NO}_2$ ,  $\text{SO}_x$ ,  $\text{SO}_2$ ,  $\text{O}_2$  y  $\text{THC}$ , mediante un analizador de gases. En la literatura se presentan sistemas de vehículo controlados por computadora CVS, por sus siglas en inglés, con bombas de desplazamiento

positivo, y diferentes implementos los cuales se utilizan para tener una mejor medición (Alves et al., 2015; Giakoumis, 2016).

Generalmente la cantidad de contaminantes que se puede medir con base en la aplicación de un CTC en un dinamómetro de chasis dependerá del analizador de gases con el cual se cuente. En Colombia se establece la cantidad de contaminantes que se deben medir en base a la norma técnica colombiana NTC 4983 (ICONTEC, 2012), donde se establece la cantidad que debe medir un analizador de gases, los rangos y unidades en las cuales debe medir.

Tabla 3.8. *Requisitos de medición de gases para un analizador de gases acorde a la norma NTC 4983 (ICONTEC, 2012)*

Canal	Rango	Exactitud	Ruido	Repetibilidad
HC (ppm)	0 - 400	± 12	6	8
	401 - 1 000	± 30	10	15
	1 001 - 2 000	± 80	20	30
CO (%)	0 - 2	± 0,06	0,02	0,03
	2,01 - 5	± 0,15	0,06	0,08
	5,01 - 10	± 0,40	0,10	0,15
CO <sub>2</sub> (%)	0 - 4	± 0,60	0,20	0,3
	4,1 - 14	± 0,60	0,20	0,3
	14,1 - 16	± 0,60	0,20	0,3
O <sub>2</sub> (%)	0 - 10,0	± 0,5	0,3	0,4
	10,1 - 22,0	± 1,3	0,6	1,0

Como se establece en la Tabla 3.8, la medición de estos gases se debe realizar durante todo el recorrido del CTC, con el objetivo de analizar la cantidad de contaminantes producidos respecto a las diferentes cargas que el motor experimenta.

En este caso, se evidencia la importancia del CTC ya que representa el comportamiento de la mayoría de los vehículos, así el vehículo estará expuesto a cargas reales, y evidenciará la emisión de contaminantes la cual tendrá relación con la realidad.

### 3.4.3. Métodos para el análisis de las emisiones registradas

Con las mediciones de emisiones de los vehículos a partir de la aplicación del CTC en el dinamómetro de chasis y el analizador de gases, se obtiene una base de datos de las emisiones generadas, es de aclarar que es necesario realizar varias mediciones con el objetivo de tener una

muestra representativa del sector del parque automotor del cual se quiere conocer los factores emisión. En el marco de este proyecto los vehículos ligeros (menores a 12 pasajeros) (Ministerio de Ambiente y Desarrollo Sostenible, 2017), de combustión interna a gasolina.

Con los datos obtenidos de la muestra representativa (esta muestra debe tener incluida en si los vehículos representativos del sector, ya que son los principales aportantes a las emisiones), es posible realizar el cálculo de los factores de emisión del sector.

**Cálculo de factores de emisión.** Un factor de emisión es un valor representativo que relaciona la cantidad de un contaminante liberado a la atmosfera con la actividad que libera el contaminante a la atmosfera, generalmente se expresa en unidades del peso del contaminante liberado a la atmosfera, sobre peso, volumen, distancia o duración de la actividad que emite el contaminante, estos factores de emisión son de suma importancia ya que permiten el cálculo de las emisiones generadas por un sector de la industria. (USEPA AP 42, 1995).

El cálculo del factor de emisión para fuentes móviles como lo es un vehículo ligero tiene diferentes componentes, no solamente la emisión de gases en los tubos de escape, se tienen diferentes emisiones que intervienen en este caso, en la Figura 3.6, se evidencian dichos contaminantes.



*Figura 3.6.* Esquema de emisiones generadas por un vehículo (Ministerio de Ambiente y Desarrollo Sostenible, 2017)

En este caso los principales contaminantes atmosféricos se emiten a través de los gases de combustión, por este motivo se realiza el enfoque en los gases del exhosto. La ecuación que

relaciona las emisiones con el factor de emisión es la Ecuación 3.6 (Ministerio de Ambiente y Desarrollo Sostenible, 2017).

$$E = FE_i * FA_t \quad \text{Ecuación 3.6}$$

Donde  $E$ , corresponde a la emisión producida,  $FE$  al factor de emisión del contaminante  $i$ , y  $FA$  el factor de la actividad en el tiempo  $t$ . Cabe aclarar que estas emisiones deben tener en consideración arranque en frío y arranque en caliente de un vehículo, el cual se encuentra contemplado en el CTC.

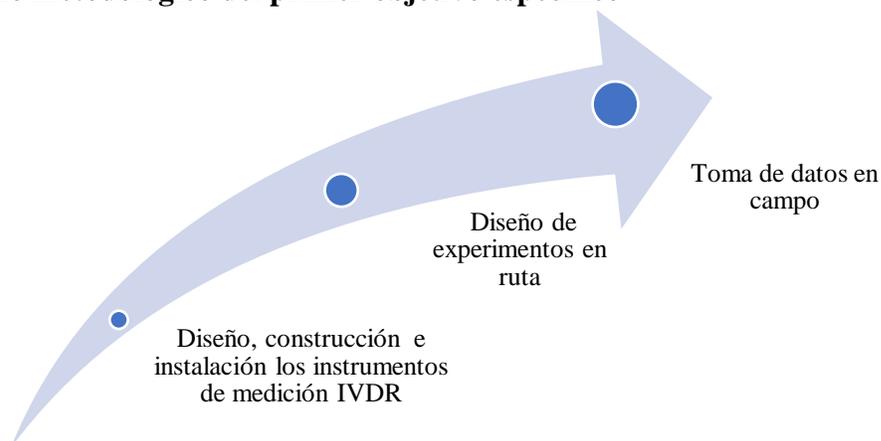
El factor de actividad, generalmente se expresa por la distancia recorrida por una población de vehículos en un tiempo determinado, y esta también puede hacer referencia al consumo del combustible.

El factor de emisión puede ser obtenido directamente mediante los resultados del analizador de gases, y/o mediante literatura, donde se contengan los factores para el sector a analizar (Ministerio de Ambiente y Desarrollo Sostenible, 2017). En el caso de la estimación del parque auto motor liviano de la zona Bogotá región se requiere establecer la relación entre cada uno de los contaminantes medidos por el analizador de gases y elaborar la ecuación, la cual defina los factores de emisión (Frey & Eichenberger, 1997), mediante métodos estadísticos debido a la gran cantidad de información que se recolectará.

#### 4. Metodología

Para el desarrollo del proyecto se establece la siguiente metodología, acorde a cada uno de los objetivos específicos.

##### 4.1. Desarrollo metodológico del primer objetivo específico



*Figura 4.1.* Esquema de la metodología para el desarrollo del primer objetivo específico

A continuación, se realiza una descripción detallada de cada uno de los pasos anteriormente presentados en la Figura 4.1.

#### **4.1.1. Diseño, construcción e instalación de los instrumentos de medición IVDR**

Para este paso se desarrolló un IVDR tal como se presenta en el diagrama modular en la Figura 3.4. Para el diseño y construcción del dispositivo IVDR acorde a los requerimientos de diseño expuestos en la sección 3.2.2, se plantean los siguientes pasos metodológicos, acorde a los trabajos previos desarrollados (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022) .

***Requerimientos de diseño del dispositivo IVDR.*** Los requerimientos generales de diseño para la construcción del IVDR fueron planteados en la sección 3.2.2 respecto a los componentes necesarios a ser utilizados en la construcción del dispositivo, así como las funciones generales principales que el dispositivo debe cumplir en su funcionamiento. Sin embargo, en el marco del desarrollo del proyecto se añaden los siguientes requerimientos de diseño (Echeverry Mejía, 2018):

- Debe ser el instrumento de medición de variables vehiculares en tiempo real, para la construcción del CTC.
- Debe ser el dispositivo base con el cual se valide la metodología para la evaluación medioambiental del parque automotor liviano, tanto como dispositivo de medición para la construcción del CTC, como dispositivo de medición de consumo de combustible durante el desarrollo de las pruebas dinámicas en el dinamómetro de chasis.

***Desarrollo de hardware del dispositivo IVDR.*** Para la construcción del dispositivo IVDR, se debe contar con los componentes necesarios para que el dispositivo cumpla con las funciones y requerimientos de diseño necesarios. Acorde con los requerimientos mencionados en la sección 3.2.2, se presentan una vez más cada uno de estos componentes, su función, y los componentes utilizados para su construcción.

- **Módulo de microcontrolador:** Es una de las partes fundamentales para la construcción del dispositivo IVDR, ya que permite la programación, e implementación de los algoritmos necesarios para que el dispositivo se comunique

con el vehículo, así mismo este microcontrolador debe cumplir con los siguientes requisitos:

- Debe contar el protocolo de comunicación CAN, para así poder interactuar con el vehículo y adquirir la información necesaria.
- Debe contar el con protocolo de comunicación UART, para así poder comunicarse con el computador para su programación, y poder comunicarse con diferentes módulos.
- Este microcontrolador debe ser validado en la industria automotriz.

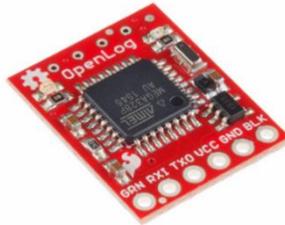
Para el desarrollo del dispositivo IVDR se seleccionó el microcontrolador SKEAZ128MLK4 con base en la tarjeta de desarrollo NXP KEAZ128 (Freescale, 2014), el cual cuenta con el protocolo de comunicación CAN, el protocolo de comunicación UART, y ha sido utilizado dentro de la industria automotriz. Este microcontrolador está diseñado, para realizar prototipados de manera rápida mediante su disposición para la conexión de cables muy similar a la configuración de conexión de cables a la de un ARDUINO UNO. En la siguiente imagen se presenta el microcontrolador utilizado.



*Figura 4.2.* Microcontrolador SKEAZ128MLK4 – Tarjeta de desarrollo NXP FRDM KEAZ128

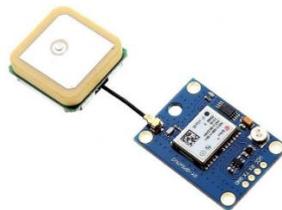
- Módulo de almacenamiento externo: Este módulo debe permitir guardar la información recolectada por el microcontrolador, y por el módulo GPS en una memoria tipo MicroSD de alta capacidad. El componente seleccionado corresponde al Datalogger SparkFun OpenLog DEV -13712, debido a que este componente no requiere programación adicional para la transmisión de datos y almacenamiento. En la memoria MicroSD, debido que en su arquitectura un microcontrolador

integrado, el microcontrolador ATmega328 siendo este el mismo microcontrolador utilizado en algunas versiones de Arduino UNO, reduciendo el tiempo de diseño del dispositivo IVDR. Este dispositivo permite una velocidad máxima de transmisión de 115200 baudios.



*Figura 4.3.* Datalogger SparkFun OpenLog DEV -137 12

- **Módulo GPS:** Este módulo permite conocer la ubicación en tiempo real del vehículo, para así conocer las rutas por las cuales se estuvo transitando, así como la inclinación de la carretera transitada. Este módulo debe contar con comunicación UART, tanto de transmisión como de recepción de datos, ya que los datos proporcionados por el módulo GPS son almacenados en la memoria MicroSD. El módulo seleccionado es el módulo GPS NEO 6M UBLOX, el cual ha sido utilizado y validado en trabajos previos (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022). Este cuenta con facilidad de conexión, transmisión y recepción de datos mediante el protocolo UART, este dispositivo cuenta con una velocidad máxima de transmisión de 9600 baudios.



*Figura 4.4.* Módulo GPS NEO 6M UBLOX

- **Módulo de alimentación:** El dispositivo IVDR es alimentado mediante el cable SAEJ1962 – DB9, el cual toma la energía eléctrica directamente del vehículo. Para

esto, es necesario utilizar convertidores DC-DC ya que el voltaje suministrado por el vehículo corresponde a  $\sim 14.2$  V. Sin embargo, el voltaje requerido para el funcionamiento del microcontrolador debe ser de 12 V, y para el módulo GPS se debe tener un suministro de 5V. En el caso del módulo de almacenamiento, este toma la energía directamente del microcontrolador. Para esto, se utilizan dos tipos de convertidores DC-DC: el convertidor DC-DC Buck XL405 de 5A para la alimentación del microcontrolador, debido a que maneja un rango mayor de corriente, y el convertidor DC-DC Buck LM 2596 de 2A, para la alimentación del módulo GPS.



*Figura 4.5.* Convertidores DC-DC BUCK XL4015 y BUCK LM2596

- Cable SAEJ1962-DB9: Este cable es indispensable ya que permite la comunicación entre el vehículo y el dispositivo IVDR mediante el puerto de conexión SAEJ1962 comúnmente encontrado en los vehículos actuales. Así mismo, permite el funcionamiento del dispositivo al ser este cable el que permite la alimentación de energía al dispositivo.

El dispositivo IVDR desarrollado en este trabajo difiere con el desarrollado y validado en trabajos previos (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022), específicamente en el uso de una diferente tarjeta de aplicación, siendo esta la tarjeta de desarrollo NXP FRDM KEAZ128. Sin embargo, esta tarjeta de desarrollo mantiene el mismo microcontrolador utilizado en trabajos previos siendo este el microcontrolador SKEAZ128MLK4. La mayor diferencia presentada corresponde a que el IVDR desarrollado en este trabajo no tiene en consideración un módulo GPRS para el almacenamiento de datos en la nube, lo cual disminuye los costos de desarrollo respecto a el no uso de un componente. Por otra parte, el uso del módulo de almacenamiento Sparkfun

OpenLog DEV -13712 no requiere de programación dentro del microcontrolador para el almacenamiento de los datos en sí, reduciendo así el tiempo de desarrollo de software.

Tabla 4.1. *Costos de cada uno de los componentes del IVDR*

Cantidad	Módulo del IVDR	Referencia	Precio (USD)	Precio (COP)
1	Módulo GPS	NEO 6M UBLOX	\$ 7.48	\$ 29,750.00
1	Microcontrolador	NXP FRDM KEAZ128	\$ 141.03	\$ 560,749.00
1	Módulo de almacenamiento	Sparkfun OpenLog DEV -13712	\$ 19.59	\$ 77,910.00
1	Módulo de alimentación	BUCK XL4105	\$ 3.50	\$ 13,924.00
1		Conector Jack 5.5 mm	\$ 0.60	\$ 2,380.00
1		BUCK LM2596	\$ 1.59	\$ 6,335.00
1	Memoria MicroSD	Memoria Micro SD de 8GB	\$ 10.10	\$ 40,161.00
1	Módulo de conexión	Cable DB9- SAE1962	\$ 12.58	\$ 50,013.00
1		Puerto DB9	\$ 0.45	\$ 1,785.00
1	Interruptor de tres pines		\$ 0.90	\$ 3,570.00
1	Interruptor de cuatro pines		\$ 1.67	\$ 6,646.00
4	LEDs		\$ 0.28	\$ 1,100.00
1	Porta fusibles		\$ 0.54	\$ 2,142.00
1	Fusible		\$ 0.09	\$ 374.00
1	Caja para el desarrollo de proyectos electrónicos		\$ 1.76	\$ 7,000.00
-	Cables de conexión		\$ 2.52	\$ 10,000.00
	Total		\$ 204.68	\$ 813,839.00

En la Tabla 4.1 se presentan los costos de cada uno de los componentes del IVDR, así como el costo final del dispositivo siendo este de ~\$205 USD, necesarios para la implementación de esta metodología a mayor escala. Si bien el costo de este dispositivo es mayor al de los dispositivos previos, es necesario considerar que el componente de mayor precio corresponde al microcontrolador. Es por este motivo que se tiene una diferencia de ~40% mayor en el costo final del IVDR desarrollado en este proyecto respecto a sus versiones previas.

El ensamblaje mecánico y el diagrama de conexiones del IVDR se presentan en la sección 5.1.

**Desarrollo de software del IVDR.** Como ya ha sido mencionado, el dispositivo IVDR de este proyecto es un avance a partir de trabajos previos, los cuales cuentan con un extenso desarrollo y validación en proyectos similares al desarrollado en esta tesis (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022). Específicamente, el software fue desarrollado en los trabajos mencionados, sin embargo, en este caso se presentan algunas modificaciones como ya fue mencionado.

Debido a que en este proyecto no es necesario el uso de un dispositivo GPRS, el cual permite el acceso a internet al IVDR para almacenar los datos en la nube, no es necesario realizar un desarrollo en el software del dispositivo para el almacenamiento de los datos alimentados debido al uso del módulo SparkFun OpenLog DEV1312. Este dispositivo permite el almacenamiento de una forma más intuitiva al contar con un microcontrolador integrado, donde solamente al enviar información a través del puerto UART en el microcontrolador KEAZ128 asignado al módulo MicroSD que guarda los datos en un archivo de texto.

Por esta razón, el software del dispositivo desarrollado tiene una menor complejidad en comparación con los dispositivos utilizados en trabajos previos. Sin embargo, los algoritmos y rutinas generales desarrolladas en trabajos previos son también utilizadas en el IVDR desarrollado para este proyecto (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022).

Para que el IVDR cumpla con su función debe recolectar datos en tiempo real de un vehículo a través del protocolo de comunicación CAN mediante el sistema de diagnóstico OBD2 utilizando PIDs específicos, así como, el registro del posicionamiento global mediante el módulo GPS, para así almacenar esta información en una memoria MicroSD.

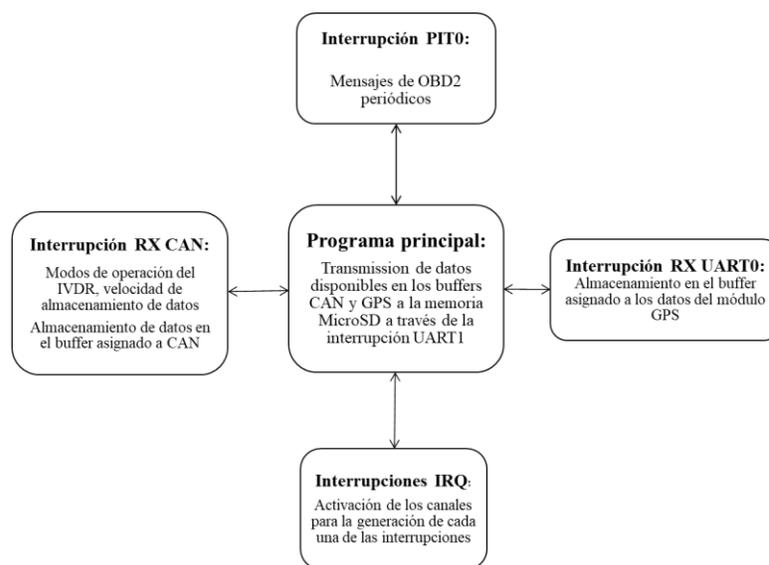
El dispositivo IVDR debe tener la capacidad de leer los diferentes tipos de información que se pueden obtener de la computadora motor de un vehículo. Por este motivo, se desarrollaron diferentes modos de operación: Uno de ellos es el modo “*sniffer*”, el cual permite leer todos los mensajes que pasan por el bus CAN; el modo que permite solicitar y leer solamente IDs específicos en protocolo CAN; y el modo OBD2, el cual permite solicitar y leer mensajes del sistema diagnóstico OBD2, con base a PIDs requeridos. En el IVDR desarrollado solo se utiliza el modo OBD2, debido a que con este modo los mensajes de solicitud de parámetros se encuentran estandarizados, a diferencia de los modos “*sniffer*”, y el modo de IDs de CAN. Estos últimos modos utilizan IDs para la solicitud de parámetros establecidos directamente por las casas automotrices los cuales son específicos para cada modelo de vehículo, los cuales se desconocen.

El software requerido para el funcionamiento del dispositivo IVDR fue desarrollado con base en la tarjeta de desarrollo NXP FRDM KEAZ128, la cual cuenta con el microcontrolador SKEAZ128MLK4, programado en el estudio de desarrollo integrado CodeWarrior basado en el lenguaje de programación C. Como se mencionó en la sección 3.2.2, el software del IVDR debe estar basado en interrupciones, lo que quiere decir que las rutinas utilizadas en el IVDR no

dependen de un programa secuencial, sino de eventos externos y del desarrollo de las mismas rutinas. Las rutinas de interrupción utilizadas por el software del dispositivo IVDR, son presentadas a continuación (Echeverry Mejía, 2018):

- Interrupción de recepción por CAN: Esta interrupción controla la petición de datos mediante al sistema de diagnóstico OBD2, así mismo lee los datos del bus CAN, y dependiendo del modo utilizado (OBD2) los almacena en el buffer asignado a los mensajes recibidos del bus CAN.
- Interrupciones PIT (*Periodic Interrupt Timer*, por sus siglas en inglés): Se establecen dos interrupciones de ese tipo, la primera para la sincronización la solicitud de datos mediante OBD2 y la segunda es utilizada para la temporización de algunas funciones del IVDR como errores y funciones que son ejecutadas periódicamente.
- Interrupciones seriales UART: Una interrupción de este tipo es establecida, aquella que recibe y almacena los datos del módulo GPS, y los almacena en un buffer asignado para este tipo de datos.
- Interrupción IRQ (*Interruption Request*, por sus siglas en inglés): Se realizan cuatro solicitudes de interrupción, una para la habilitación del canal de interrupción PIT, una para la habilitación del canal UART asignado a la recepción de datos del módulo GPS, una para la habilitación del canal UART asignado a la transmisión de datos al módulo de memoria MicroSD, y una solicitud de interrupción es utilizada para para el protocolo CAN.
- Rutina de programa principal: Esta rutina inicializa cada uno de los comandos principales del microcontrolador y periféricos (módulo GPS), su función principal es la de transmitir asincrónicamente los datos almacenados en los buffers a la memoria MicroSD.

En la Figura 4.6 se presenta la máquina de estados del dispositivo IVDR, donde se evidencia que cuando se está en una interrupción, no es posible pasar a otra sin antes haber pasado por el programa principal. En caso de que se presenten varias interrupciones al mismo tiempo, el software respetará la jerarquía de las interrupciones en este orden: IRQ, UART, PIT, CAN (Echeverry Mejía, 2018).



*Figura 4.6.* Máquina de estados del dispositivo IVDR

El código desarrollado se encuentra en ciclo infinito en el programa principal, esperando a que los datos de los buffers sean completados para que puedan ser transmitidos y almacenados en la memoria MicroSD. En caso de tener un evento externo como un mensaje de GPS o un mensaje de CAN, se ejecuta interrupción activando la función a la cual esta interrupción lleva, y una vez finalizada la función se regresa al programa principal para continuar con el ciclo infinito.

Es importante que, para validar el correcto funcionamiento de la recepción de datos de GPS y CAN, se agregó un LED indicador por cada una de estas señales, los cuales parpadean cada vez que se recibe una trama de datos. Así mismo, se agregó un LED indicador cada vez que se envía esta información a la memoria MicroSD.

El dispositivo IVDR maneja tres velocidades de lectura de datos para el sistema de diagnóstico OBD2, basado en el protocolo CAN, permitiendo así leer datos a la tasa máxima permitida de 100 Hz y a 10 Hz. En este proyecto, se escogió la velocidad de 1000 datos por segundo, ya que la mayoría de los parámetros seleccionados no tienen una variación en el tiempo tan alta. Sin embargo, el módulo GPS registra datos a 1 Hz.

Los algoritmos de las rutinas presentadas en la máquina de estados del dispositivo en la Figura 4.6, son muy similares a los algoritmos desarrollados con anterioridad (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022). Sin embargo, debido al uso de un módulo de almacenamiento

diferente, y que no se utilizó el módulo GPRS, se presentan modificaciones para estos algoritmos. En el apéndice A se analizan los algoritmos de cada una de las rutinas y subrutinas utilizadas en el IVDR a modo de diagrama de flujo.

Es importante mencionar que, para el desarrollo del proyecto de construyeron 9 dispositivos IVDRs.

**Validación del dispositivo IVDR.** La validación del dispositivo en aplicaciones similares a las presentadas (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022), donde se analiza la pérdida de datos del dispositivo, así como su validación respecto a la construcción de un ciclo típico de conducción. Sin embargo, para este proyecto solamente se realizó la validación de:

- Timestamp, análisis del tiempo de cada dato almacenado por OBDII.
- PIDs necesarios para el cálculo de los parámetros requeridos. Principalmente los sensores MAF (Mass Air Flow, por sus siglas en ingles), MAP (Mass Air Pressure, por sus siglas en ingles), y los sensores disponibles de oxígeno.

Es importante mencionar que cada IVDR instalado en cada vehículo tuvo que ser programado individualmente, debido a que todos los vehículos no cuentan con los mismos sensores, y tampoco con la misma tasa de transmisión de datos. Esta validación es importante, ya que no todos los vehículos cuentan con el sensor MAF, o con el sensor MAP, sensores importantes para el cálculo del consumo de combustible. Además, dependiendo de cada sensor disponible en el vehículo el cálculo se realiza con relaciones matemáticas diferentes.

Los resultados de la validación se presentan en la sección 5.1.3.

#### **4.1.2. Diseño de experimentos en ruta**

Acorde a lo expuesto en la sección 3.1.1, se realizará un CTC basado en el método de rutas variables, para lo cual se realizarán recorridos aleatorios en las principales vías y autopistas de la ciudad de Bogotá y sus municipios aledaños (Chía, Cota), siendo estas vías las que se mencionan a continuación.

- Avenida Pradilla
- Autopista Norte
- Autopista Sur

- Avenida Boyacá
- Avenida Ciudad de Cali
- Avenida Suba
- Avenida el Dorado
- Calle 80
- Carrera séptima
- Calle 13
- Calle 170
- Carrera novena
- Avenida Caracas
- Calle 26
- Variante Cota - Chía

Con base en estas principales vías, se plantearon recorridos en diferentes horarios, para así, capturar la información de la variación del tráfico. Los recorridos realizados corresponden, a rutas recorridas de manera aleatoria acorde al uso del del vehículo.

Las vías mencionadas representan un alto tráfico a pesar de ser catalogadas como vías principales, por tanto, a través de los IVDRs se obtienen los datos para determinar el comportamiento de los vehículos en la zona y la información necesaria para la estimación de los factores de emisión.

#### **4.1.3. Toma de datos en campo**

Para la toma de datos en campo se utilizaron dos vehículos, a los cuales se les instaló un dispositivo IVDR programado específicamente para cada uno, con el objetivo de leer las variables de interés, que se muestran en la Tabla 5.3. Estas variables incluyen los parámetros necesarios para la medición de consumo de combustible como el sensor MAF, el sensor MAP, y el sensor de oxígeno, así como la carga del motor.

Los vehículos utilizados para la recolección de datos realizaron diferentes recorridos en las rutas presentadas en la sección 4.1.2, durante los meses de Abril, mayo y Junio del año 2023 de manera aleatoria, acorde al enfoque de construcción del CTC. Estos vehículos debían cumplir con

las características de un vehículo liviano, siendo un en general un vehículo familiar. Las características técnicas de los vehículos disponibles para la toma de datos fueron los siguientes:

Tabla 4.2. *Ficha técnica general de los vehículos experimentales para la toma de datos*

Características	Vehículo 1	Vehículo 2
Modelo	Nissan Versa 2013	Ford Fiesta 2017
Cilindraje (L)	1.6	1.6
Peso (kg)	1445.7	1642
Potencia (hp @ rpm)	106 @ 5600	120 @ 6350
Torque (lb - ft @ rpm)	105 @ 4000	112 @ 5000
Transmisión	Automática - 4 velocidades	Automática - 6 velocidades
Número de válvulas	4 por cada pistón	4 por cada pistón
Sistema de inyección	Secuencial multipunto	Directa
Sistema de carga	Naturalmente aspirado	Naturalmente aspirado



Figura 4.7. Vehículos experimentales para la toma de datos. De izquierda a derecha, Nissan Versa 2013 y Ford Fiesta 2017

Para la medición de los datos experimentales en cada de uno de los vehículos fue necesario realizar una validación de los sensores que el vehículo puede tener, para así realizar la programación final del dispositivo IVDR de cada vehículo. La validación se realiza utilizando una lista de PIDs, con los cuales es posible calcular los parámetros de la Tabla 5.3, incluyendo consumo de combustible y carga del motor. Sin embargo, como se mencionó anteriormente no todos los vehículos cuentan con todos los sensores necesarios.

La validación consiste en programar toda la lista de PIDs en el IVDR, en la rutina RX CAN presentada en el anexo A.3. *Interrupción RX CAN*, y así validar los sensores con los cuales cuenta el vehículo, mediante los PIDs, directamente relacionados con los sensores. Una vez se conoce la lista de estos sensores, se programa una vez más el dispositivo IVDR, para así medir en cada vehículo solamente los PIDs, con los que el vehículo cuenta. De esta manera, se minimizar el

tiempo de ejecución reduciendo el envío de mensajes que no tendrán una respuesta. La lista de PIDs solicitados, y los PIDs, con los cuales cuenta cada vehículo se presentan en la Tabla 4.3.

Tabla 4.3. PIDs solicitados – PIDs de cada vehículo experimental

No.	PIDs solicitados			Vehículo experimental	
	PID Hexadecimal	PID Decimal	Sensor	Nissan Versa 2013	Ford Fiesta 2017
1	04	4	Carga del motor	x	x
2	0B	11	MAP (Mass Air Flow)	-	x
3	0C	12	RPM	x	x
4	0D	13	Velocidad	x	x
5	0F	15	IAT (Intake Air Temperature)	x	x
6	11	17	TPS (Throttle Position Sensor)	x	x
7	15	21	O2 -2 (Sensor de oxígeno 2)	x	x
8	24	36	Relación Aire - Combustible en el primer sensor de oxígeno	x	-
9	33	51	Presión Atmosférica	-	x
10	49	73	Posición del pedal	x	x
11	14	20	O2 -1 (Sensor de oxígeno - 2)	-	-
12	10	16	MAF (Mass Air Flow)	x	-
13	61	97	Porcentaje de torque solicitado por el conductor	-	-
14	62	98	Porcentaje de torque actual del motor	-	-
15	63	99	Torque de referencia del motor	-	-
16	5E	94	Flujo del combustible	-	-

Se solicitaron en total 16 PIDS, de los cuales el vehículo Nissan versa cuenta con nueve de estos sensores, al igual que el vehículo Ford Fiesta 2017. Ninguno de los vehículos cuenta con los últimos PIDs solicitados, ya que son muy pocos los vehículos que cuentan con estos tipos de sensores actualmente en Colombia. Como se observa se solicitaron dos sensores de oxígeno, ya que algunos vehículos cuentan con dos o más sensores de oxígeno en el escape, antes del catalizador y después del catalizador. El valor de estos sensores es importante para el cálculo de la relación aire – combustible, la cual es necesaria para el cálculo de flujo de combustible. De igual manera, se analiza para el vehículo Nissan Versa el PID el cual contiene directamente la relación de aire combustible.

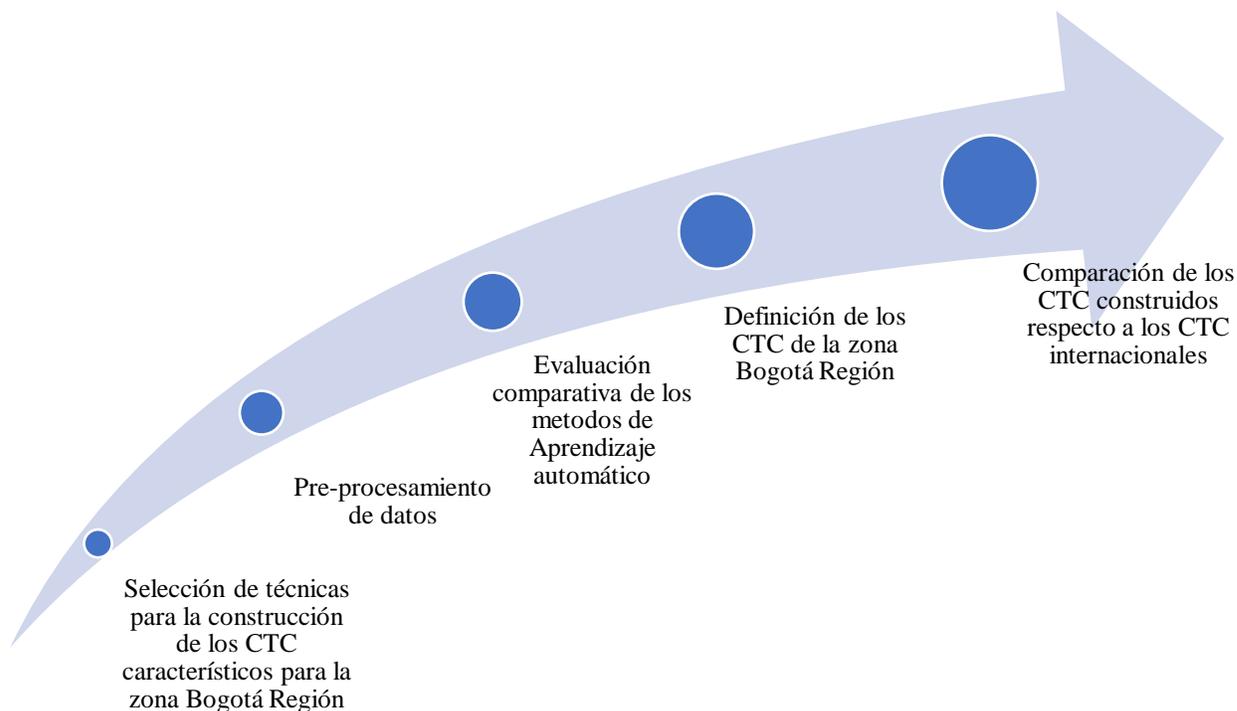
Cada uno de estos PIDs fueron solicitados acorde a la norma SAE J1979 (Society of Automotive Engineers, 2007), la cual contiene tanto los PIDs necesarios, como las ecuaciones para el cálculo de los valores solicitados. Como se mencionó en la sección 3.2.3, los datos obtenidos

pueden venir divididos en diferentes bytes, el cálculo de cada uno de los parámetros y sus respectivas formulas se analizará en la sección 4.2.2.

Los PIDs con los cuales cuenta los vehículos fueron analizados y almacenados durante los recorridos realizados, obteniendo así 59 archivos, de 59 recorridos. Los cuales contienen los parámetros solicitados mediante OBDII y los mensajes del módulo GPS.

El dispositivo IVDR, fue encendido durante cada uno de los viajes, una vez el vehículo estuviese en funcionamiento. Desde el momento en el que el vehículo empieza a funcionar, debido a que los encendidos en frío contribuyen a el análisis de consumo de combustible y emisiones. Además, no se presentan restricciones de peso, ni de posicionamiento geográfico para los recorridos. Una vez se finaliza el recorrido el IVDR fue apagado, el cual cierra el archivo de esos recorrido, y una vez se encienda de nuevo guardará los datos del nuevo recorrido en un archivo diferente.

#### 4.2. Metodología para el desarrollo del segundo objetivo específico



*Figura 4.8.* Esquema de la metodología para desarrollar el segundo objetivo específico

En las siguientes secciones se realiza una descripción de cada uno de los pasos metodológicos para el desarrollo del segundo objetivo específico.

#### **4.2.1. Selección de técnicas para la construcción de los ciclos típicos de conducción característicos de la zona (Bogotá – Región)**

Para el análisis de los CTCs y acorde a la literatura presentada, se utilizaron tres métodos y se realizó la respectiva comparación.

- Método MWD-CP (Echeverry Mejía, 2018; Huertas et al., 2018)
- Método de aprendizaje de automático basado en el método K-means, unificado con el método de cadenas de Markov y el método PCA (Desineedi et al., 2020; J. Liu et al., 2016; Roy & Morency, 2020; Yuhui et al., 2019)
- Una combinación del método de aprendizaje de máquina y el método MWD-CP

Basados en el enfoque de microciclos, donde cada ciclo medido es segmentado en pequeños ciclos (microciclos), existen diferentes definiciones de un microciclo (Nouri & Morency, 2017). Para este proyecto, un microciclo es definido como el periodo de tiempo entre dos paradas del vehículo tendiendo en consideración su periodo inicial de ralentí, donde se tienen combinaciones de microciclos y el método k-means (Fotouhi & Montazeri-Gh, 2013). Para la construcción del CTC se utilizará la combinación del método k-means y el método de cadenas de Markov, donde el método k-means realiza la clusterización de los microciclos segmentados, y con base en este, se calcula la TPM para así construir el ciclo típico (Desineedi et al., 2020). La combinación de estos dos últimos métodos basados en el enfoque de microciclos con el método PCA para la construcción del CTC aumentan el espacio de análisis del método k-means. Con el objetivo de reducir este espacio, y así el tiempo computacional, se realiza una reducción de dimensionalidad donde los componentes principales resultantes deben representar en un 98% la variabilidad de los parámetros seleccionados (Yang et al., 2019; Yuan et al., 2021; Yuhui et al., 2019).

En la sección 3.3.1, se presentó de manera general el funcionamiento de los algoritmos k-means, y PCA, sin embargo, no se presentaron los pseudocódigos del algoritmo. Teniendo en consideración que estos algoritmos presentan procesos metodológicos importantes, se presentan aquí los pseudocódigos donde se visualiza el funcionamiento general de estos algoritmos.

##### ***Análisis de componentes Principales (PCA)***

Como se mencionó, este algoritmo cumple con la funcionalidad de realizar una disminución de dimensionalidad de los parámetros calculados a los microciclos segmentados, conservando en

un 98% la variabilidad de los parámetros originales. El pseudocódigo de este algoritmo es presentado a continuación.

---

#### **Algoritmo 4.1** Análisis de Componentes Principales (PCA)

---

**Entrada:**  $X_{m \times n}$

**Procedimiento:**

1. Calcular la media de cada parámetro para todos los microciclos, guardar los datos en un vector con el mismo número de elementos, que columnas de la matriz X
2. Calcular la desviación estándar de cada parámetro para todos los microciclos, guardar los datos en un vector con el mismo número de elementos, que columnas de la matriz X

3. **Para**  $j = 1 \leq n$  **Hacer**

**Para**  $i = 1 \leq m$  **Hacer**

$$X_{\text{norm}ij} = X_{ij} - \mu_j / s_j$$

**Fin para**

**Fin para**

4. Calcular la matriz de covarianza de la matriz de datos X

$$C = (1/n) * X_{\text{norm}}^T X_{\text{norm}}$$

5. Calcular la descomposición de valores singulares para la matriz C

**Para**  $k = 1 \leq n$  **Hacer**

6. Calcular la suma de la diagonal de la matriz de valores singulares

$$\text{SUM} = \text{SUM} + S_{kk}$$

**Fin para**

7. Seleccionar la cantidad de componentes principales que representa el 98% de la variación de los datos

PCAq=0

var=0

**Para**  $l \leq n$  **Hacer**

$$\text{SUM2} = \text{SUM2} + S_{ll}$$

$$\text{var} = \text{SUM2} / \text{SUM}$$

**Si**  $\text{var} \geq 0.98$

PCAq=l

Salir del ciclo

**Fin si**

**Fin Para**

8. Seleccionar los componentes principales con base en la cantidad de componentes principales calculados con anterioridad.

Seleccionar la cantidad de columnas de la matriz de vectores singulares izquierdos calculada en el paso 5.

$$\text{PCA} = U_{n \times \text{PCAq}}$$

9. Calcular la matriz con la reducción de dimensionalidad de la matriz original.

$$Z_{m \times k} = X_{m \times n} * \text{PCA}_{n \times \text{PCAq}}$$

**Salida:** Matriz con la reducción de dimensionalidad de los parámetros originales calculados de los microciclos, lo cual representa el 98% de la variación de los datos originales.

---

Los datos de entrada del algoritmo PCA corresponden a los parámetros calculados de cada uno de los microciclos, los cuales se presentan en la Tabla 5.3. Teniendo esto en consideración, es necesario realizar la reducción de dimensionalidad debido a que cada parámetro representa una dimensión de análisis, y estos parámetros pueden presentar colinealidad entre sí, y error puede ser

añadido en la clusterización de los microciclos. La salida del algoritmo PCA, corresponde a la matriz de componentes principales con un tamaño  $m \times PCAq(k)$ , donde  $m$  corresponde a la cantidad de microciclos segmentados en la base de datos, y  $PCAq$  o  $k$  a la cantidad de componentes principales.

La matriz resultante corresponde a los datos de entrada necesarios para el algoritmo k-means.

### ***Algoritmo k-means***

Los datos de entrada del algoritmo k-means corresponden a la matriz obtenida mediante con reducción de dimensionalidad del algoritmo PCA, y la cantidad de clusters en los cuales asignar los microciclos. La cantidad de clusters (K) fue calculada mediante el método del codo de acuerdo con el análisis de la literatura (Yang et al., 2019; Yuan et al., 2021), seleccionando así la cantidad de clusters como 3.

Las salidas de este algoritmo corresponden a la solución óptima encontrada, incluyendo el vector de asignación de cada microciclo a un cluster, y las coordenadas de los centroides óptimos calculados. La asignación de clusters es de suma importancia para el desarrollo de la TPM, y posteriormente la construcción del CTC.

---

#### **Algoritmo 4.2** Algoritmo K-means

---

**Entrada:**  $Z_{m \times PCAq}$ ,  $K$   $K$  es el número de clusters

**Procedimiento:**

1. Inicializar los centroides iniciales de los clusters
2. Asignar cada microciclo a cada clúster hasta que los centroides converjan.
- Mientras (1) Hacer**
  - Para  $i=1 \leq m$  Hacer**
    3. Calcular la distancia euclidiana de cada microciclo a cada centroide
    4. Asignar cada microciclo al centroide cuya distancia euclidiana sea la menor
  - Fin para**
  - Para  $j=1 \leq m$  Hacer**
    5. Calcular la distancia promedio de los microciclos asignados a cada cluster  $K$
    6. Asignar las coordenadas de la distancia promedio calculada como los nuevos centroides
  - Fin para**
  7. Comparar el valor de la función de costo de la iteración anterior con el valor de la iteración actual
    - Si valor de la función de costo en la iteración previa – valor de la función de costo en la iteración actual  $\leq 0.01$  Hacer**
      - Salir del ciclo
    - Fin si**

**Fin mientras**

**Output:** Vector con la asignación de los ciclos a cada clúster  
La matriz con las coordenadas de los centroides calculados

---

Debido a que el método k-means es una heurística la cual generalmente se queda estancada en óptimos locales, es necesario el uso de metaheurísticas para obtener mejores soluciones. Las metaheurísticas seleccionadas son la metaheurística GRASP, la metaheurística ILS, y la metaheurística GKA, las cuales son combinadas con el algoritmo k-means, donde se presentan sus respectivos algoritmos en pseudocódigo.

Se utilizaron dos métodos de inicialización para la metaheurística GRASP y la metaheurística ILS. (Peña et al., 1999)

1. la inicialización aleatoria (Random) la cual consiste en seleccionar los datos para los centroides del algoritmo k-means de manera aleatoria del espacio de análisis, para su respectiva inicialización.
2. La inicialización Forgy, la cual consiste en seleccionar los datos los centroides del algoritmo k-means de manera aleatoria de los datos disponibles de la segmentación y reducción de dimensionalidad de microciclos.

***Greedy Randomized Search Procedure (GRASP)*** (Cano et al., 2002).

Los parámetros del algoritmo se establecen de la siguiente manera 100 iteraciones para la convergencia del resultado, y 400 iteraciones como límite máximo de análisis. El pseudocódigo del algoritmo se presenta a continuación.

---

#### **Algoritmo 4.3** Metaheurística GRASP combinada con k-means

---

**Entradas:**  $Z_{mxk}$ ,  $K$                        $K$  número de clusters

**Procedimiento:**

1. Asignar un parámetro para la validación del valor final de la función de costo  
 $FCV = \infty$
2. Establecer el número iteraciones para la validación de convergencia  
 $n = 100$
3. Establecer el número máximo de iteraciones para el algoritmo  
 $m = 400$
4. Inicializar el contador para el ciclo mientras  
 $counter1 = 0$
5. Inicializar el contador para la validación de convergencia  
 $counter2 = 0$
6. Desarrollar la metaheurística GRASP junto con el algoritmo k-means  
**Mientras (1)**  
 $counter1 := counter1 + 1$
7. Desarrollar la búsqueda local mediante el algoritmo k-means teniendo en consideración el método de inicialización (RANDOM, FORGY)
8. Analizar si la función de costo encontrada es menor que la función de costo encontrada en la iteración anterior

**Si Mejor solución encontrada por k-means < FCV hacer**



```

Si Mejor valor de la función de costo encontrada en la iteración anterior = Mejor valor de la función
de costo encontrada en la iteración actual Hacer
    counter2:=counter2+1
Sino
    counter2=0
Fin si
11. Comparar el Contador para la validación de convergencia y el contador para la validación de
iteraciones máximas
Si counter2 >= n o counter1 >= m Hacer
    Salir del ciclo
Fin si
Fin mientras
Salida: Best solution

```

---

*Algoritmo genético k-means (GKA, Genetic k-means, por sus siglas en inglés)* (Yuan et al., 2021)

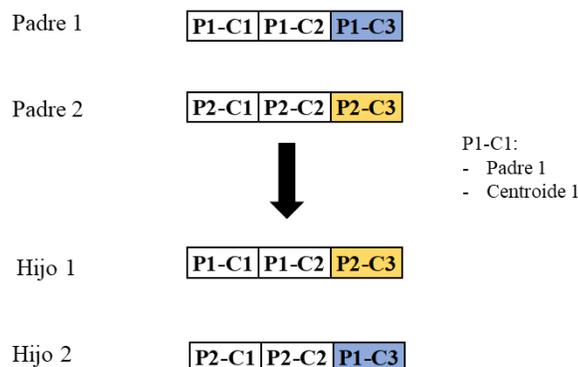
La metodología del algoritmo genético aquí propuesto sigue la metodología general de los algoritmos genéticos, sin embargo, presenta algunas modificaciones como se describe a continuación. La población inicial (primera generación) se selecciona aleatoriamente de la base de datos de los microciclos, y un individuo de la población se define como el conjunto de tres centroides. El algoritmo k-means es utilizado como la función de costo de esta metaheurística, así mismo cada individuo es utilizado como los centroides iniciales. El inverso de la función del valor de la función de costo encontrada por el algoritmo k-means se selecciona como el valor fitness teniendo como centroides iniciales. Un 10% de la población que tengan los mejores valores fitness son directamente introducidos en la siguiente generación, y el 90% restante se completa mediante la operación de crossover la cual será explicada más adelante. La mejor solución encontrada dentro de toda la población se guarda como la mejor solución hasta encontrar una mejor, en las siguientes generaciones. Si, la mejor solución encontrada no cambie durante cuatro generaciones el algoritmo para, así mismo si se han completado 10 generaciones.

La operación de crossover desarrollada por (Yuan et al., 2021), se define de la siguiente manera:

1. Seleccionar aleatoriamente dos individuos de la población como padres para desarrollar la operación crossover. Los padres son seleccionados aleatoriamente, teniendo como criterio de ponderación el valor del fitness de cada uno.
2. Seleccionar aleatoriamente el número de centroides a ser intercambiados de cada individuo.

3. Seleccionar aleatoriamente la posición de los centroides a ser intercambiados.
4. Intercambiar los centroides entre los individuos seleccionados
5. Los “hijos” resultantes nuevos conjunto de centroides, son introducidos en la población, hasta completar la cantidad de individuos requerida.

Un ejemplo gráfico de la operación de crossover se presenta en la Figura 4.9.



*Figura 4.9.* Ejemplo de operación Crossover

Posterior a la operación de crossover, un número es aleatoriamente generado entre cero y uno y es comparado con una probabilidad de mutación establecida, la cual es un parámetro de entrada del algoritmo genético establecido como 0,01. Si el número generado es menor o igual a la probabilidad de mutación, uno de los hijos generados en la operación de crossover es seleccionado aleatoriamente y reemplazado con un conjunto de centroides totalmente nuevo generado mediante el método de inicialización RANDOM. El funcionamiento general de este algoritmo es presentado en forma de pseudocódigo.

---

**Algoritmo 4.5 Genetic k-means**


---

**Entrada:**  $Zmxk, K$   $K$  es el número de clusters

**Procedimiento:**

1. Asignar el número de individuos en la población  
 $M=100$
2. Asignar la probabilidad de mutación  
 $P=0.01$
3. Asignar el número máximo de generaciones para la convergencia  
 $N=4$
4. Asignar el número máximo de generaciones  
 $G=10$
5. Asignar una variable para guardar el mejor valor de la función de costo interna  
 $FCVin=\infty$
6. Asignar una variable para guardar la mejor función de costo encontrada  
 $FCV=\infty$
7. Inicializar el contador para el ciclo mientras  
 $n=0$
8. Inicializar el contador para la validación de convergencia  
 $counter=0$
9. Desarrollar el algoritmo genético k-means
  - Mientras (1)**
  - $n:=n+1$
  - Si  $n = 1$  Hacer**
  - 10. Generar la población inicial
  - Sino**
  - 16. Seleccionar el 10% de los individuos con el mejor valor de la función de costo de la generación anterior e insertarlos en la generación actual
  - 17. Desarrollar la operación de crossover acorde a la **Figura 4.9** hasta completar el número de la población, considerando la mutación
  - 18. Generar un número aleatorio entre 0 - 1 ( $p$ )
  - 19. Comparar el número generado con la probabilidad de mutación, y realizar el proceso de mutación según corresponda
    - Si  $p \leq P$  Hacer**
    - Seleccionar aleatoriamente uno de los hijos de la previa operación de crossover , y generar aleatoriamente (Random) un individuo completamente nuevo (grupo de centroides), para así reemplazar el hijo seleccionado con el individuo generado.
    - Fin si**
    - Fin si**
  - 11. Desarrollar el algoritmo k-means con cada uno de los individuos de la población (centroides iniciales)
  - 12. Calcular el valor fitness para cada individuo de la generación
  - 13. Calcular la mejor solución para la generación
    - Si Mejor solución encontrada por el algoritmo k-means para el individuo  $< FCVin$  Hacer**
    - $FCVin:=$  Mejor solución encontrada por el algoritmo k-means para el individuo
    - Best solution in:= Solución encontrada por el algoritmo k-means para el individuo
    - Fin si**
  - 14. Calcular la mejor solución para todas las generaciones
    - Si  $FCVin < FCV$  Hacer**
    - $FCV:= FCVin$
    - Best solution:=Best solution in
    - Fin si**
  - 15. Comparar los valores para convergencia o validación del límite máximo de generaciones
    - Si  $n > 1$  Hacer**
    - Si** Mejor solución en la generación anterior = Mejor solución en la generación actual en comparación
    - $counter:=counter+1$
    - Sino**
    - $counter:=0$
    - Fin si**

**Fin si**  
**Si counter > = Non >=G Hacer**  
 break  
**Fin si**  
**Fin mientras**  
**Salida:** Best solution

---

Los parámetros seleccionados para el análisis son obtenidos del estado del arte, con base en que estos parámetros tienen una influencia en las emisiones y el consumo de combustible del vehículo. Los parámetros seleccionados se presentan en la Tabla 4.4.

Tabla 4.4. *Parámetros seleccionados para el análisis del ciclo típico de conducción.*

Parámetros	Unidad	Yang et al., 2019	Huertas et al., 2018	Ho et al., 2014	Yuan et al., 2021
Velocidad máxima	km/h	x		x	x
Velocidad promedio	km/h	x		x	x
Velocidad media en movimiento	km/h	x		x	
Aceleración máxima	m/s <sup>2</sup>	x	x	x	x
Desaceleración máxima	m/s <sup>2</sup>	x			
Aceleración promedio	m/s <sup>2</sup>	x		x	x
Desaceleración promedio	m/s <sup>2</sup>	x		x	x
Distancia promedio	km	x			
Tiempo en movimiento	s	x			
Proporción de aceleración	%	x		x	x
Proporción de desaceleración	%	x		x	x
Proporción de velocidad crucero	%	x		x	x
Proporción de ralentí	%	x		x	x
Rango de velocidad (0-20 km/h)	%	x			
Rango de velocidad (20-40 km/h)	%	x			
Rango de velocidad (40-60 km/h)	%	x			
Rango de velocidad (60-80 km/h)	%	x			
Rango de velocidad (>80 km/h)	%	x			
Velocidad en terreno montañoso	m/s		x		
Tiempo de ralentí	s		x		
Número de paradas	-		x		
Inclinación de la carretera	-		x		
Fuerzas aerodinámicas	m <sup>3</sup> /s <sup>3</sup>		x		
Altitud máxima	m		x		
Duración media de un ciclo de medición	s			x	
Proporción de arrastre	%			x	
Promedio de paradas por kilometro	km <sup>-1</sup>			x	
Promedio de cambios de aceleración y desaceleración	-			x	
Aceleración mínima	m/s <sup>2</sup>			x	
Desaceleración mínima	m/s <sup>2</sup>			x	x
Desviación estándar de la velocidad	km/h				x
Desviación estándar de la desaceleración	m/s <sup>2</sup>				x
Velocidad promedio del motor	rpm				x
Flujo de combustible promedio	L/h				x
Desviación estándar de la aceleración	m/s <sup>2</sup>				x
Torque promedio neto de salida	%				x

Para el caso de los parámetros correspondientes a modos de conducción (proporción de aceleración, proporción de desaceleración, proporción de ralentí y proporción de crucero), se realizó un análisis del estado del arte respecto a la definición de estos parámetros para su respectivo calculo. Este análisis se presenta en la Tabla 4.5.

Tabla 4.5, *Revisión de la literatura acerca de la definición de los modos de conducción*

Referencia	Ralentí	Crucero	Aceleración	Desaceleración
Liu et al., 2020	$v < 0.139 \text{ m/s}$ $-0.15 \text{ m/s}^2 < a < 0.15 \text{ m/s}^2$	$v \geq 0.139 \text{ m/s}$ $-0.15 \text{ m/s}^2 < a < 0.15 \text{ m/s}^2$	$a \geq 0.15 \text{ m/s}^2$	$a \leq -0.15 \text{ m/s}^2$
Jia et al., 2021	$v < 0.139 \text{ m/s}$ $-0.15 \text{ m/s}^2 < a < 0.15 \text{ m/s}^2$	$v \geq 0.139 \text{ m/s}$ $-0.15 \text{ m/s}^2 < a < 0.15 \text{ m/s}^2$	$a \geq 0.15 \text{ m/s}^2$	$a \leq -0.15 \text{ m/s}^2$
Yuan et al., 2021	$v < 0.278 \text{ m/s}$ $-0.1 \text{ m/s}^2 < a < 0.1 \text{ m/s}^2$	$v \geq 0.278 \text{ m/s}$ $-0.1 \text{ m/s}^2 < a < 0.1 \text{ m/s}^2$	$a \geq 0.1 \text{ m/s}^2$	$a \leq -0.1 \text{ m/s}^2$
Chauhan et al., 2020	$v < 1.39 \text{ m/s}$ $-0.1 \text{ m/s}^2 \leq a \leq 0.1 \text{ m/s}^2$	-	$a > 0.1 \text{ m/s}^2$	$a < -0.1 \text{ m/s}^2$
Desineedi et al., 2020	$v = 0 \text{ m/s}$	$v > 1.39 \text{ m/s}$ $-0.1 \text{ m/s}^2 < a < 0.1 \text{ m/s}^2$	$a \geq 0.1 \text{ m/s}^2$	$a \leq -0.1 \text{ m/s}^2$
Yuhui et al., 2019	$v = 0 \text{ m/s}$	$v > 0 \text{ m/s}$ $-0.1 \text{ m/s}^2 < a < 0.1 \text{ m/s}^2$	$a > 0.1 \text{ m/s}^2$	$a < -0.1 \text{ m/s}^2$
Yang et al., 2019	$v < 0.278 \text{ m/s}$ $-0.14 \text{ m/s}^2 < a < 0.14 \text{ m/s}^2$	$v \geq 0.278 \text{ m/s}$ $-0.14 \text{ m/s}^2 \leq a \leq 0.14 \text{ m/s}^2$	$a > 0.14 \text{ m/s}^2$	$a < -0.14 \text{ m/s}^2$
Ho et al., 2014	-	$-0.28 \text{ m/s}^2 \leq a \leq 0.28 \text{ m/s}^2$	$a > 0.28 \text{ m/s}^2$	$a < -0.28 \text{ m/s}^2$

La definición utilizada en el cálculo de cada uno de los modos de conducción se utiliza basado en los trabajos realizados por Y. Liu et al., 2020, debido a que específicamente el CTC desarrollado se realizó para un país completo en este caso China.

Estos parámetros son calculados para cada uno de los microciclos y es la base la base para el algoritmo PCA y el método k-means. Así mismo, los ciclos construidos son comparados respecto a estos parámetros, basado en el promedio del error medio tal como se define en la Ecuación 4.1.

$$AE = 100 \frac{\sum_{i=1}^n \left| \frac{\hat{P}_i - P_i}{\hat{P}_i} \right|}{n} \quad \text{Ecuación 4.1}$$

Donde  $\hat{P}_i$  es el parámetro promedio ponderado de todos los microciclos,  $P_i$  es el parámetro  $i$ , del CTC construido, y  $n$  es la cantidad de parámetros.

#### 4.2.2. Preprocesamiento de datos el dispositivo IVDR

Para la realización del preprocesamiento de datos, se propone la siguiente metodología, donde se tienen en consideración todos los pasos necesarios para obtener datos listos para su procesamiento.

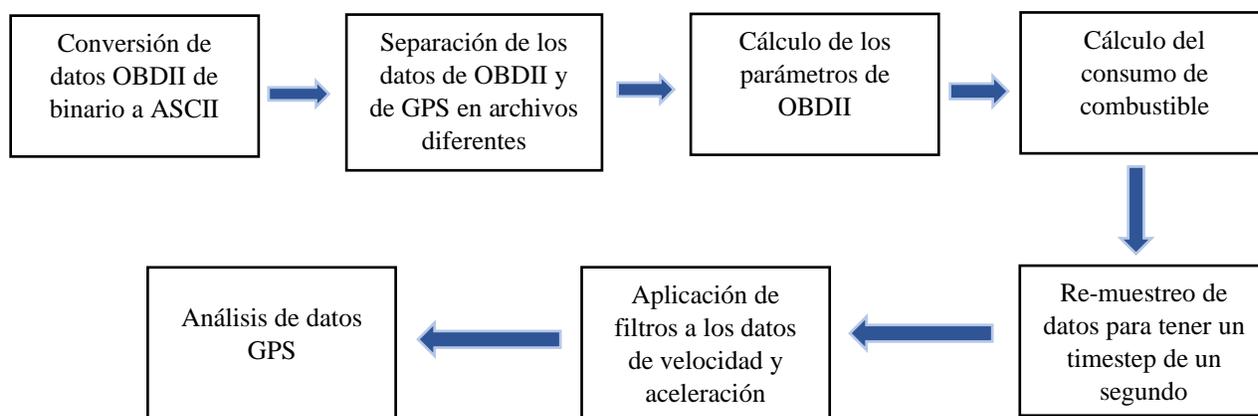


Figura 4.10. Metodología para el Desarrollo del preprocesamiento de datos

##### ***Conversión de datos OBDII de binario a ASCII y separación de datos***

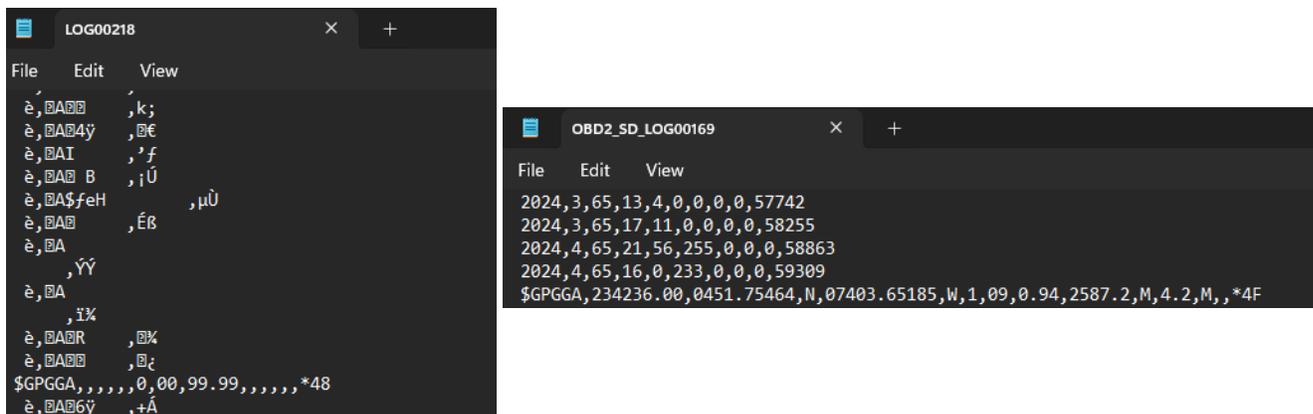
Los datos recolectados por el dispositivo IVDR corresponden a dos tipos de datos, los datos del vehículo recolectados a través del protocolo OBDII y los datos de GPS.

Los datos obtenidos por el vehículo vienen en formato binario y no en formato ASCII (*American Standard code for Information Interchange*, por sus siglas en inglés), esto con el objetivo de ahorrar tiempo de procesamiento del microcontrolador. Sin embargo, los datos en este formato no pueden ser procesados, por lo que es necesario desarrollar un programa en el software MATLAB para convertir estos datos de Binario a ASCII. De igual manera los datos de GPS son entregados en el formato GPGGA (*Global Positioning System Fix Data*, por sus siglas en inglés), los cuales deben ser analizados para conocer la posición donde los vehículos estuvieron.

Dado que los datos tanto de OBDII y GPS se encuentran en el mismo documento, el programa de preprocesamiento primero debe convertir los datos de OBDII de binario a ASCII para su

respectiva manipulación, y así mismo separar en diferentes documentos los datos de OBDII y GPS para su análisis.

El algoritmo utilizado para la conversión de datos de OBDII de binario a ASCII es el mismo utilizado en trabajos previos (Echeverry Mejía, 2018), sin embargo, se modifica para que este algoritmo funcione como función de un script general, y así mismo separe los datos de GPS y OBDII en archivos diferentes.



*Figura 4.11.* Comparación entre datos de OBDII en formato binario (izquierda) y datos de OBDII en formato ASCII (derecha)

En la Figura 4.11, se presenta un archivo del dispositivo IVDR antes de que los datos de OBDII pasen por el proceso de conversión de formato binario a ASCII, y el resultado después de la conversión, junto con los mensajes de GPS identificados por el inicio del mensaje como \$GPGGA. Después de que los datos de OBDII se encuentren en formato ASCII, son separados de los datos GPS y almacenados en diferentes espacios de memoria para su posterior análisis.

### **Cálculo de los parámetros de OBDII**

Como se analizó en las secciones 3.2.3 y 4.1.3, los datos obtenidos por los parámetros deben ser preprocesados para obtener los valores reales medidos, como se analizó el ejemplo de las revoluciones por minuto del motor. Esto debido a que los datos son almacenados en diferentes bytes y cada parámetro tiene una fórmula característica para ser calculada (Society of Automotive Engineers, 2007).

A partir de los parámetros disponibles para el cálculo de los parámetros de interés en el análisis del CTC observados en la

Tabla 4.3, en la Tabla 4.6 se presentan sus respectivas fórmulas para su cálculo.

Tabla 4.6. Descripción de las fórmulas para el cálculo de los parámetros de OBDII

PIDs solicitados			Cálculo de Parámetros		
No.	PID Hexadecimal	Sensor	No. de bytes	Datos entregados	Formula
1	04	Carga del motor	1	Carga del motor (%)	$ML = \frac{A}{2.55}$
2	0B	MAP (Mass Air Pressure)	1	MAP (kPa)	-
3	0C	RPM	2	RPM	$RPM = \frac{256A + B}{4}$
4	0D	Velocidad	1	Velocidad del vehículo (km/h)	-
5	0F	IAT (Intake Air Temperature)	1	Temperatura del aire del colector de admisión (°C)	$IAT = A - 40$
6	11	TPS (Throttle Position Sensor)	1	Posición del acelerador (%)	$TPS = \frac{A}{2.55}$
7	15	O2 -2 (Sensor de oxígeno 2)	1	A: Voltaje del sensor de oxígeno(V) B: Ajuste de combustible a corto plazo (%)	$VSO = \frac{A}{200}$ $ACC = \frac{B}{1.28 - 100}$
8	24	Relación Aire - Combustible en el primer sensor de oxígeno	4	A,B : Relación aire combustible ( $\lambda$ ) C,D: Voltaje del sensor de oxígeno (V)	$\lambda = \frac{256A + B}{32768}$ $VSO = \frac{256C + D}{8192}$
9	33	Presión Atmosférica	1	Presión barométrica absoluta (kPa)	-
10	49	Posición del pedal	1	Posición del pedal del acelerador (%)	$PPA = \frac{A}{2.55}$
12	10	MAF (Mass Air Flow)	2	Flujo de aire a la entrada del colector de admisión (gr/seg)	$MAF = \frac{256A + B}{100}$

Cada parámetro solicitado tiene una cantidad específica de bytes debido a los valores máximos que pueden representar: el primer byte entregado se denomina con la letra A, el segundo con la letra B, y así sucesivamente. El algoritmo desarrollado tiene la capacidad de leer los datos de OBD2 y leer el respectivo PID en decimal, para así realizar el cálculo del parámetro en cuestión y almacenar cada parámetro junto con su timestamp. Su validación se presentará en la sección 5.1.3, realizando el cálculo general del timestamp en un vector para independiente por cada parámetro para su posterior análisis.

### *Cálculo del consumo de combustible*

El consumo de combustible es un parámetro bastante importante, ya que esta está relacionado con la cantidad de contaminantes liberados al ambiente, así como con la eficiencia energética de los vehículos a analizar. Es necesario que CTC construido tenga en consideración esta variable. Sin embargo, no es posible realizar la medición directa en los vehículos experimentales debido a que la mayoría de los vehículos comerciales no cuentan con sensores de flujo de combustible, por lo que es necesario calcular el flujo de combustible que entra al motor mediante parámetros alternos.

Una de las formas de calcular el consumo de combustible es mediante el sensor MAF, el cual mide el flujo de aire en la entrada del colector de admisión, el cual está relacionado directamente con el flujo de combustible mediante la Ecuación 4.2 (Moradi & Miranda-Moreno, 2020)

$$FC_t = \frac{MAF_t}{\lambda AF_{ideal}} \quad \text{Ecuación 4.2}$$

Donde  $FC_t$  representa el consumo de combustible en gramos/segundo,  $MAF_t$  el valor del sensor MAF en gramos/segundo,  $AF_{ideal}$  representa la relación de aire combustible ideal, la cual es de 14.7, y el valor  $\lambda$  (lambda).  $\lambda$  relaciona la relación aire combustible real, con la relación aire combustible ideal, teniendo valores entre 0 y 1.9. Si el valor de lambda es menor a uno, se tiene una mezcla rica en combustible, y si el valor mayor a uno se tiene una mezcla pobre en combustible.

Sin embargo, muchos vehículos comerciales no cuentan con el sensor MAF, sino que cuentan con el sensor MAP, como en el caso de los vehículos experimentales donde cada uno de ellos cuenta con un sensor diferente (Nissan Versa 2013 – MAF, Ford Fiesta 2017 – MAP). Por este motivo es necesario recurrir a una ecuación diferente para realizar el cálculo de consumo de combustible mediante el sensor MAP. En la Ecuación 4.3 (Moradi & Miranda-Moreno, 2020), se presenta como calcular el valor del flujo de aire o el valor MAF con base en el sensor MAP para que una vez se tenga este valor se pueda reemplazar en la Ecuación 4.2.

$$MAF_t = \frac{RPM}{120} * \frac{MAP_t}{IAT_t} * \frac{VE}{100} * ED * \frac{MM}{R} \quad \text{Ecuación 4.3}$$

Donde:

- $MAF_t$ : Flujo de aire en el colector de admisión (gr/seg)
- $MAP_t$ : Presión del aire en el colector de admisión (kPa)
- $IAT_t$ : Temperatura del aire en el colector de admisión (K)
- $VE$ : Eficiencia volumétrica (65%)
- $ED$ : Desplazamiento del motor (1.6 L)
- $MM$ : Masa molar del aire (28.97 gr/mol)
- $R$ : Constante ideal de los gases (8.314 J/mol\*K)

En ambos casos el subíndice  $t$ , representa los datos tomados en tiempo real a través del dispositivo IVDR. En ambos métodos, para calcular el consumo de combustible es necesario conocer el valor  $\lambda$  (lambda). Para determinar este valor, existen dos formas de realizarlo, mediante los datos obtenidos por el dispositivo IVDR. El primero, consiste en solicitar directamente este valor mediante el PID 0x24, y calcularlo a través de la fórmula descrita en la Tabla 4.6, mediante los dos primeros bytes entregados por este PID. El segundo método consiste en establecer una relación entre el voltaje del sensor de oxígeno obtenido mediante el PID 0x15, a través del cual se relaciona el voltaje obtenido con el valor lambda.

Esta relación se establece mediante el análisis de la función se establece al graficar el voltaje en función de lambda, para esto es posible analizarlo mediante una función logística decreciente, mediante la Ecuación 4.4.

$$V_t = V_{min} + \frac{V_{max} - V_{min}}{1 + e^{\frac{\lambda-1}{0.1}}} \quad \text{Ecuación 4.4}$$

Donde:

- $V_t$ : Voltaje medido por el PID 0x015
- $V_{min}$ : Voltaje mínimo medido
- $V_{max}$ : Voltaje máximo medido

Esta ecuación, que es directamente deducida a partir del comportamiento entre estos dos valores, establece el valor medio de lambda como 1 y una tasa de decrecimiento de 0.1. Reemplazando  $V_{min}$ , como 0 V,  $V_{max}$  como 0.9 V y despejando Lambda, se obtiene la Ecuación 4.5, la cual establece la relación entre el valor lambda y el voltaje entregado por el PID 0x015 obtenido mediante el dispositivo IVDR.

$$\lambda = \text{Ln}\left(\frac{0.9}{V_t} - 1\right) 0.1 + 1 \quad \text{Ecuación 4.5}$$

Mediante la ecuación descrita, o el valor calculado a través del PID 0x24, es posible realizar el cálculo del consumo de combustible del vehículo en tiempo real, para que al igual que los demás parámetros sean re muestreados. Cabe aclarar que los resultados de la Ecuación 4.5, fueron validados en el vehículo experimental Nissan Versa – 2013, debido a que en este vehículo se midieron ambos PIDs y se compararon los resultados obtenidos, encontrando resultados similares en los valores del valor lambda.

En la se presenta a manera de diagrama de flujo el proceso realizado por el algoritmo para el cálculo del consumo de combustible.

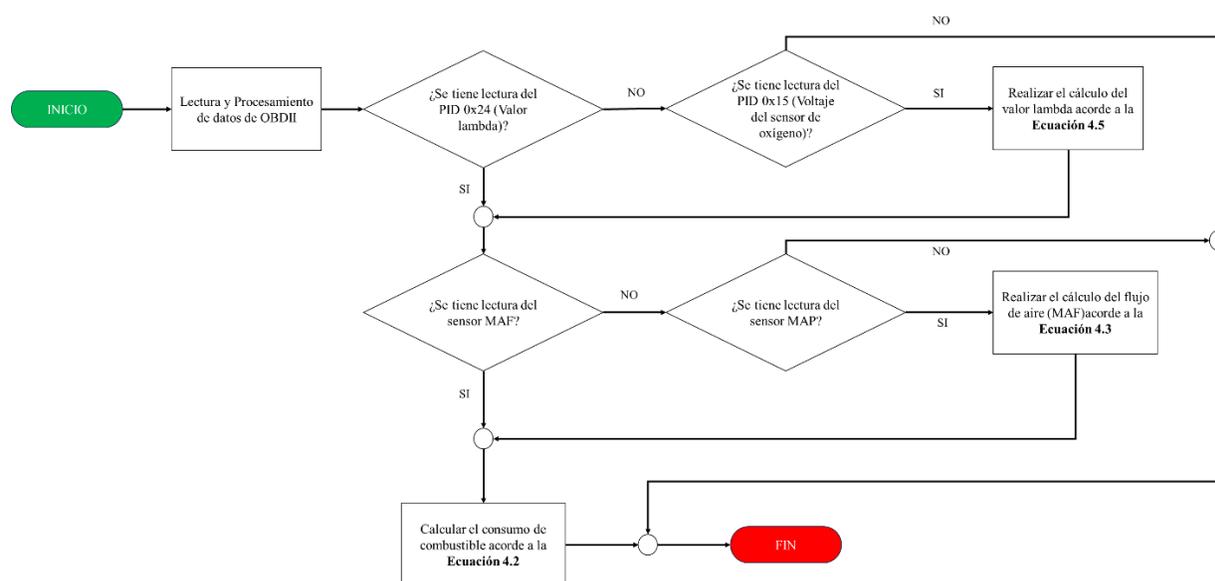


Figura 4.12. Metodología para el cálculo del consumo de combustible

### ***Remuestreo de datos para tener un timestep de un segundo***

En la sección 4.1.1, se evidenció que el dispositivo IVDR tiene diferentes velocidades de transmisión de datos, 10 mseg, 100 mseg, y 1 seg. Durante el desarrollo del proyecto se seleccionó una velocidad de transmisión de datos de 1 dato cada 10 mseg, teniendo así una cantidad de 100 datos por segundo. Esto en específico para el vehículo experimental Nissan Versa 2013, ya que el vehículo experimental modelo Ford Fiesta 2017, solamente permite adquirir datos a una tasa de 1 dato cada 30 mseg, es decir ~33 datos por segundo. Teniendo en consideración que para cada

vehículo se solicitaron 9 parámetros, se tendrá al menos 3 datos de cada parámetro por segundo. Sin embargo, la mayoría de los parámetros no presentan una variabilidad tan alta.

Por lo anteriormente mencionado, es de suma importancia realizar un remuestreo de los datos, con el objetivo de tener muestras de los parámetros cada segundo. Para esto, se realiza una interpolación mediante dos métodos, el método spline el cual asegura una mejor aproximación de los datos acorde al timestep requerido, y en dado caso que alguno de los datos interpolados presente un NaN (*Not a number*, por sus siglas en ingles), se procede a utilizar un método menos estricto, en este caso una interpolación cubica. La interpolación se realiza, tomando el menor valor en segundos en el cual se tenga el dato de alguno de los parámetros, se realiza un vector espaciado a un segundo hasta el valor en segundos anteriormente mencionado, y se solicita la interpolación con los métodos descritos.

### ***Aplicación de filtros a los datos de velocidad y aceleración***

Se realiza un filtrado y suavizado de datos, específicamente para los datos de velocidad y aceleración, de los cuales se obtendrán la mayoría de los parámetros de influencia en el CTC. Este filtrado y suavizado, sigue la misma metodología propuesta por Duran & Earleywine y Nguyen et al. (Duran & Earleywine, 2012; Nguyen et al., 2020).

Donde se realizaron los siguientes pasos:

- Remoción de pequeñas variaciones de velocidad cero cuando el vehículo se encuentra en reposo (motor en ralentí)
- Reemplazo de valores de velocidad cero falsos
- Reemplazo de valores anormales de aceleración y desaceleración
- Filtrado de los datos de velocidad mediante un filtro Kalman

***Remoción de pequeñas variaciones de velocidad cero cuando el vehículo se encuentra en reposo.*** Para la remoción de pequeñas variaciones en la velocidad cuando el vehículo se encuentra en reposo, se analizó el modo de conducción establecido para el estado de ralentí especificado en la Tabla 4.5. De acuerdo con esto, si la velocidad es menor a 0.139 m/s, el punto de velocidad es reemplazado por un valor de 0 m/s. Estos errores en la medición cuando el vehículo se encuentra en reposo con el motor encendido se pueden dar por errores en el sensor de medición y/o por la interpolación realizada en pasos previos.

**Reemplazo de valores de velocidad cero falsos.** Es posible tener valores de velocidad cero falsos debido a un error en la medición del sensor del vehículo. Por este motivo, si se tiene una señal de velocidad con valor de 0 m/s y los datos adyacentes previo y posterior son diferentes de 0 m/s, se genera un vacío en el vector de datos el cual será completado con el algoritmo de Selesnick (Selesnick, 2022). Este algoritmo permite completar datos faltantes mediante el método de mínimos cuadrados.

**Reemplazo de valores anormales de aceleración y desaceleración.** Posterior al análisis de los valores de velocidad cero falsos, se realizó el cálculo de la aceleración acorde a la definición de esta, como el cambio en el tiempo de la velocidad tomando la ecuación de aceleración instantánea como se puede ver en la Ecuación 4.6 (Nguyen et al., 2020) .

$$a(t) = \frac{V_i - V_{i-1}}{3.6Ts} \quad \text{Ecuación 4.6}$$

Donde  $V_i$  y  $V_{i-1}$  corresponden a la velocidad en el punto tiempo actual y el anterior respectivamente, los cuales se encuentran en km/h;  $Ts$  corresponde al paso de tiempo, donde acorde a la interpolación realizada corresponde a 1 segundo; y el valor de 3.6, ubicado en el denominador corresponde al factor de conversión para que el valor de la aceleración este en  $\text{m/s}^2$ .

Es posible que, durante la medición en los recorridos realizados, se pudiesen realizar mediciones incorrectas respecto a los cambios de velocidad presentando aceleraciones anormales, o debido a los tratamientos de preprocesamiento anteriormente realizados. Por este motivo, el algoritmo analiza si se presentan aceleraciones anormales acorde a los límites de aceleración propuestos por Duran & Earleywine (Duran & Earleywine, 2012), siendo el límite inferior  $-7.875 \text{ m/s}^2$  ( $-17.5 \text{ mph/s}$ ), y el límite superior de  $7.875 \text{ m/s}^2$  ( $17.5 \text{ mph/s}$ ) para vehículos livianos. Si la aceleración calculada excede alguno de límites establecidos, los puntos de velocidad correspondientes a dichas aceleraciones son eliminados, y posteriormente completados con el algoritmo de Selesnick. Este proceso se realiza hasta que no haya aceleraciones anormales dentro del recorrido.

**Filtrado de los datos de velocidad mediante un filtro Kalman.** Con el objetivo de suavizar y quitar el ruido de los datos de velocidad obtenidos, junto con los preprocesamientos se realiza la aplicación de un filtro Kalman, el cual es una técnica comúnmente utilizada para estos fines (Jun

et al., 2006; Kim & Bang, n.d.). Basado en teoría de control, este filtro tiene dos procesos, un proceso de predicción (actualización de tiempo) y proceso de corrección (actualización de las mediciones). Teniendo en consideración que los datos de velocidad se encuentran a 1Hz, las ecuaciones del filtro Kalman para el suavizado de velocidad son las siguientes (Jun et al., 2006; Nguyen et al., 2020):

Proceso de predicción:

$$\hat{x}_k^- = \hat{x}_{k-1} \quad \text{Ecuación 4.7}$$

$$P_k^- = P_{k-1} + W \quad \text{Ecuación 4.8}$$

Proceso de corrección:

$$K_k = P_k^- (P_k^- - V)^{-1} \quad \text{Ecuación 4.9}$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - \hat{x}_k^-) \quad \text{Ecuación 4.10}$$

$$P_k = (I - K_k) P_k^- \quad \text{Ecuación 4.11}$$

Donde:

- $k$ : timestep;
- $\hat{x}_{k-1}$  y  $P_{k-1}$ : Predictor inicial y ruido de error inicial, respectivamente;
- $W$ : Variación del error de la predicción, el cual es un ruido Gaussiano,  $W \sim N(0, Q)$ ;
- $K_k$ : Matriz de ganancia del filtro Kalman;
- $z_k$ : Datos observados (velocidad medida);
- $P_k$ : Varianza del error modificado;
- $\hat{x}_k$ : Predictor posterior,
- $V$ : Variación del error de la medición, el cual es un ruido Gaussiano,  $V \sim N(0, R)$  e;
- $I$ : Lo cual denota la matriz identidad.

Acorde a la teoría del filtro Kalman, la variación del error de la medición  $V$  y la variación del error de la predicción  $W$ , deben comportarse como una distribución normal con media 0 y con desviación estándar  $R$  y  $Q$ , respectivamente, los cuales denotan el ruido de la medición y el ruido del proceso. Acorde con Jun et al. (Jun et al., 2006), ambos valores son iguales a 0.25 mph (0.402 km/h).

Después de aplicado el filtro Kalman los datos filtrados corresponden a aquellos guardados en el vector  $\hat{x}_k$ , a los cuales se les aplican los pasos de filtrado aplicado antes del filtro Kalman, con el fin de evitar tener valores anormales generados por el filtro Kalman. Los valores de velocidad y aceleración que se obtienen de todos los procesos de filtrado son aquellos que se utilizan para la construcción del CTC de la zona Bogotá región.

### *Análisis de datos GPS*

Los datos separados de GPS de los archivos originales son almacenados en un archivo diferente y con estos datos se realizará una gráfica donde se visualicen los recorridos realizados. Como se mencionó anteriormente, los datos GPS son almacenados en el formato GPGG, el cual almacena los datos de posición satelital de manera específica. En la Tabla 4.7 se presenta un ejemplo de un mensaje de GPS y como se encuentran organizados los datos dentro de este mensaje.

Tabla 4.7. *Organización de un mensaje GPS*

<b>Mensaje de ejemplo</b>		
\$GPGGA,220418.00,0451.91698,N,07403.39842,W,1,08,1.10,2596.2,M,4.2,M,,*4C		
<b>Nombre de la variable</b>	<b>Parte del mensaje de ejemplo</b>	<b>Descripción</b>
Identificador	\$GPGGA	Global Positioning System Fix Data
Hora UTC	220418	22:04:18 (hhmmss.ss)
Latitud	0451.91698,N	4° 51.91698' N
Longitud	07403.39842,W	74° 3.39842 W
Calidad fijada:		
- 0 = Invalido		
- 1 = GPS fijado	1	Los datos son obtenidos de GPS
- 2 = DGPS fijado		
Número de satélites	08	8 satélites están siendo contactados
Dilución de la precisión horizontal	1.1	Exactitud relativa de la posición horizontal
Altitud	2596.2, M	2596.2 metros sobre el nivel del mar
Altura del geoide sobre el elipsoide WGS84	4.2, M	4.2 metros
Tiempo desde la última actualización del DGPS	-	Sin última actualización
Id de la estación de referencia del DGPS	-	Sin estación
Revisión de suma	*4C	Suma utilizada por el programa para la revisión de errores de transmisión

Si bien, se tiene una gran variedad de datos dentro de los mensajes de GPS, los únicos datos que se tienen en consideración son los datos de longitud y latitud, con el fin de verificar la posición del vehículo y visualizar los recorridos realizados durante los recorridos de recolección de datos.

La visualización de la posición de los vehículos se realiza a través de una función en MATLAB, la cual toma los datos de GPS, y mediante la función `geoplot`, utilizando la visualización de calles se realiza el análisis de la posición satelital de los vehículos durante los recorridos.

Todo el preprocesamiento de datos presentado en esta sección fue desarrollado por completo en MATLAB, sin el uso de librerías específicas con excepción de las ya mencionadas, en un script general elaborando cada etapa del preprocesamiento de datos como una función. Este algoritmo solicita los archivos recolectados del dispositivo IVDR, y entrega los datos de OBD2 en archivos con la extensión `.csv`. En la primera columna de archivo `.csv` se almacena el tiempo en segundos, en la segunda columna la velocidad procesada en  $m/s$ , en la tercera columna la aceleración en  $m/s^2$ , en la cuarta columna la carga del motor, y en la quinta columna el consumo de combustible en  $gr/s$ . Estos datos son los principalmente requeridos para la construcción del CTC. Los PIDs no procesados, como la posición del pedal y la presión atmosférica, fueron almacenados para futuros estudios. Así mismo se genera un archivo con extensión `.txt`, con todos los datos de longitud y latitud obtenidos de los datos GPS.

#### **4.2.3. Evaluación comparativa de los métodos de análisis de Aprendizaje automático**

Se realiza la aplicación de los métodos de aprendizaje automático y la combinación de cadenas de Márkov y el método `k-means` utilizando el método PCA para la reducción de dimensionalidad para cada una de las metaheurísticas mencionadas y sus métodos de inicialización, realizando una comparación estadística entre los métodos utilizados. De esta manera, se validará si el método utilizado tiene efecto sobre el error medio y el tiempo de ejecución del algoritmo. Validando los supuestos de normalidad, independencia y homocedasticidad para los resultados de cada uno de estos algoritmos tanto en error como en tiempo. Para esto se utilizan las siguientes pruebas estadísticas (Khadse et al., 2020):

- Normalidad
  - Shapiro - Wilks test
  - Kolmogorov -Smirnov with Lilliefors correction test
  - D'Agostino - Pearson test
  - Anderson - Darling test
- Homocedasticidad
  - Fligner-Killen test
  - Leven's test
- Independencia
  - Durbin Watson test

Esto con el objetivo de validar los supuestos se realizan pruebas paramétricas o no paramétricas. En el caso de poder realizar pruebas paramétricas, se realiza un Análisis de la varianza (ANOVA, por sus siglas en ingles) y se realiza una prueba paramétrica Post-Hoc de Tukey. Si los supuestos no son validados, se realiza una prueba no paramétrica de Kruskal-Wallis y una prueba Post-Hoc no paramétrica de comparación de suma exacta de rangos pareados de Wilcoxon con correcciones de Holm y Bonferroni. Esta validación se realizó con bases de datos del Laboratorio de Energías renovables de Estados Unidos (NREL, por sus siglas en ingles), para El Paso, San Antonio y Houston (The National Renewable Energy Laboratory (NREL), 2015).

#### **4.2.4. Definición del ciclo típico de conducción para Bogotá - Región.**

Para la construcción del CTC correspondiente a la zona Bogotá región, se construyeron o seleccionaron diferentes ciclos de acuerdo con los diferentes métodos propuestos, introduciendo un parámetro más a los descritos anteriormente. Este parámetro adicional es una variable dicotómica (0 – Ford Fiesta 2017, 1 – Nissan Versa 2013), la cual corresponde al vehículo utilizado durante la medición, con el objetivo de analizar la influencia de este sobre la construcción del CTC. De igual manera, se toma la metaheurística con el mejor desempeño acorde a los resultados obtenidos al desarrollar la sección anterior y una combinación de la metaheurística seleccionada junto con el método estadístico MWD-CP.

En general, los métodos utilizados para la construcción del CTC para la zona Bogotá región son los siguientes, obteniendo un ciclo por cada uno de estos métodos:

- MWD-CP – sin la variable dicotómica del tipo de vehículo
- MWD-CP – con la variable dicotómica del tipo de vehículo
- Metaheurística seleccionada – sin la variable dicotómica del tipo de vehículo
- Metaheurística seleccionada – con la variable dicotómica del tipo de vehículo
- Metaheurística seleccionada en combinación con el método MWD-CP – sin la variable dicotómica del tipo de vehículo
- Metaheurística seleccionada en combinación con el método MWD-CP – con la variable dicotómica del tipo de vehículo

El ciclo construido con base en la metaheurística utilizada en combinación del método MWD-CP consiste en la construcción del CTC con base en el algoritmo k-means, el algoritmo PCA y el método de cadenas de Markov, generando la TPM para la construcción del CTC. Sin embargo, la Ecuación 4.1 modificada al añadir los pesos asignados a cada parámetro mediante el método RWA con base en que la variable a ser representada sea el consumo de combustible. Esto genera un CTC más enfocado tanto a la representatividad de los contaminantes como al consumo de combustible. La Ecuación 4.12 presenta las modificaciones realizadas al error medio para la construcción de este CTC.

$$AE = 100 \sum_{i=1}^n W_i \left| \frac{\hat{P}_i - P_i}{\hat{P}_i} \right| \quad \text{Ecuación 4.12}$$

Donde  $W_i$  corresponde a los pesos calculados mediante el método RWA presentado en la sección 3.3.1 de los parámetros utilizados para la construcción del ciclo CTC, con base en el análisis de la variable consumo de combustible. El denominador ya no es asignado como el número de parámetros, ya que este valor asigna un peso equivalente a todos los parámetros.

Finalmente el ciclo con el menor error medio, tanto de la ecuación Ecuación 4.11 y Ecuación 4.12, se selecciona como el ciclo de la zona Bogotá región, teniendo en consideración que se tienen unos CTCs que representan de manera general la base de datos con base a todos los parámetros, y

otros que representan de manera general la base datos con el enfoque de representar mejor el consumo de combustible, que está directamente relacionado con las emisiones generadas.

#### 4.2.5. Comparación del CTC encontrado con los ciclos estándar internacionales

Se realiza una comparación respecto al error medio de los CTCs calculados y los CTCs estándar, entre estos el ciclo FTP-75, el ciclo WLTC (con cada una de sus clases), y el ciclo (UC)/LA 92 (*California Unified Cycle*, por sus siglas en inglés) (The National Renewable Energy Laboratory, n.d.) . Para esta comparación, se desarrolla un análisis de representatividad de la base de datos adquirida respecto a los ciclos construidos y los ciclos estándar.

El algoritmo de preprocesamiento y construcción del CTC fue desarrollado en el software MATLAB® (The Mathworks Inc, 2022a), junto con el Statistics and Machine learning toolbox de este software (The Mathworks Inc, 2022b), mientras que los análisis estadísticos fueron realizados en el software R (R Core Team, 2022). Se utilizó un computador con las siguientes características: Procesador Intel® Core™ i7 - 9700K CPU @ 3.6 GHz, con una memoria RAM de 16GB con 8 núcleos.

#### 4.3. Desarrollo metodológico del tercer objetivo específico

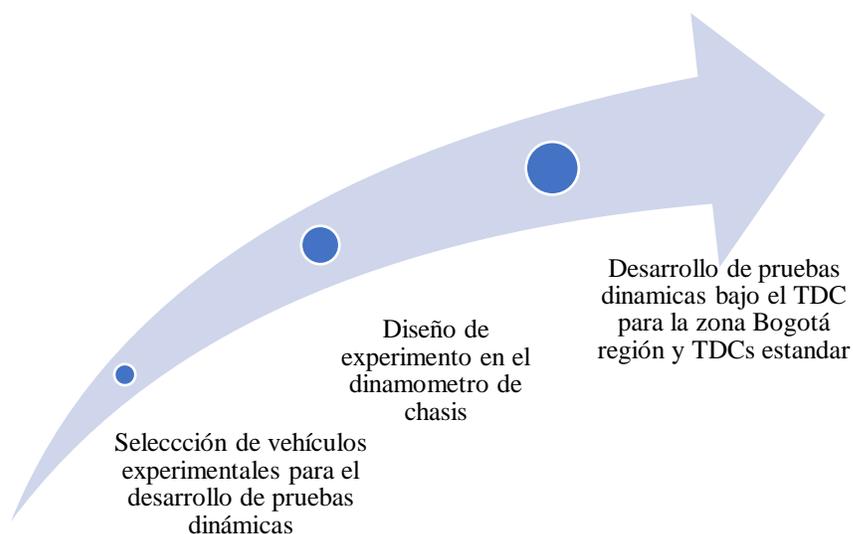


Figura 4.13. Proceso metodológico para el desarrollo del tercer objetivo específico

#### **4.3.1. Selección de vehículos experimentales para el desarrollo de pruebas dinámicas**

Debido a que el objetivo principal de este proyecto consiste en validar la metodología para la evaluación medioambiental del parque automotor liviano de la zona Bogotá región, se selecciona el vehículo experimental Nissan Versa – 2013, cuyas especificaciones se encuentran en la Tabla 4.2. Este vehículo se selecciona debido a su disponibilidad para la realización de las pruebas.

#### **4.3.2. Diseño de experimento en dinamómetro de chasis**

El diseño de experimentos para el desarrollo de las pruebas consiste en una serie de experimentos por triplicado, realizando tres pruebas en el dinamómetro de chasis por cada CTC a analizar, los cuales incluyen:

- Ciclo desarrollado para la zona Bogotá región – CTC Bogotá
- FTP-75
- WLTC – Clase 1

Sin embargo, debido a la disponibilidad del vehículo y de equipos de medición, solamente fue posible realizar una medición de cada CTC. Además, se tomaron datos de contaminantes gaseosos generados como como  $\text{NO}_x$  ( $\text{NO}_x$  y  $\text{NO}$ ),  $\text{CO}_2$ ,  $\text{CO}$  y  $\text{THC}$  (Hidrocarburos no quemados) y el respectivo consumo de combustible (gal/km) para el ciclo desarrollado para la zona Bogotá región (CTC Bogotá) con el equipo analizador de gases descrito en la sección 4.3.4. Para todos los CTC estándar evaluados se realizaron mediciones de consumo de combustible, para los CTCs estándar se obtuvieron los datos a través del dispositivo IVDR.

Las normas que se utilizaran para las pruebas en el dinamómetro de chasis son las siguientes: SAE J2951(Society of Automotive Engineers (SAE), 2014), SAE J1263(Society of Automotive Engineers (SAE), 2010), SAE J254 (Society of Automotive Engineers (SAE), 2011).

#### **4.3.3. Desarrollo de pruebas dinámicas bajo el CTC Bogotá y CTCs estándar.**

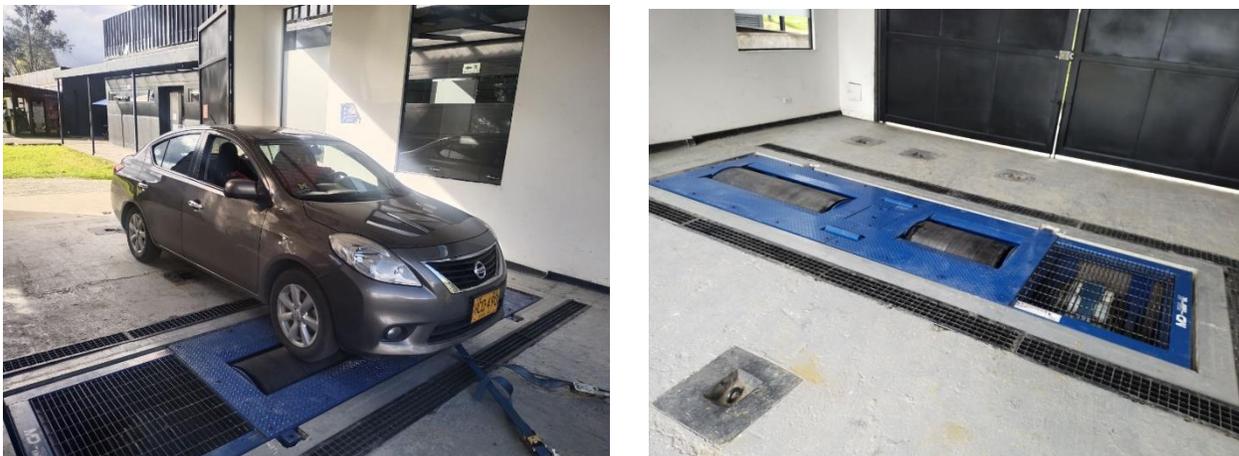
Las pruebas se desarrollarán en el dinamómetro de chasis MD1100-SE, el cual se encuentra en los laboratorios de la facultad de ingeniería ubicados en FabLab, acorde a las normas de la SAE

anteriormente mencionadas. El procedimiento de las pruebas se llevó a cabo de la siguiente manera:

1. Posicionamiento del vehículo sobre los rodillos del dinamómetro con los rodillos frenados.
2. Anclaje del vehículo mediante eslingas de posicionamiento desde el chasis del vehículo a los anclajes del dinamómetro.
3. Posicionamiento de topes para evitar el deslizamiento del vehículo
4. Desactivación de freno al dinamómetro
5. Pruebas de validación de la sujeción del vehículo a velocidades bajas
6. Conexión del analizador de gases al tubo de escape del vehículo (medición de gases de escape).
7. Conexión del dispositivo IVDR para el almacenamiento de datos de consumo de combustible.
8. Selección de CTC a desarrollar en la prueba mediante el software PowerDyne, del dinamómetro.
9. Realización de la prueba
10. Validación de almacenamiento de datos.
11. Frenado de rodillos
12. Desanclaje del vehículo
13. Salida del vehículo del laboratorio

Es importante recalcar que todos los CTCs tanto los construidos como los estándar fueron introducidos en el formato requerido por el software PowerDyne para su lectura y realización de la prueba.

El desarrollo de la prueba consiste en que el conductor siga la curva de velocidad del CTC mientras el dinamómetro de chasis monitorea los valores de velocidad, carga e inercia de los rodillos, para así validar el correcto seguimiento de las curvas del CTC.



*Figura 4.14.* Izquierda: Vehículo experimental durante una prueba dinámica. Derecha: Dinamómetro de chasis Mustang Dyno MD-110SE

#### **4.3.4. Medición de diferentes contaminantes con el analizador de gases.**

Se realiza la medición de diferentes gases mediante el analizador de gases portable par SYNC FLEX PNC, el cual realizará las mediciones de cuatro gases de combustión ( $O_2$ ,  $CO_2$ ,  $CO$ ,  $NO_x$  ( $NO$  y  $NO_2$ ) y  $THC$ ), así como consumo de combustible. Estos gases son medidos directamente en el tubo de escape del vehículo mediante una sonda la cual recolecta muestras donde son transportadas a la unidad analizadora, la cual reporta la concentración de los gases con un periodo de un segundo entre cada reporte, acorde a las mediciones realizadas. Esta actividad se realizará de manera simultánea con el desarrollo de pruebas en el dinamómetro de chasis. En la

Figura 4.15 y Figura 4.16 se presentan las especificaciones generales de este analizador de gases (3DATX, 2023).

GasMOD™ Cartridge - N	3 Electrode Electro-Chemical	
	Nitric Oxide (NO)	Nitrogen Dioxide (NO <sub>2</sub> )
Linear Measurement Range	0-5000ppm	0-300ppm
T <sub>90</sub> Response Time	<5 seconds	<35 seconds
Resolution	1-2ppm	0.1ppm
Repeatability	2% of signal	2% of signal

Figura 4.15. Características de los sensores de NO<sub>x</sub> del analizador de gases.

GasMOD™ Cartridge - C	Non-Dispersive Infrared Spectrometer (NDIR)			Electro-galvanic
	Carbon Dioxide (CO <sub>2</sub> )	Carbon Monoxide (CO)	Hydrocarbon (HC)	Oxygen (O <sub>2</sub> )
Measurement Range	0-20%	0-15%	0-4000ppm (extended range up to 30,000ppm)	0-100%
T <sub>90</sub> Response Time	<3.5 seconds	<3.5 seconds	<3.5 seconds	<6 seconds
Accuracy	±0.3% absolute or ±3% relative	±0.02% absolute or ±3% relative	±8ppm absolute or 3% relative	±0.1% absolute or 2% relative
Repeatability	±0.1% absolute or ±2% relative	±0.02% absolute or ±2% relative	±6ppm absolute or ±2% relative	±0.01% absolute or ±2% relative

Figura 4.16. Características de los sensores de CO<sub>2</sub>, CO, THC, O<sub>2</sub> y del analizador de gases



Figura 4.17. Conexión del analizador de gases al vehículo experimental.

#### 4.4. Desarrollo metodológico del objetivo cuarto específico

##### 4.4.1. Cálculo de los factores de emisión del CTC Bogotá

Posteriormente se realiza el respectivo análisis de las mediciones realizadas para el CTC Bogotá, para la determinación de los factores de emisión del vehículo. Esto se realiza teniendo como la unidad de actividad la distancia recorrida durante el ciclo en km, determinando así los factores de emisión como *unidad de contaminante/km*. De esta forma se desarrolla la

comparación de cada uno de los gases medidos para el CTC Bogotá, así como el análisis de los valores de eficiencia energética para cada uno de los CTCs seleccionados en gal/km.

#### **4.4.2. Comparación de los factores de emisión del CTC Bogotá respecto a los CTCs estándar.**

Es necesario realizar la comparación de los factores de emisión calculados para el CTC Bogotá respecto a los factores de emisión presentados por la Secretaría Distrital de Ambiente (Secretaría de Ambiente de Bogotá, 2022) y los presentados en la resolución 762 del Ministerio de Ambiente (Ministerio Ambiente Y Desarrollo Sostenible, 2022). Esto permite realizar una evaluación de la relación con la realidad que representan los factores de emisión encontrados a partir del CTC construido.

Finalmente se realiza una comparación respecto a los valores de eficiencia energética entre los resultados obtenidos por el CTC Bogotá, respecto a los CTCs estándar.

## **5. Resultados**

### **5.1. Desarrollo del dispositivo IVDR**

#### **5.1.1. Ensamblaje mecánico del dispositivo IVDR**

El resultado final del dispositivo IVDR incluye varios LEDs indicadores con el objetivo de validar el correcto funcionamiento de la recepción y emisión de datos tanto de GPS como de OBD2. El IVDR presenta los siguientes componentes.

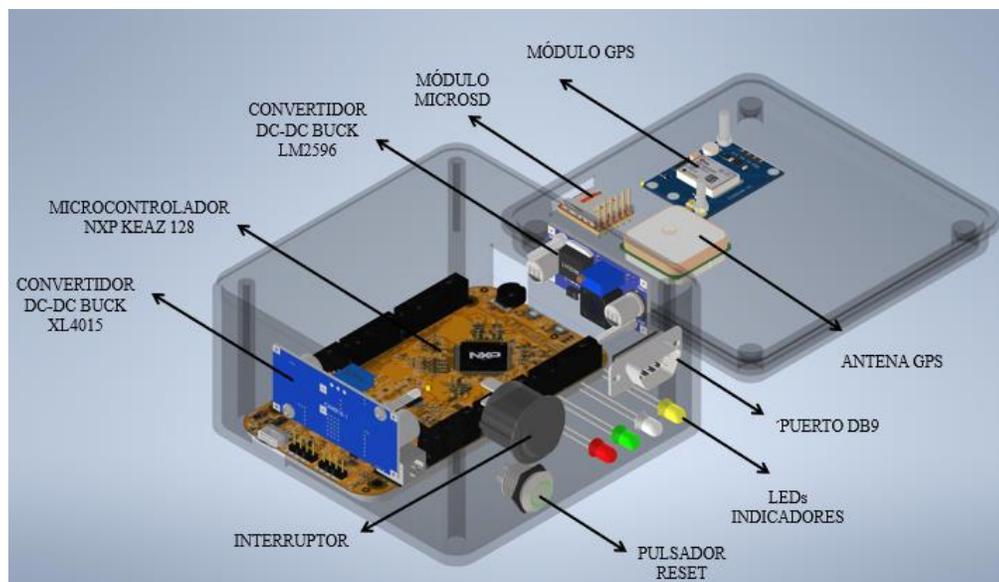
- Cuatro LEDs indicadores para la validación de la recepción y emisión de los datos de OBD2 y los datos de GPS
- Un interruptor de encendido/apagado del dispositivo IVDR
- Un pulsador de reset
- Una ranura para el ingreso/retiro de la memoria MicroSD
- Una ranura de acceso al puerto micro USB del microcontrolador, para realizar modificaciones al software del dispositivo.
- Una ranura al potenciómetro del microcontrolador en caso de que en futuras actualizaciones se necesite este periférico.

- Un puerto DB9, para la conexión con el puerto SAEJ 1962 del vehículo.

El ensamblaje mecánico del dispositivo IVDR, se desarrolló en una caja para proyectos electrónicos, donde mediante corte laser se realizaron los cortes para la inserción de los periféricos del dispositivo. Posterior a la fijación de cada uno de los elementos necesarios del dispositivo, se procede a realizar las conexiones eléctricas mediante cable calibre 22, soldado con estaño. Los cables de transmisión de datos fueron conectados mediante jumpers de Arduino para facilidad de conexión.

Un primer desarrollo del dispositivo se realizó con base en una protoboard y un dispositivo Arduino el cual simulaba el vehículo. Este prototipado inicial, permitió conocer el funcionamiento de cada uno de los periféricos del IVDR y su respectiva validación para posteriormente realizar ensamblaje final.

En la Figura 5.1 se puede visualizar el ensamblaje mecánico del dispositivo IVDR, donde se especifican cada uno de sus componentes. El ensamblaje desarrollado fue realizado en el software CAD Autodesk Inventor 2024.



*Figura 5.1.* Ensamblaje mecánico del dispositivo IVDR y sus componentes.

De igual manera, en la Figura 5.2 se puede visualizar el ensamblaje completo del dispositivo IVDR desarrollado en el mismo software CAD. Aquí se observa que incluyeron indicativos de cada uno de los puertos, dispositivos e interruptores los cuales contiene el dispositivo IVDR.



Figura 5.2. Ensamblaje mecánico del dispositivo IVDR.

Finalmente, en la Figura 5.3 se presenta el dispositivo IVDR construido, ensamblado y totalmente funcional. Para este proyecto se construyeron 9 dispositivos IVDR, de los cuales solamente se utilizaron 2 debido a la disponibilidad de vehículos.



Figura 5.3. Dispositivo IVDR final

### 5.1.2. Esquemático de conexiones del dispositivo IVDR

Para el correcto funcionamiento del dispositivo IVDR, se desarrollaron las respectivas conexiones entre cada uno de los módulos para transferencia de datos, y así mismo para las conexiones eléctricas de cada uno de los módulos. En la Figura 5.4 se presenta el esquemático de conexiones del dispositivo IVDR.

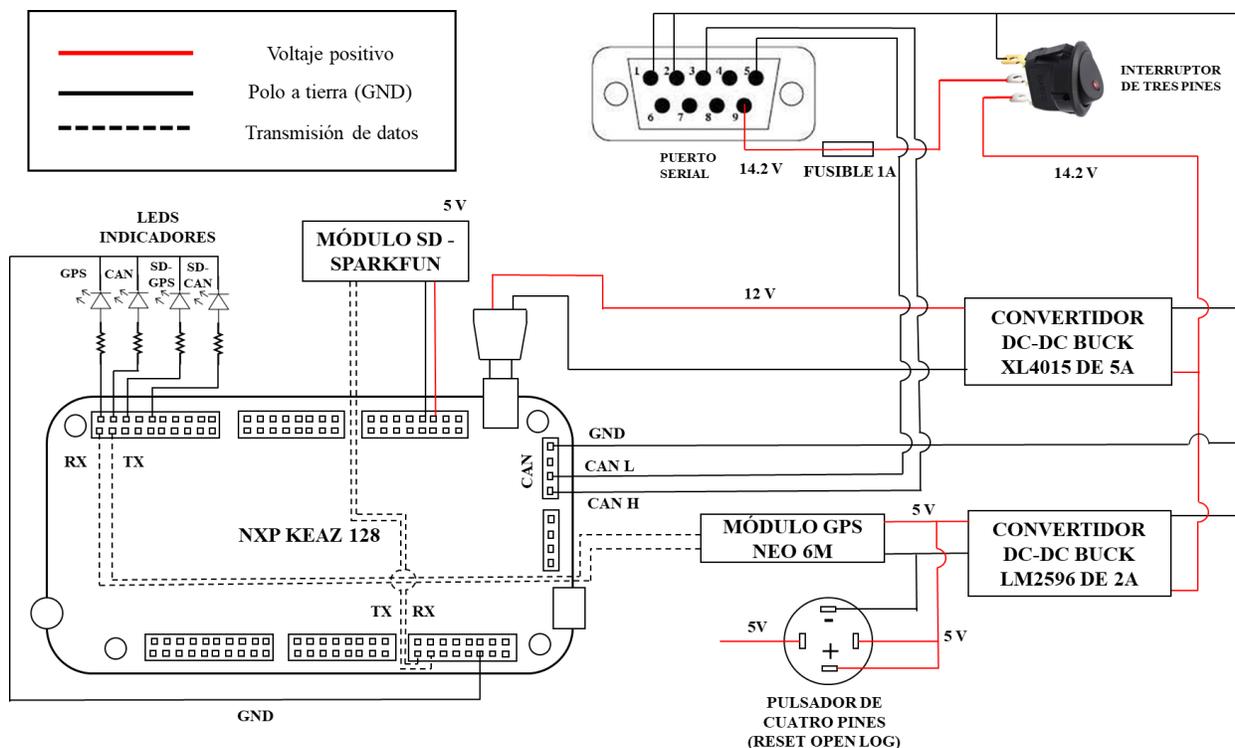


Figura 5.4. Esquemático de conexiones del dispositivo IVDR

A partir del puerto DB9 del vehículo se obtiene la energía eléctrica para alimentar el dispositivo IVDR y cada uno de sus módulos, así como los datos de OBDII. El vehículo proporciona un voltaje de corriente directa (*Direct Current*, por sus siglas en inglés) de ~14.2 V, sin embargo, los módulos requieren 12 V y 5 V para su funcionamiento, específicamente para el microcontrolador NXP KEAZ128 y el módulo GPS, respectivamente. Por este motivo, el convertidor DC-DC XL4015 convierte el voltaje de ~14.2 V a 12 V para que sea conectado al microcontrolador, y así mismo el convertidor DC-DC LM2596 convierte el voltaje de ~14.2 V a 5V para el GPS. Todos los polos a tierra se encuentran unidos al polo a tierra general provisto por el puerto DB9.

Es importante mencionar que, antes de que el voltaje sea convertido, se pasa por un fusible de 1 A y al interruptor con el fin encender y apagar el dispositivo. Si la corriente entregada al dispositivo sobre pasa los 1 A, el fusible se quema con el fin de guardar la integridad del dispositivo.

El módulo MicroSD es alimentado directamente desde el microcontrolador a 5 V, y cada uno de los LEDs indicativos es conectado a una resistencia de 1 k $\Omega$ . Entonces, los LEDs son conectados a cada uno de los puertos del microcontrolador, donde se obtiene la señal digital programada para el indicativo de cada LED. Finalmente, cada módulo es conectado a los puertos de lectura y escritura (RX, TX) del microcontrolador asignado.

### **5.1.3. Validación del dispositivo IVDR**

Como en la metodología, específicamente en la sección 4.2.1, es necesario realizar una validación del dispositivo IVDR. Esta validación se hizo específicamente en el vehículo experimental Nissan Versa – 2013, donde se analizó el correcto funcionamiento de cada uno de los módulos y de la recepción de los datos de OBDII. La trama de datos de CAN contiene, específicamente mediante el protocolo de diagnóstico OBDII, los siguientes datos ID: trama de datos (8 bytes) y timestamp. La validación del timestamp es de suma importancia, ya que permite ubicar los mensajes recibidos de OBDII en el tiempo. Esta validación fue desarrollada por los autores de los trabajos previos (Echeverry Mejía, 2018).

El timestamp no corresponde a un dato entregado por el bus CAN, sino que este dato es asignado directamente por el dispositivo que lee los datos del bus CAN, en este caso el microcontrolador NXP KEAZ128. En el caso del microcontrolador utilizado aquí, este guarda el timestamp en un contador de 16 bits, el cual aumenta en una unidad cada 2 $\mu$ s de acuerdo con el time step de transferencia de datos del protocolo CAN (0.5 Mbps). Cada vez que se lee un dato de OBDII, en microcontrolador guarda dentro del mensaje el timestamp correspondiente.

Tabla 5.1. Validación de tiempo del dispositivo IVDR

ID	DLC	MODO PID	PID	DATA				TIME- STAMP	TIME - STAMP CORREGIDO	TIEMPO (s)	TIME STEP (ms)	
2024	3	65	13	0	0	0	0	0	<b>20086</b>	0	0	
2024	4	65	12	10	240	0	0	0	<b>25205</b>	<b>5119</b>	0.01024	10.238
2024	3	65	13	0	0	0	0	0	30328	10242	0.02048	10.246
2024	4	65	12	10	240	0	0	0	35447	15361	0.03072	10.238
2024	3	65	13	0	0	0	0	0	40061	19975	0.03995	9.228
2024	4	65	12	10	240	0	0	0	45182	25096	0.05019	10.242
2024	3	65	13	0	0	0	0	0	50299	30213	0.06043	10.234
2024	4	65	12	10	240	0	0	0	55418	35332	0.07066	10.238
2024	3	65	13	0	0	0	0	0	60541	40455	0.08091	10.246
2024	4	65	12	10	240	0	0	0	<b>65190</b>	45104	0.09021	9.298
2024	3	65	13	0	0	0	0	0	<b>4735</b>	<b>50185</b>	0.10037	10.162
2024	4	65	12	10	240	0	0	0	9854	55304	0.11061	10.238
2024	3	65	13	0	0	0	0	0	14978	60428	0.12086	10.248

Como se mencionó anteriormente, el dato de timestamp es almacenado en un contador de 16 bits. Por este motivo, el número máximo que puede almacenar este contador es  $2^{16}$  (65536), por lo que al llega a este valor el controlador se reinicia, lo que es importante tener en consideración para el cálculo del timestamp. En la Tabla 5.1 se evidencia a grandes rasgos el proceso de análisis de timestamp. En la columna con título TIME-STAMP, se evidencia el timestamp almacenado, mientras que en la columna inmediatamente siguiente se calcula la diferencia acumulada. Esto tiene en consideración que si el valor anterior al analizado es menor que el actualmente analizado, se resta el número anterior a 65536 y se suma el valor de análisis actual, más el aculado, como se observa en los dos primeros mensajes la diferencia es de 5119. Sin embargo, entre los mensajes 10 y 11, el contador es reiniciado aplicando el proceso anteriormente mencionado.

A los datos de timestamp procesados para su análisis respecto al tamaño del número de bits se multiplican por  $2\mu\text{s}$  para así obtener el valor en el tiempo del mensaje respectivo como se muestra en la columna TIEMPO, donde se puede evidenciar el timestamp de cada dato en segundos. En la última columna se valida el timestep de los datos, siendo este de en promedio 10ms, que velocidad de transmisión de datos seleccionada para la lectura y almacenamiento de datos del dispositivo OBD2. Las pruebas fueron desarrolladas en ambos vehículos de prueba en ruta, sin embargo, los datos de la Tabla 5.1 corresponden a las pruebas realizadas en el vehículo Nissan Versa 2013. El vehículo experimental Ford Fiesta 2017, solo permite un timestep mínimo en OBDII de 30ms.

#### 5.1.4. Resultados de los datos obtenidos por el módulo GPS

A partir de los datos obtenidos mediante el módulo GPS y almacenados en por el microcontrolador en el módulo MicroSD, se procede a realizar el procesamiento de datos para obtener la visualización de los recorridos realizados mediante los datos de latitud y longitud. En la Figura 5.5 se presentan los recorridos realizados durante el proceso de recolección de datos.

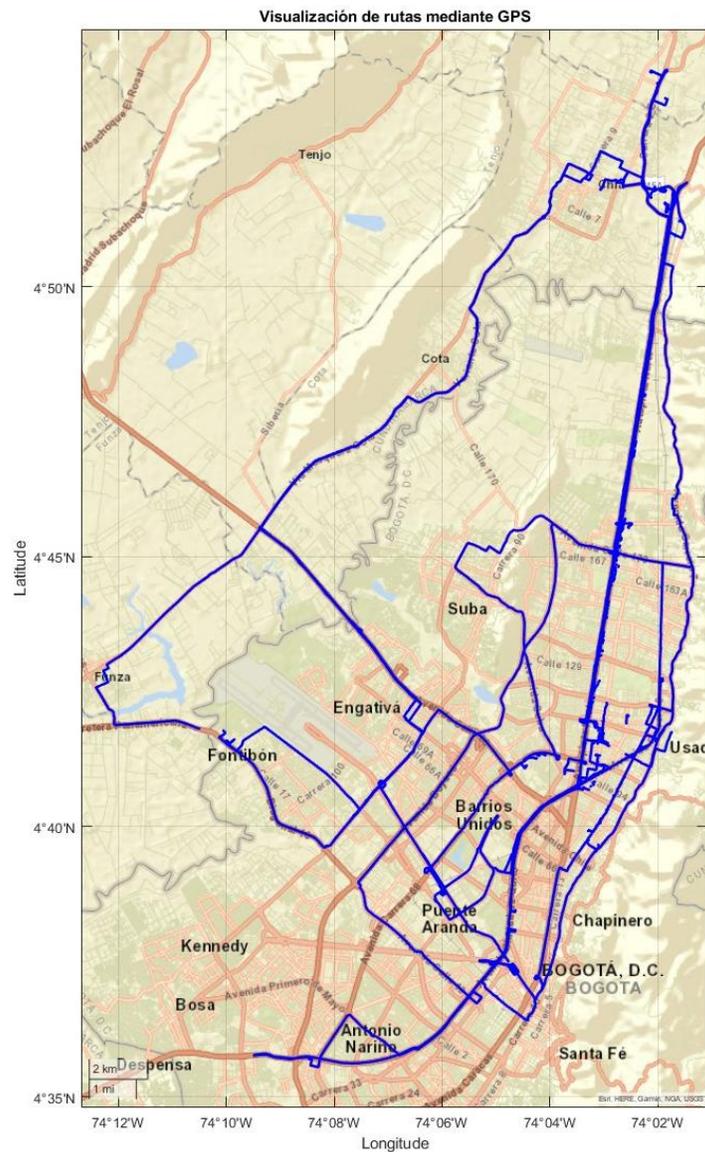


Figura 5.5. Recorridos realizados durante el proceso de recolección de datos

Como fue mencionado en la sección 4.1.1, los recorridos se realizaron en las principales rutas de la ciudad de Bogotá y sus alrededores. En la figura anterior, es posible visualizar todos los recorridos realizados con el fin de garantizar la mayor representatividad posible en la toma de datos de ruta en la zona Bogotá – región.

## **5.2. Desarrollo de algoritmos para la construcción de CTCs (Evaluación comparativa de los métodos de aprendizaje automático)**

### **5.2.1. Análisis general de las metaheurísticas utilizadas**

El algoritmo para la construcción de CTCs, fue desarrollado acorde a como se mencionó en la sección 4.2.1, basado en el algoritmo k-means y tres diferentes variaciones utilizando metaheurísticas, en esta sección se presenta la validación de este algoritmo, disponiendo de tres bases de datos de mediciones de ciclos de conducción, para las ciudades de El Paso - Texas, San Antonio – Texas, y Houston/Galveston – Texas, proporcionadas por el NREL (*National Renewable Energy Laboratory*, por sus siglas en ingles). Estas bases de datos fueron preprocesadas, donde se realizó la respectiva limpieza de datos erróneos, adición de tiempos de ralentí (debido a la naturaleza de la bases de datos, donde para disminuir el tamaño de los archivos, los periodos de ralentí fueron eliminados), si un periodo de ralentí tiene una duración menor a 300 segundos, el periodo es restaurado, si este periodo tiene una duración mayor, no se añade al ciclo de conducción. Posterior a esto, todas las bases de datos fueron segmentadas en microciclos los resultados se presentan en la Tabla 5.2.

Tabla 5.2. *Descripción de las base de datos*

<b>Base de datos</b>	<b>Recorridos</b>	<b>Microciclos</b>
Houston/Galveston	745	15113
El Paso	1382	19821
San Antonio	624	11202

Cada una de las bases de datos, una vez segmentada en microciclos, fueron calculados los parámetros mencionados en la sección metodológica, sin embargo, debido a que estas bases de datos cuentan con información limitada, de los 36 parámetros analizados, solamente 27 pudieron ser calculados, los cuales se presentan en la Tabla 5.3.

Tabla 5.3. *Parámetros calculados para las bases de datos seleccionadas.*

Parámetro	Unidad	Parámetro	Unidad	Parámetro	Unidad
Velocidad máxima	km/h	Proporción de desaceleración	%	Número de paradas	-
Velocidad promedio	km/h	Proporción de crucero	%	Fuerzas aerodinámicas	m <sup>3</sup> /s <sup>3</sup>
Velocidad media en movimiento	km/h	Proporción de ralentí	%	Duración media de un ciclo de conducción	s
Aceleración máxima	m/s <sup>2</sup>	Rango de velocidad (0-20 km/h)	%	Número promedio de cambios de aceleración y desaceleración	-
Desaceleración máxima	m/s <sup>2</sup>	Rango de velocidad (20-40 km/h)	%	Aceleración mínima	m/s <sup>2</sup>
Aceleración promedio	m/s <sup>2</sup>	Rango de velocidad (40-60 km/h)	%	Desaceleración mínima	m/s <sup>2</sup>
Desaceleración promedio	m/s <sup>2</sup>	Rango de velocidad (60-80 km/h)	%	Desviación estándar de la velocidad	km/h
Tiempo en movimiento	s	Rango de velocidad (>80 km/h)	%	Desviación estándar de la desaceleración	m/s <sup>2</sup>
Proporción de aceleración	%	Tiempo de ralentí	s	Desviación estándar de la aceleración	m/s <sup>2</sup>

La matriz con cada uno de los parámetros calculados, para cada microciclo es la entrada al algoritmo PCA, el cual redujo la dimensionalidad del análisis de 27 (teniendo una dimensión por cada parámetro), a 13, teniendo 98% de la variación de los datos. Así mismo, la matriz de componentes principales corresponde a la entrada para la clusterización de los microciclos, del algoritmo k-means. Para el desarrollo del algoritmo k-means respecto a cada una de las metaheurísticas aplicadas, juntos con sus métodos de inicialización se utiliza un tamaño de clusters definido como tres acorde a un análisis realizado mediante el método del codo, donde en esta cantidad de clusters se evidencia el punto de inflexión respecto a la función de costo para el algoritmo k-means.

En la Tabla 5.4 se presentan los resultados generales obtenidos por cada una de las metaheurísticas utilizadas, en combinación con el algoritmo k-means, para cada una de las bases de datos.

Tabla 5.4. *Resultados generales de las metaheurísticas utilizadas.*

Metaheurística	EL PASO		SAN ANTONIO		HOUSTON/GALVESTON	
	Valor de la función de costo	Iteraciones	Valor de la función de costo	Iteraciones	Valor de la función de costo	Iteraciones
GRASP - RANDOM	3.48606	148	3.18262	130	5.28726	217
GRASP - FORGY	3.48526	127	3.18262	101	5.28701	124
ILS - RANDOM	3.48642	145	3.18256	101	5.99471	107
ILS - FORGY	3.48498	172	3.18262	104	5.28579	107
GENETIC K-MEANS	3.48495	13107	3.18262	13805	5.28578	15389

En general es posible evidenciar que para cada una de las bases de datos las metaheurísticas convergen a valores similares de la función de costo. Los algoritmos que evidencian menores valores de la función de costo, lo cual es bueno ya que el algoritmo k-means es un algoritmo que busca minimizar el valor de la función de costo, son los algoritmos ILS. En la base de datos de El Paso, el algoritmo ILS-FORGY presenta menores valores de la función de costo en comparación con el algoritmo ILS-RANDOM. Resultados similares se evidencian respecto a la base de datos de Houston/Galveston, sin embargo para esta base de datos presenta valores más altos de la función de costo respecto a las demás metaheurísticas, en comparación con la base de datos de San Antonio, donde se evidenció que los valores de la función de costo obtenidos por el algoritmo ILS-RANDOM son menores, respecto a los valores encontrados por los valores de otras metaheurísticas.

En términos de iteraciones, al igual que la función de costo las metaheurísticas ILS, presentan menor número de iteraciones, para la base de datos de San Antonio y Houston/Galveston, sin embargo para la base de datos de El Paso, un mayor número de iteraciones fueron necesarias.

El algoritmo genético k-means de ahora en adelante mencionado como GKA, si bien presenta menores valores de función de costo para las bases de datos de El Paso y Houston/Galveston, la cantidad de iteraciones necesarias para llegar a esta solución es entre ~76 y ~143 veces mayor en comparación con otras metaheurísticas.

Los algoritmos ILS, presentan menores valores de función de costo respecto a las demás metaheurísticas e iteraciones debido a la construcción de este algoritmo, permitiendo tener un proceso de diversificación en el espacio de análisis mayor en comparación las otras metaheurísticas, esto debido al proceso de perturbación de la solución encontrada. Esto permite al algoritmo tener diferentes posibilidades de inicialización, con base en resultados previos para así llegar a soluciones óptimas locales en un menor tiempo, en comparación con el algoritmo GRASP, el cual solamente corre el algoritmo ciertas veces hasta encontrar una mejor solución. El algoritmo GKA, presenta buenos resultados debido al proceso de diversificación incluido en la operación de crossover y en el proceso de mutación, sin embargo la cantidad de iteraciones requeridas son demasiadas en comparación con otras metaheurísticas.

La distribución de microciclos en cada uno de los clusters se presenta en la Tabla 5.5 , donde se evidencia que el porcentaje de microciclos asignado a cada cluster es similar para todas las bases

de datos. Es importante aclarar que cada cluster fue organizado teniendo en consideración la velocidad media de cada clúster, donde el cluster 1 corresponde a la mínima velocidad, el cluster 2 a una velocidad media y el cluster 3 a la máxima velocidad media, esto para cada base de datos.

La metaheurística ILS-RANOM presenta resultados diferentes para la base de datos de Houston/Galveston en comparación las otras metaheurísticas analizadas. Como se mencionó anteriormente este algoritmo presenta resultados mostrando menor desempeño del algoritmo para esta base de datos.

Tabla 5.5. *Distribución de microciclos en cada uno de los clusters para cada base de datos*

Metaheurística	EL PASO			SAN ANTONIO			HOUSTON/GALVESTON		
	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
GRASP - RANDOM	83.08	11.95	4.97	89.47	10.28	0.25	82.83	12.53	4.64
GRASP - FORGY	82.11	12.81	5.08	89.47	10.28	0.25	81.68	13.52	4.80
ILS - RANDOM	83.15	11.89	4.96	89.47	10.28	0.25	91.81	8.10	0.09
ILS - FORGY	82.39	12.55	5.07	89.47	10.28	0.25	82.25	13.03	4.72
GENETIC K-MEANS	82.53	12.42	5.05	89.47	10.28	0.25	82.25	13.03	4.72

Como se mencionó anteriormente los clusters fueron organizados de menor a mayor velocidad media, en la Tabla 5.6 es posible evidenciar la velocidad media de cada uno de los clusters, acorde a los microciclos clusterizados y para cada base de datos. Donde el primer cluster contiene los microciclos de menor velocidad media, el segundo cluster contiene los microciclos con una velocidad media intermedia, y el tercer cluster contiene los microciclos velocidad media alta. De igual manera el algoritmo ILS-RANDOM, presenta diferentes resultados para la base de datos Houston/Galveston.

Tabla 5.6. *Velocidad promedio en cada clúster para cada base de datos.*

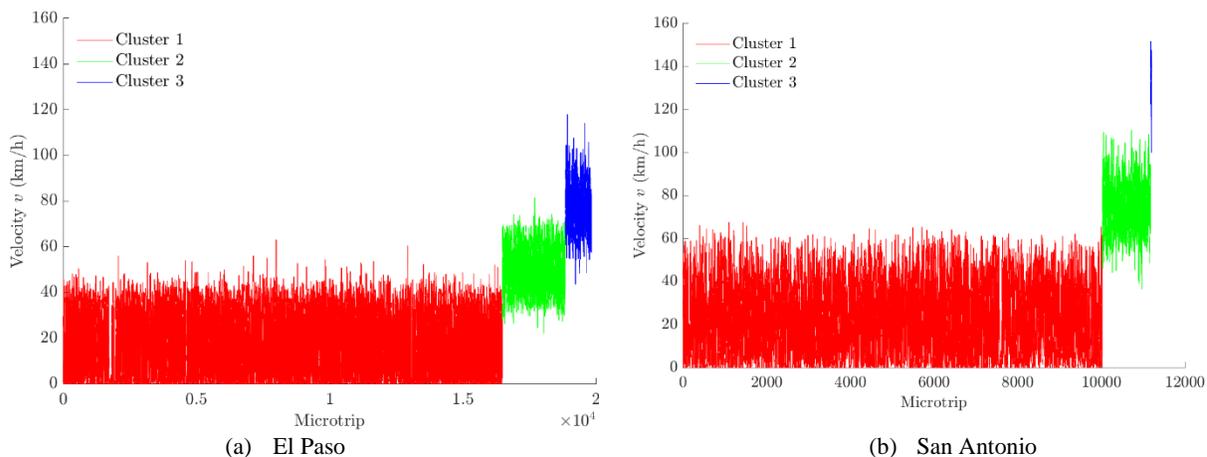
Metaheurística	EL PASO			SAN ANTONIO			HOUSTON		
	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
	(km/h)	(km/h)	(km/h)	(km/h)	(km/h)	(km/h)	(km/h)	(km/h)	(km/h)
GRASP - RANDOM	16.91	48.77	77.00	23.63	74.10	133.93	20.50	53.70	84.21
GRASP - FORGY	16.65	47.90	76.76	23.63	74.10	133.93	20.18	52.63	83.75
ILS - RANDOM	16.93	48.83	77.02	23.63	74.10	133.93	23.33	75.26	115.71
ILS - FORGY	16.72	48.13	76.79	23.63	74.10	133.93	20.35	53.12	83.99
GENETIC K-MEANS	16.76	48.24	76.83	23.63	74.10	133.93	20.35	53.12	83.99

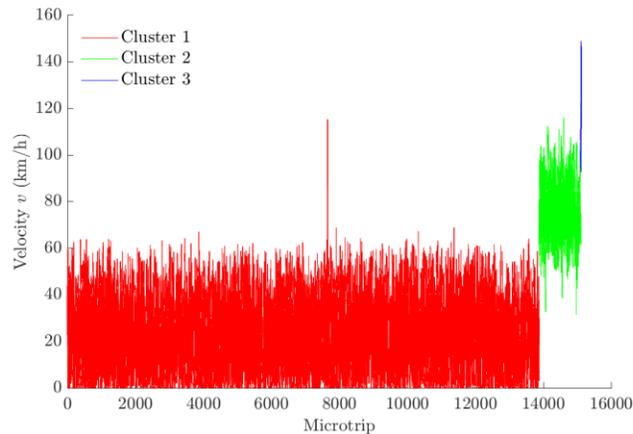
En general el cluster con mayor cantidad de microciclos asignados es el cluster 1, el cual contiene una velocidad media baja, indicando condiciones de tráfico altas, seguido de este cluster

se encuentra el cluster 2, el cual cuenta con velocidades medias intermedias, lo cual indica condiciones de tráfico medio, y por último se encuentra el cluster 3, con velocidades medias altas, indicando condiciones de tráfico bajas.

Para evaluar las metaheurísticas de manera más intuitiva, una representación gráfica de la distribución de los microciclos similar a la presentada por Yuan et al. (Yuan et al., 2021), es presentada en la Figura 5.6. Tomando la clusterización realizada por el algoritmo ILS-RANDOM, y como variable de visualización la velocidad media de cada microciclo.

En esta figura se evidencia que el cluster 1, contiene la mayor cantidad de microciclos, para todas las bases de datos, seguido de cluster 2 y el cluster 3, respectivamente. La superposición de microciclos se evidencia en espacio de transición entre un cluster y otro. En la Figura 5.6(c) se evidencia un microciclo atípico asignado al cluster 1, el cual acorde a su velocidad media debería estar asignado al cluster 3, sin embargo es importante recalcar que la velocidad media es solamente uno de los parámetros de análisis utilizados para el proceso de clusterización. En apéndice **Distribución de velocidades para todas las metaheurísticas sin refinamiento de parámetros.**, se presentan las distribuciones para los diferentes algoritmos y cada base de datos.





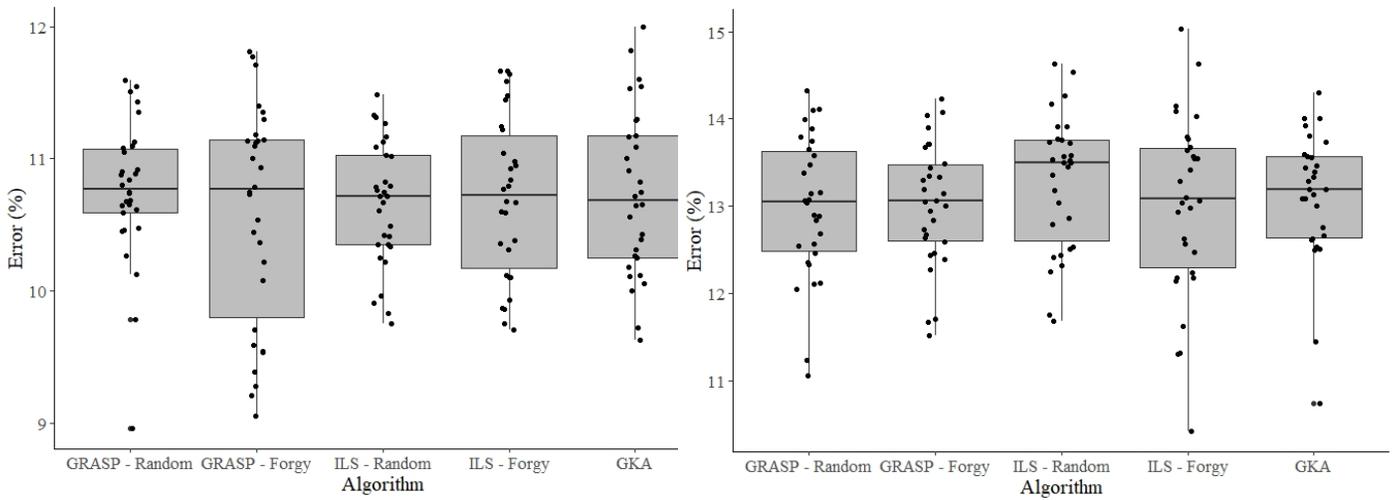
(c) San Antonio

*Figura 5.6.* Distribución de los microciclos acorde a la velocidad media para la metaheurística ILS-RANDOM.

### 5.2.2. Análisis estadístico de las metaheurísticas utilizadas

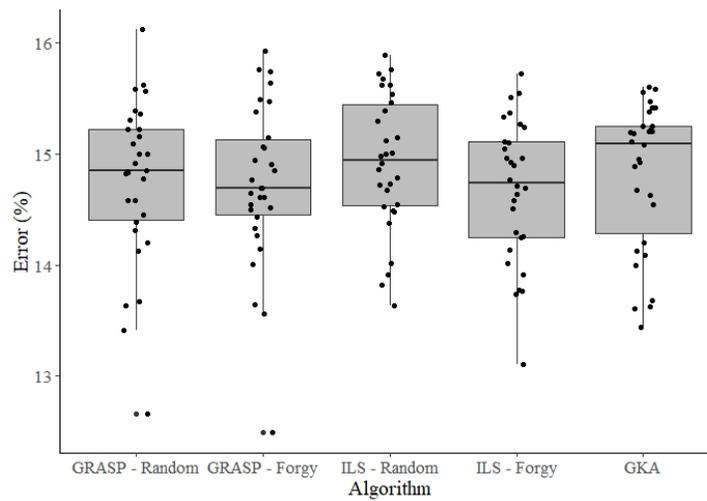
Como se mencionó en la sección 4.2.1, las metaheurísticas utilizadas tienen componentes aleatorios, lo cual hace que la solución encontrada en una corrida del algoritmo no sea la misma en una diferente, esto respecto a las iteraciones utilizadas, llegando en sí a soluciones muy parecidas entre sí para los valores de función de costo, o error medio. Es por esto por lo que se realiza un análisis estadístico para evidenciar el efecto de las metaheurísticas utilizadas en el error medio de los CTCs y el tiempo de ejecución para la construcción del CTC. Esto es realizado para cada base de datos, y así realizar las respectivas comparaciones.

En primer lugar, se realiza un análisis gráfico previo, a proceder con las pruebas estadísticas robustas. En la Figura 5.7 y Figura 5.8 se encuentran las gráficas box-plot, para el análisis del promedio del error medio, y el tiempo de ejecución.



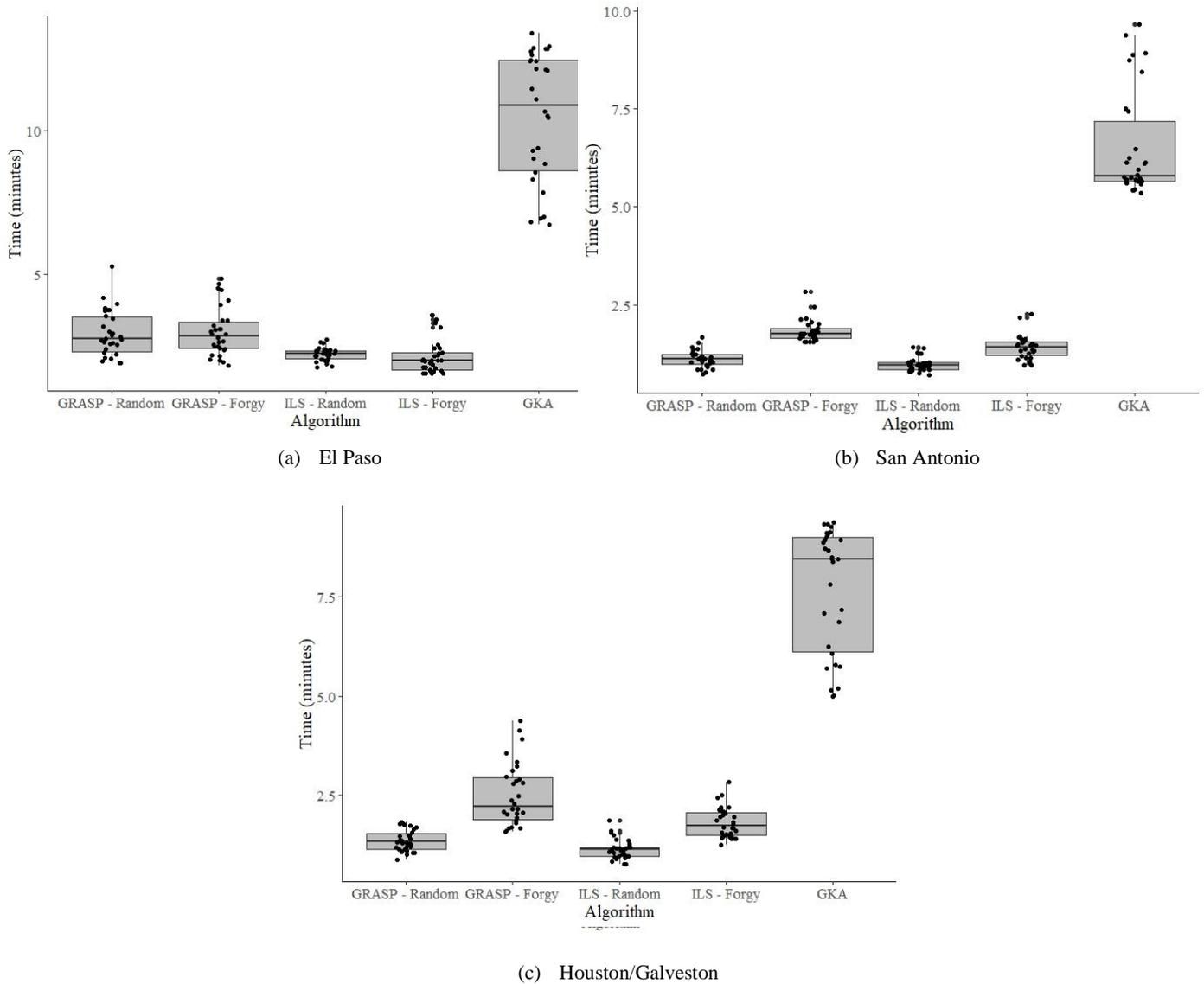
(a) El Paso

(b) San Antonio



(c) Houston/Galveston

Figura 5.7. Box-plot del error medio para todas las metaheurísticas en todas las bases de datos



*Figura 5.8.* Box-plot del tiempo de ejecución para todas las metaheurísticas en todas las bases de datos

El análisis gráfico presentado anteriormente, presenta que, para el promedio del error medio de los ciclos típicos de conducción construidos, no hay un efecto significativo de la metaheurística utilizada demostrando así la robustez de la metodología utilizada. Sin embargo, para el tiempo de ejecución es importante notar que, si hay un efecto significativo, respecto al algoritmo GKA, el cual presenta un mayor tiempo de procesamiento, para llegar a errores similares. De manera general, el algoritmo ILS-RANDOM aparenta tener un mejor desempeño respecto al tiempo de ejecución, teniendo una desviación estándar baja, y tiempos de ejecución bajos.

Con el objetivo de evidenciar los efectos de cada metaheurística, se plantean un análisis estadístico mediante pruebas paramétricas y no paramétricas. En primero lugar se realiza el análisis de la validación de supuestos (normalidad, independencia, y homocedasticidad), con el objetivo de realizar una prueba ANOVA, en dado caso que los supuestos no se cumplieren se realizan pruebas no paramétricas, en este caso una prueba de Kruskal-Wallis. En la Tabla 5.7 y la Tabla 5.8 se presentan los resultados de la validación de supuestos, para cada base de datos.

Tabla 5.7. *Validación de supuestos para el promedio del error medio para cada base de datos*

<b>Validación de supuestos para el error medio</b>			
	<b>EL PASO</b>	<b>SAN ANTONIO</b>	<b>HOUSTON</b>
<b>Normalidad</b>	<b>p-valor</b>	<b>p-valor</b>	<b>p-valor</b>
Shapiro - Wilk	0.13	0.02	$1.04 \times 10^{-3}$
Lilliefors (Kolmogorov - Smirnov)	0.13	0.13	0.11
D'Agostino	0.05	$5.83 \times 10^{-3}$	$8.38 \times 10^{-4}$
Anderson - Darling	0.22	0.02	$3.82 \times 10^{-3}$
<b>Homocedasticidad</b>	<b>p-valor</b>	<b>p-valor</b>	<b>p-valor</b>
Fligner - Killen	$8.31 \times 10^{-3}$	0.31	0.99
Levene	$9.16 \times 10^{-3}$	0.30	0.99
<b>Independencia</b>	<b>p-valor</b>	<b>p-valor</b>	<b>p-valor</b>
Durbin-Watson	0.34	0.75	0.98

Tabla 5.8. *Validación de supuestos para el tiempo de ejecución para cada base de datos*

<b>Validación de supuestos para el tiempo de ejecución</b>			
	<b>EL PASO</b>	<b>SAN ANTONIO</b>	<b>HOUSTON</b>
<b>Normalidad</b>	<b>p-valor</b>	<b>p-valor</b>	<b>p-valor</b>
Shapiro - Wilk	$2.20 \times 10^{-16}$	$2.20 \times 10^{-16}$	$5.67 \times 10^{-16}$
Lilliefors (Kolmogorov - Smirnov)	$2.20 \times 10^{-16}$	$2.20 \times 10^{-16}$	$2.20 \times 10^{-16}$
D'Agostino	$2.23 \times 10^{-10}$	$6.85 \times 10^{-11}$	$1.24 \times 10^{-9}$
Anderson - Darling	$2.20 \times 10^{-16}$	$2.20 \times 10^{-16}$	$2.20 \times 10^{-16}$
<b>Homocedasticidad</b>	<b>p-valor</b>	<b>p-valor</b>	<b>p-valor</b>
Fligner - Killen	$8.87 \times 10^{-15}$	$7.47 \times 10^{-5}$	$5.37 \times 10^{-11}$
Levene	$2.20 \times 10^{-16}$	$3.66 \times 10^{-7}$	$1.66 \times 10^{-12}$
<b>Independencia</b>	<b>p-valor</b>	<b>p-valor</b>	<b>p-valor</b>
Durbin-Watson	0.24	0.99	0.16

Al analizar la validación de supuestos ninguna base de datos los cumple en totalidad con el objetivo de realizar pruebas paramétricas, por esta razón un se realiza un análisis no paramétrico basado en la prueba de Kruskal-Wallis; los resultados se encuentran en la Tabla 5.9, donde se evidencia acorde al p-valor que para el error medio no hay una diferencia significativa respecto a los algoritmos utilizados para la construcción del CTC, sin embargo si hay una diferencia significativa respecto al tiempo de ejecución para la construcción del CTC, como se evidenció en el análisis gráfico.

Tabla 5.9. *Resultados de las pruebas no paramétricas para el promedio del error medio y para tiempo de ejecución de los CTCs para cada base de datos*

<b>Kruskal - Wallis</b>	<b>EL PASO</b>	<b>SAN ANTONIO</b>	<b>HOUSTON</b>
p-valor tiempo de ejecución	$2.20 \times 10^{-16}$	$2.20 \times 10^{-16}$	$2.20 \times 10^{-16}$
p-value error medio	0.94	0.65	0.64

Pruebas Post-Hoc fueran desarrolladas, las cuales demuestran que en general el efecto significativo el tiempo de ejecución se presenta para la metaheurística GKA, como se evidenció en el análisis gráfico.

El error medio no se ve afectado por el algoritmo la metaheurística utilizada para la clusterización de los datos, debido a que el CTC no depende directamente de este algoritmo para su construcción, ya que el algoritmo de clusterización solamente cumple con el objetivo de asignar los microciclos acorde a sus características, para posteriormente realizar la TPM, y a partir de esta el CTC. Como se vio en la Tabla 5.4, los algoritmos presentan valores similares de la función de costo presentando así TPMs muy similares, lo cual no afecta la construcción del CTC.

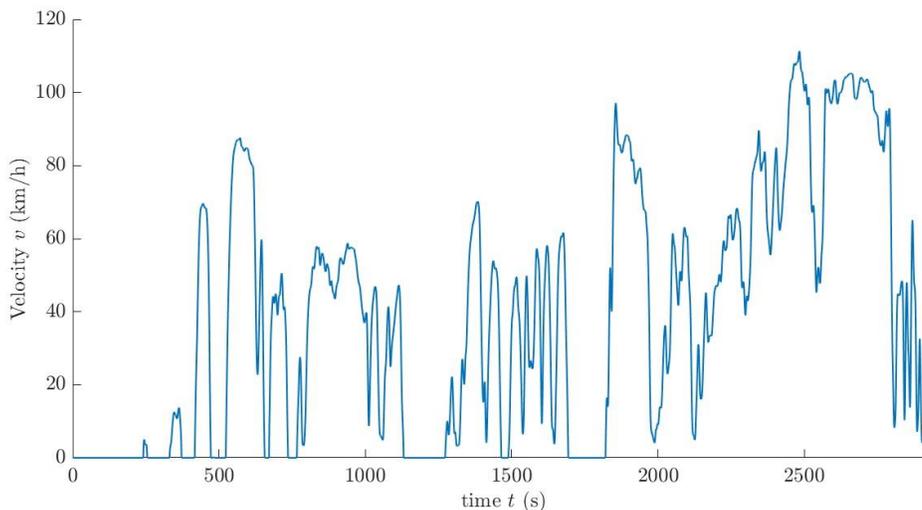
En contraste el tiempo de ejecución para la construcción del CTC, si es afectado por el tipo de metaheurística utilizada, debido al tiempo empleado en la búsqueda de una solución óptima local de la función de costo. Esto es evidenciado específicamente para el algoritmo GKA, el cual debido a la cantidad de iteraciones empleadas presenta un mucho mayor tiempo de ejecución para encontrar la misma solución.

Teniendo en consideración el análisis presentado, y el comportamiento de la metaheurística ILS-RANDOM, respecto al valor medio del tiempo de ejecución y su desviación estándar, esta es seleccionada como la metaheurística a utilizar para la construcción de los respectivos CTCs, sin embargo es importante mencionar que para la base de datos de Bogotá región se debe analizar debido a su inconsistencia en los resultados de clusterización encontrados para la base de datos Houston/Galveston.

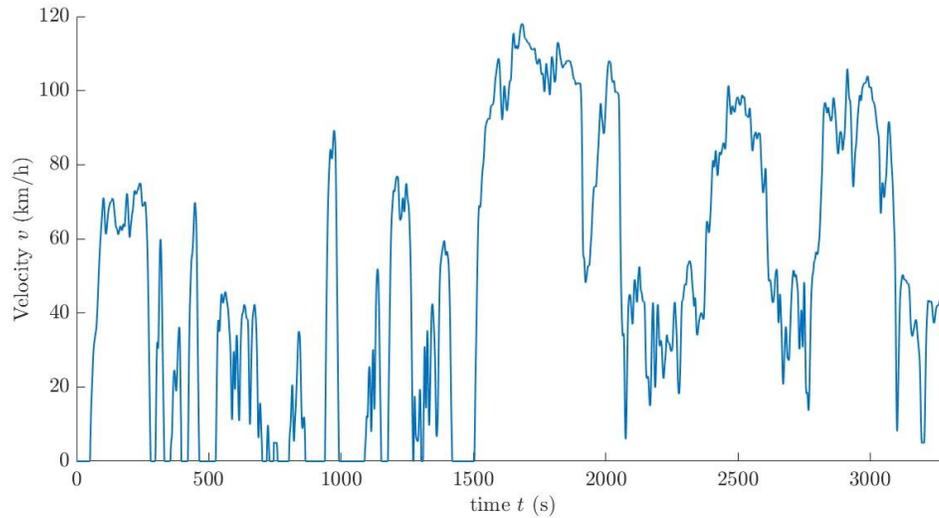
### 5.2.3. Ciclos típicos desarrollados para cada base de datos

Como resultado del análisis estadístico se obtiene la metaheurística con un desempeño destacado, y se utiliza como el algoritmo para la construcción de los CTCs, siguiendo los pasos descritos en la metodología. En la Figura 5.9 se presenta el CTC construido para la base de datos de El Paso, en la

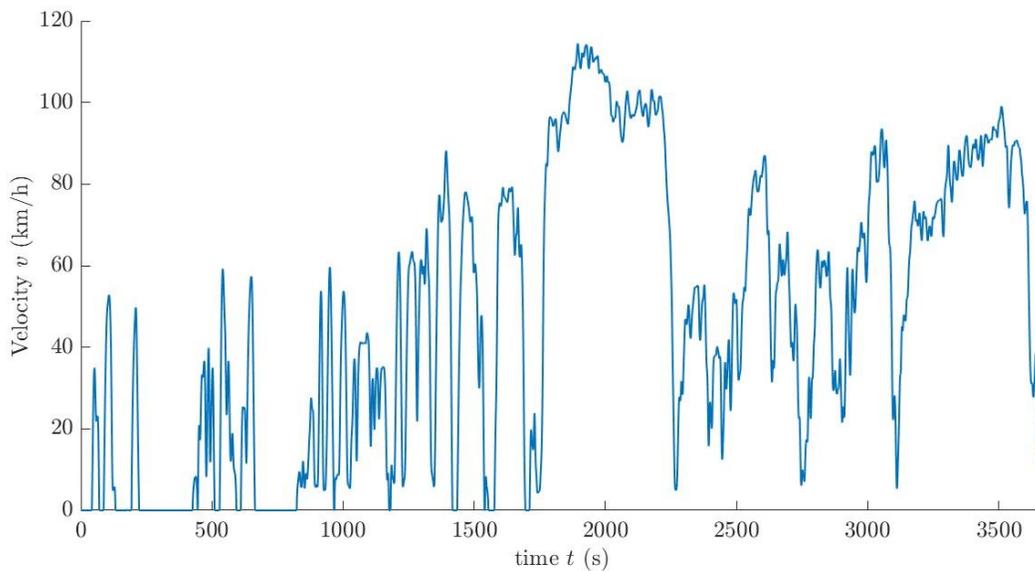
Figura 5.10 se presenta el CTC para la base de datos de San Antonio, y en la Figura 5.11 se presenta el CTC para la base de datos Houston/Galveston. El CTC construido para El Paso, es considerado como un CTC de tráfico alto-medio debido a que los modos de conducción, y rangos de velocidad se encuentran en los rangos para los cuales los microciclos tienen esta característica, así mismo los microciclos utilizados para la construcción de este CTC, corresponden a este tipo de microciclos, teniendo en consideración el uso de microciclos de tráfico alto y medio, este comportamiento también se evidencia para los CTCs de cada base de datos. Sin embargo estos CTCs construidos también consideran en los rangos de alta velocidad condiciones de tráfico medio.



*Figura 5.9.* Ciclo típico de conducción construido para El Paso



*Figura 5.10.* Ciclo típico de conducción para San Antonio



*Figura 5.11.* Ciclo típico de conducción para Houston/Galveston

Cada uno de los ciclos anteriormente presentados, presenta el análisis de encendido del motor en frío, teniendo un periodo de ralentí inicial, así mismo el encendido en caliente es tenido en consideración, debido a los periodos de ralentí intermedios. Estos dos tipos de encendido del motor son de gran importancia ya que estos evidencian características en emisiones importantes a considerar. Los parámetros generales de los CTCs construidos se encuentran en la Tabla 5.10.

Tabla 5.10. Comparación de los parámetros de cada ciclo típico de conducción construido

Comparación de los parámetros de los CTCs				
Parámetro	Unidad	EL PASO	SAN ANTONIO	HOUSTON
Velocidad máxima	km/h	111.32	117.99	114.45
Velocidad promedio	km/h	38.66	49.94	47.20
Velocidad promedio en movimiento	km/h	52.08	60.10	56.46
Aceleración máxima	m/s <sup>2</sup>	6.05	5.38	4.97
Desaceleración mínima	m/s <sup>2</sup>	-0.15	-0.15	-0.15
Aceleración promedio	m/s <sup>2</sup>	1.06	0.98	1.02
Desaceleración promedio	m/s <sup>2</sup>	-1.24	-1.08	-1.12
Tiempo de recorrido	s	2187	2741	3067
Proporción de aceleración	-	32.04%	34.80%	35.38%
Proporción de desaceleración	-	27.39%	31.62%	32.38%
Proporción de cruce	-	12.93%	14.03%	12.48%
Proporción de ralentí	-	25.39%	16.43%	15.94%
Rango de velocidad (0-20 km/h)	-	38.66%	26.46%	29.38%
Rango de velocidad (20-40 km/h)	-	12.42%	16.34%	16.11%
Rango de velocidad (40-60 km/h)	-	22.30%	17.19%	15.29%
Rango de velocidad (60-80 km/h)	-	9.98%	13.16%	15.92%
Rango de velocidad (>80 km/h)	-	16.63%	26.86%	23.30%
Tiempo de ralentí	s	748.00	542	585
Número de paradas	-	10	13	13
Fuerzas aerodinámicas	m <sup>3</sup> /s <sup>3</sup>	4452.81	7069.03	6111.98
Duración del microciclo	s	2946	3299	3669
Número de cambios de aceleración y desaceleración	-	49	37	68
Aceleración mínima	m/s <sup>2</sup>	0.15	0.15	0.15
Desaceleración máxima	m/s <sup>2</sup>	-5.66	-5.07	-5.24
Desviación estándar de la velocidad	km/h	34.04	36.46	35.17
Desviación estándar de la desaceleración	m/s <sup>2</sup>	1.15	1.06	1.03
Desviación estándar de la aceleración	m/s <sup>2</sup>	0.95	0.88	0.90

#### 5.2.4. Comparación de los ciclos construidos con los ciclos internacionales

Los ciclos construidos presentados anteriormente son comparados con los ciclos estándar FTP-75, WLTC con cada una de sus tres clases, y el ciclo (UC)/LA 92, respecto al error medio de cada ciclo, para cada base de datos, teniendo en consideración los 27 parámetros analizados. En la Tabla 5.11 se presenta la comparación del promedio de los errores medios para cada base de datos.

Tabla 5.11. Comparación de los errores medios para cada base de datos.

	EL PASO	SAN ANTONIO	HOUSTON
Ciclo típico de conducción	Error (%)	Error (%)	Error (%)
CTC construido	11.15	12.48	14.38
FTP-75	53.71	57.41	57.97
WLTC Clase 3	39.34	35.84	37.31
WLTC Clase 2	50.04	50.65	51.52
WLTC Clase 1	62.64	65.17	64.87
(UC)/LA 92	37.62	42.40	43.27

Teniendo en consideración los resultados presentados, es evidente que los ciclos estándar presentan errores bastante altos respecto a los ciclos construidos, si bien los errores de los ciclos construidos no son demasiado bajos, presentan un error aceptable, respecto a los ciclos estándar. Demostrando la baja representatividad del comportamiento real de los vehículos en una zona específica.

### 5.2.5. Mejoramiento de los CTCs desarrollados

Con el objetivo de disminuir el error, y así mejorar la representatividad de CTC construido para cada base de datos se procede a calcular el error medio respecto a cada parámetro para cada base de datos y compararlo con el error medio del CTC, con esto se evidenciará cuales parámetros contribuyen más al error medio del ciclo. En la Figura 5.12 se presenta el análisis de la comparación de los errores medios de los parámetros para cada base datos.

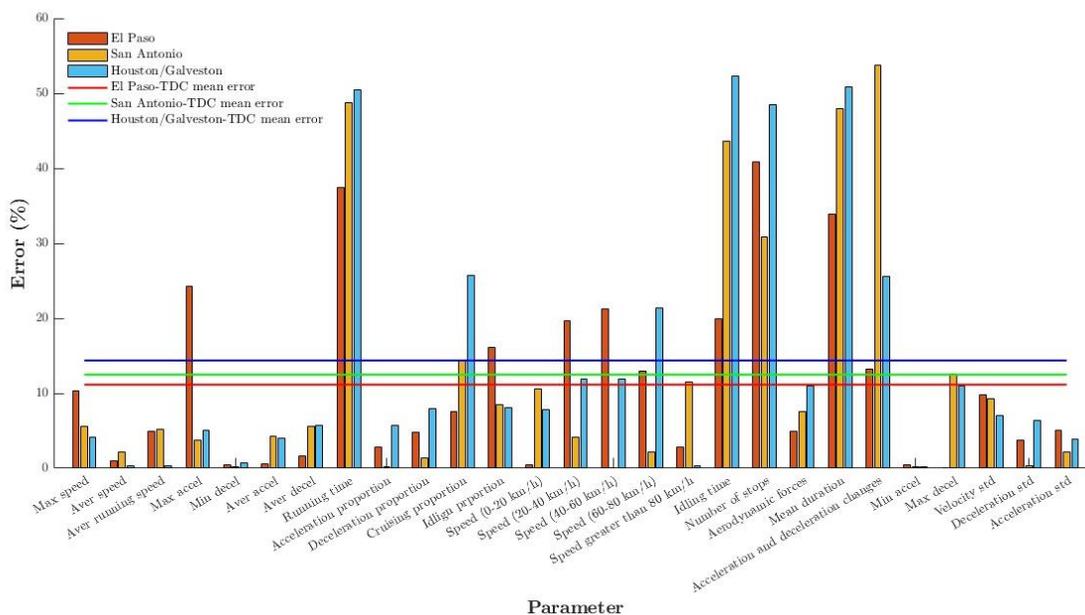


Figura 5.12. Comparación del error medio de los parámetros para cada base de datos.

En la Figura 5.12 es posible evidenciar que el error medio de algunos parámetros excede el error medio del CTC para la base de datos respectiva, en este caso los parámetros que exceden en mayor medida el error medio del CTC construido para cada base de datos son: Tiempo en movimiento, proporción de crucero, tiempo en ralentí, número de paradas, duración media del ciclo, y los cambios en aceleraciones y desaceleraciones. En sí estos parámetros, son aquellos que

dependen de la longitud del CTC, y debido a la diferencia en tamaños de microciclos, estos parámetros aumentan el error medio del CTC respecto a la base de datos. Por esta razón los parámetros que se tendrán en consideración tendrán las siguientes características.

- Deben poder ser calculados con base en los datos disponibles
- Deben ser independientes del tamaño de los microciclos.

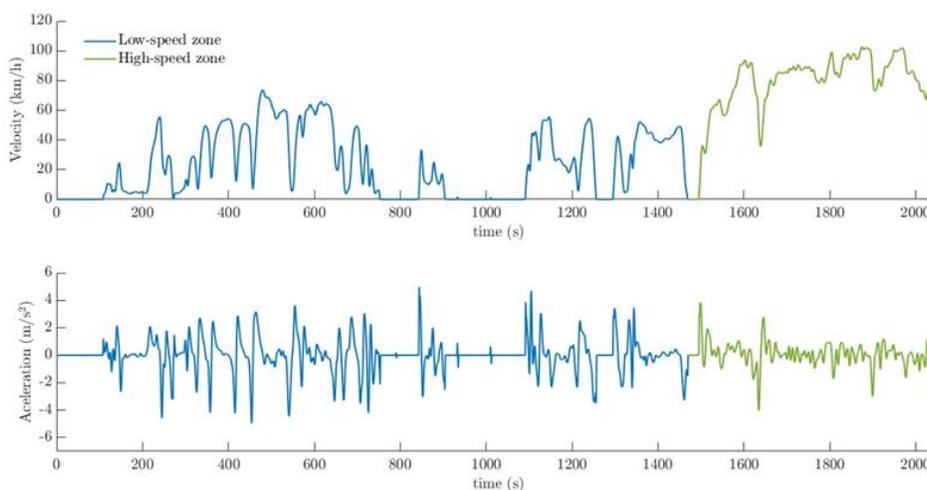
Por este motivo la cantidad de parámetros es reducida de 27 parámetros a 21 parámetros, realizando un refinamiento, no teniendo en consideración los parámetros que excedieron en mayor medida el error medio de cada CTC construido. Los parámetros que se tienen en consideración son los siguientes.

Tabla 5.12. *Parámetros seleccionados para el mejoramiento de los CTCs*

Parámetro	Unidad	Parámetro	Unidad	Parámetro	Unidad
Velocidad máxima	km/h	Proporción de desaceleración	%	Aceleración mínima	m/s <sup>2</sup>
Velocidad promedio	km/h	Proporción de ralentí	%	Desaceleración mínima	m/s <sup>2</sup>
Velocidad media en movimiento	km/h	Rango de velocidad (0-20 km/h)	%	Desviación estándar de la velocidad	km/h
Aceleración máxima	m/s <sup>2</sup>	Rango de velocidad (20-40 km/h)	%	Desviación estándar de la desaceleración	m/s <sup>2</sup>
Desaceleración máxima	m/s <sup>2</sup>	Rango de velocidad (40-60 km/h)	%	Desviación estándar de la aceleración	m/s <sup>2</sup>
Aceleración promedio	m/s <sup>2</sup>	Rango de velocidad (60-80 km/h)	%		
Desaceleración promedio	m/s <sup>2</sup>	Rango de velocidad (>80 km/h)	%		
Proporción de aceleración	%	Fuerzas aerodinámicas	m <sup>3</sup> /s <sup>3</sup>		

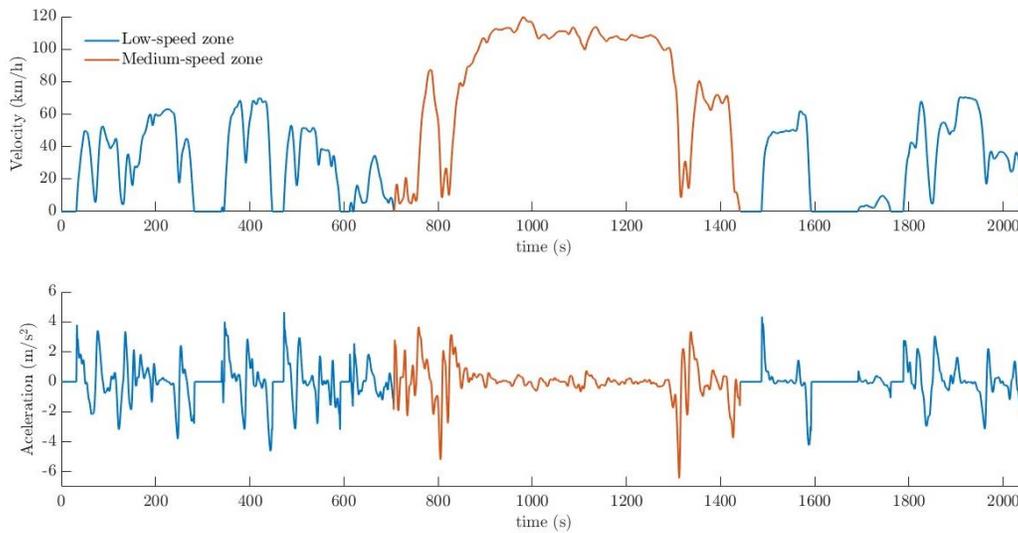
Teniendo en consideración el refinamiento de parámetros, se realizaron los análisis de variabilidad de los algoritmos, desarrollando 30 corridas de cada algoritmo, para así analizar la variabilidad respecto a la función de costo, velocidad promedio en cada cluster, distribución de microciclos en cada cluster, error medio del CTC y tiempo de construcción del CTC. Estos resultados se encuentran en la sección C del apéndice, donde se obtienen los resultados en los cuales el algoritmo con mejor consistencia es el algoritmo ILS-RANDOM respecto a la variabilidad de cada aspecto anteriormente mencionados, y respecto al tiempo de construcción del CTC, de igual manera con el refinamiento de parámetros el algoritmo clasifica de una mejor manera los microciclos en cada cluster, evitando los microciclos atípicos anteriormente vistos.

Teniendo esto en consideración, procedió a realizar la construcción de los CTCs para cada una de las bases de datos, realizando un análisis del tipo de microciclos utilizados para su construcción. Los CTCs construidos son presentados a continuación.



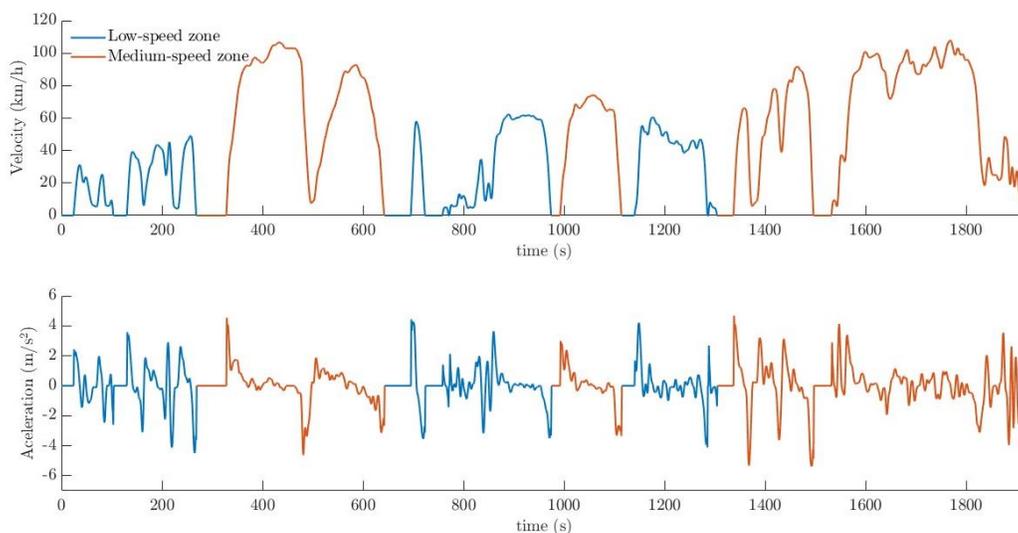
*Figura 5.13.* CTC construido para El Paso con el refinamiento de parámetros

Este CTC fue encontrado en 22945 iteraciones, las cuales corresponden a la cantidad de iteraciones necesarias para encontrar el CTC con el menor error medio posible de los candidatos construidos, presentando un error medio de  $\sim 3.83\%$ , y una duración 2056 segundos  $\sim 34.3$  minutos. Las proporciones de velocidad están enfocadas principalmente en rangos de velocidad media y alta, sin embargo, también considera velocidades medias en menor proporción. El Paso CTC, solamente cuenta con dos tipos de microciclos en su construcción, por esta razón este CTC es considerado como un CTC de baja y alta velocidad, con una considerable cantidad de microciclos de baja velocidad en el inicio, y un microciclo de alta velocidad al final.



*Figura 5.14.* CTC construido para San Antonio con el refinamiento de parámetros

El CTC presentado en la Figura 5.14, corresponde al CTC desarrollado para la base de datos de San Antonio, en general Este CTC fue construido en 47834 iteraciones, tiene un error medio  $\sim 4.15\%$ , y una duración de  $\sim 2048$  segundos,  $\sim 34.1$  minutos. La mayor cantidad de microciclos seleccionados es en este CTC corresponden a microciclos de baja velocidad, y solamente un microciclo de media velocidad. El CTC desarrollado para la base de datos de San Antonio se considera como un CTC, de condiciones tráficas medias y altas, el cual también tiene el consideración, específicamente entre los segundos  $\sim 900$  y  $\sim 1300$ .



*Figura 5.15.* CTC construido para Houston/Galveston con el refinamiento de parámetros

El CTC desarrollado para la base de datos de San Antonio, es presentado en la Figura 5.15, el cual fue construido en 67195 iteraciones, con un error medio del 4.28%, y una duración de ~1919 segundos ~32 minutos. Este CTC es considerado como un CTC de condiciones de tráfico alta y media, debido a los tipos de microciclos utilizados en su construcción, así mismo los rangos de velocidad entre 0 – 20 km/h, presentan el mayor porcentaje de la proporción de velocidad, sin embargo en los microciclos de velocidad promedio intermedia, se consideran condiciones de tráfico bajas.

En la Tabla 5.13, se presenta el resumen de cada uno de los parámetros de los CTCs desarrollados con el refinamiento de parámetros, para cada base de datos.

Tabla 5.13. *Comparación de parámetros de los CTCs desarrollados con el refinamiento de parámetros*

Comparación de parámetros de los CTCs				
Parámetro	Unidad	EL PASO	SAN ANTONIO	HOUSTON
Velocidad máxima	km/h	102.84	119.89	107.88
Velocidad Promedio	km/h	37.94	48.02	45.28
Velocidad promedio en movimiento	km/h	48.72	57.02	54.62
Aceleración máxima	m/s <sup>2</sup>	4.97	4.65	4.68
Desaceleración mínima	m/s <sup>2</sup>	-0.15	-0.15	-0.15
Aceleración promedio	m/s <sup>2</sup>	1.01	1.07	1.07
Desaceleración promedio	m/s <sup>2</sup>	-1.18	-1.12	-1.23
Proporción de aceleración	-	33.12%	32.23%	35.64%
Proporción de desaceleración	-	28.21%	30.71%	30.75%
Proporción de ralenti	-	21.64%	15.23%	16.52%
Rango de velocidad (0-20 km/h)	-	39.06%	31.30%	32.15%
Rango de velocidad (20-40 km/h)	-	14.25%	15.28%	13.76%
Rango de velocidad (40-60 km/h)	-	18.53%	17.43%	17.14%
Rango de velocidad (60-80 km/h)	-	11.82%	13.38%	13.65%
Rango de velocidad (>80 km/h)	-	16.34%	22.61%	23.29%
Fuerzas aerodinámicas	m <sup>3</sup> /s <sup>3</sup>	4089.96	7423.15	5766.13
Aceleración mínima	m/s <sup>2</sup>	0.15	0.15	0.15
Desaceleración máxima	m/s <sup>2</sup>	-4.97	-6.44	-5.38
Desviación estándar de la velocidad	km/h	32.94	38.63	35.26
Desviación estándar de la desaceleración	m/s <sup>2</sup>	1.06	1.10	1.18
Desviación estándar de la aceleración	m/s <sup>2</sup>	0.88	0.92	0.96

Al igual, que los CTCs calculados con todos los parámetros también se realizó la comparación respecto del error medio de los CTCs construidos para cada base de datos respecto al error medio de los CTCs estándar respecto a cada base de datos, esta información es presentada en la Tabla 5.14.

Tabla 5.14. Comparación de los errores medios para cada base de datos para los CTC con refinamiento de parámetros.

	EL PASO	SAN ANTONIO	HOUSTON
Ciclo típico de conducción	Error (%)	Error (%)	Error (%)
CTC construido	3.51	4.15	4.28
FTP-75	53.61	58.07	57.31
WLTC Clase 3	30.85	25.78	26.60
WLTC Clase 2	41.91	42.77	43.06
WLTC Clase 1	52.47	56.84	55.95
LA 92	35.27	39.71	39.16

Los resultados obtenidos muestran que los ciclos construidos con base en métodos de aprendizaje automático específicamente el algoritmo ILS-RANDOM, presentan errores ~10 veces menores, respecto al error más bajo de los ciclos estándar analizado. Valores tan bajos de error medio respecto a cada base de datos muestra la viabilidad y robustez de la metodología planteada para la construcción de CTCs, los cuales representes la realidad de los vehículos.

De igual manera se realizó un análisis del error medio de los parámetros respecto al error medio de cada CTC construido, para así visualizar el aporte de cada parámetro al error medio del CTC.

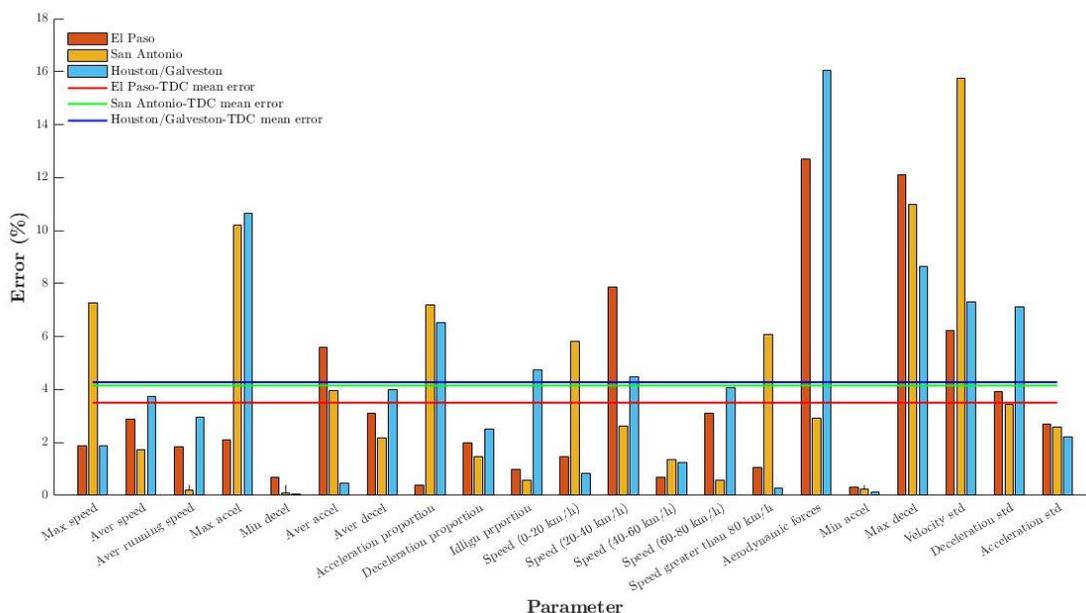


Figura 5.16. Comparación del error medio de los parámetros refinados, respecto al error medio de los CTCs construidos.

Debido a que el error medio de los CTCs disminuyo, una mayor cantidad de parámetros exceden el error medio de cada CTC, sin embargo los resultados presentan errores aceptables razón por la cual no se realiza otro refinamiento de parámetros. En este caso el aporte del error medio de

los parámetros en su valor máximo es ~17%, a comparación del ~50% presentado en la Figura 5.12.

En general los CTCs construidos con el refinamiento de parámetros presentan buenos resultados de error medio, corroborando así una mayor representatividad de los vehículos en la zona de estudio.

### **5.3. Construcción del CTC para la zona Bogotá Región – CTC Bogotá**

#### **5.3.1. Validación de la metaheurística ILS-RANDOM para la base de datos Bogotá-región.**

Teniendo en consideración que para la construcción de los CTCs de la sección anterior se utilizó el algoritmo ILS-RANDOM, este presenta ciertas diferencias en sus soluciones, específicamente para la base de datos Houston/Galveston, para los datos sin parametrizar. Se realiza la respectiva validación para la base de datos de Bogotá Región.

A los parámetros refinados de la Tabla 5.12 se agregan dos parámetros más, los cuales están disponibles a partir de los datos medidos por el dispositivo IVDR: la carga del motor (%) y consumo de combustible en L/km. Para realizar el cálculo de este último parámetro se evalúa la integral a través del tiempo del consumo de combustible a lo largo del CTC o microciclo y se divide por la distancia del CTC o microciclo, según sea el caso. Un parámetro más es considerado, el tipo de vehículo utilizado, que para este caso se analizará con y sin este parámetro.

La base de datos de Bogotá Región consta de 59 recorridos y 1152 microciclos, de los cuales solo se tienen en consideración 961 microciclos. Para la construcción del CTC, se retiran los microciclos atípicos que presentan consumos de combustible mayor 1 L/km, lo cual es demasiado combustible para un vehículo ligero.

El comportamiento de las metaheurísticas respecto al valor de la función de costo encontrada para la base de datos, incluyendo la variable del tipo de vehículo y sin la variable del tipo del vehículo, se presenta en la Tabla 5.15.

Tabla 5.15. *Resultados generales de las metaheurísticas utilizadas para zona Bogotá región*

Metaheurística	BOGOTÁ REGIÓN SIN VARIABLE VEHÍCULO		BOGOTÁ REGIÓN CON VARIABLE VEHÍCULO	
	Valor de la función de costo	iteraciones	Valor de la función de costo	iteraciones
GRASP - RANDOM	3.02111	108	2.37982	127
GRASP - FORGY	3.02111	154	2.37960	104
ILS - RANDOM	3.02111	136	2.38266	101
ILS - FORGY	3.02111	171	2.37944	106
GENETIC K-MEANS	3.02111	4555	2.37944	4911

Los resultados obtenidos por cada una de las metaheurísticas utilizadas son similares entre sí, presentando los mismos valores de función de costo, lo que ratifica los resultados de la sección anterior, donde se obtienen resultados similares sin importar la metaheurística utilizada. Sin embargo, las metaheurísticas ILS presentan un mejor comportamiento respecto a las demás para la base de datos la cual incluye la variable tipo de vehículo, con resultados ligeramente diferentes para la metaheurística ILS-RANDOM.

Lo mismo sucede para la distribución de los microciclos de cada base en los diferentes clusters y la distribución de velocidad de la velocidad media para cada cluster en cada base de datos, tal como se presenta en la Tabla 5.16 y la Tabla 5.17.

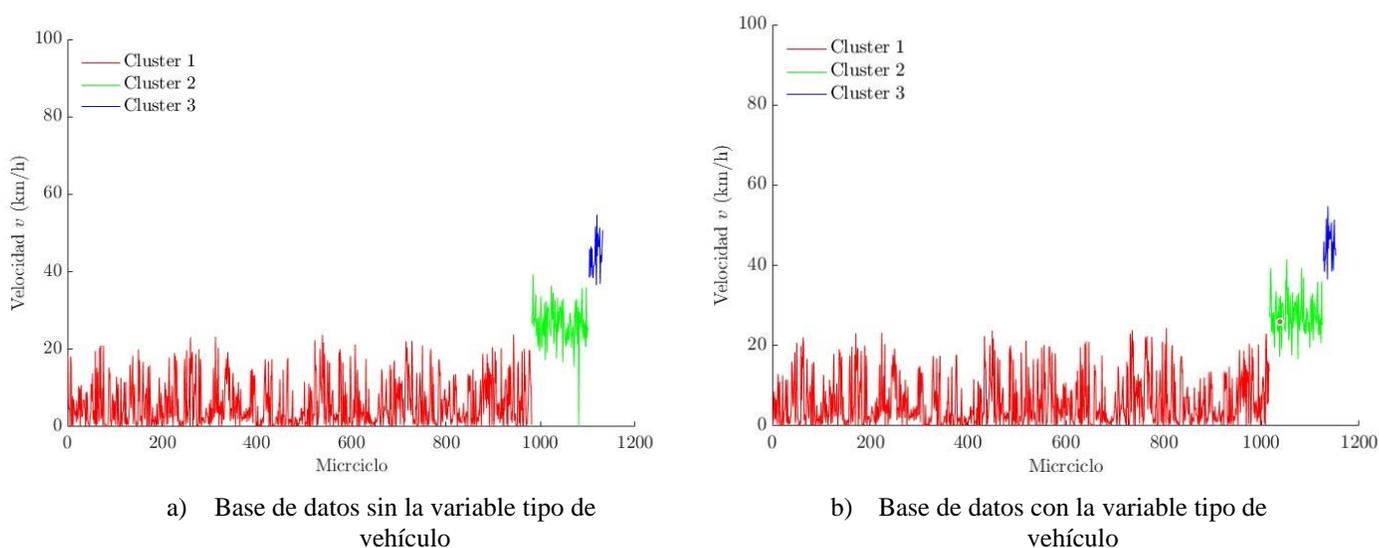
Tabla 5.16. *Distribución de porcentajes en cada microciclo para cada metaheurística en la base de datos Bogotá región.*

Metaheurística	BOGOTÁ REGIÓN SIN VARIABLE VEHÍCULO			BOGOTÁ REGIÓN CON VARIABLE VEHÍCULO		
	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)
GRASP - RANDOM	86.70	10.70	2.56	86.68	10.76	2.52
GRASP - FORGY	86.70	10.70	2.56	86.77	10.67	2.52
ILS - RANDOM	86.70	10.70	2.56	88.16	9.55	2.26
ILS - FORGY	86.70	10.70	2.56	86.86	10.59	2.52
GENETIC K-MEANS	86.70	10.70	2.56	86.86	10.59	2.52

Tabla 5.17. *Velocidad media de los microciclos asignados en cada clúster para la base de datos Bogotá región.*

Metaheurística	BOGOTÁ REGIÓN SIN VARIABLE VEHÍCULO			BOGOTÁ REGIÓN CON VARIABLE VEHÍCULO		
	Cluster 1 (km/h)	Cluster 2 (km/h)	Cluster 3 (km/h)	Cluster 1 (km/h)	Cluster 2 (km/h)	Cluster 3 (km/h)
GRASP - RANDOM	5.68	26.37	44.44	5.54	26.36	44.44
GRASP - FORGY	5.68	26.37	44.44	5.56	26.40	44.44
ILS - RANDOM	5.68	26.37	44.44	5.80	27.52	45.03
ILS - FORGY	5.68	26.37	44.44	5.57	26.43	44.44
GENETIC K-MEANS	5.68	26.37	44.44	5.57	26.43	44.44

En la Figura 5.17 también se analiza de manera gráfica la distribución de microciclos en cada cluster acorde a la velocidad media de cada uno de estos. Para el algoritmo ILS-RANDOM, los resultados muestran que, al igual que las bases de datos de EL Paso, San Antonio y Houston/Galveston, la mayoría de los microciclos corresponden al primer cluster de velocidades promedio bajas. Luego, este es seguido por el cluster 2 de velocidad intermedia, y el cluster 3, de velocidad promedio. Sin embargo, se presenta un microciclo atípico para la base de datos que no tiene en consideración el vehículo utilizado, el cual debería estar clusterizado en el cluster 1, pero se encuentra en el cluster 2. Este resultado, similar al obtenido para la base de datos Houston/Galveston, en este caso se genera por la combinación de parámetros utilizados, además de la velocidad, para asignar este microciclo a dicho cluster. Es importante mencionar que no se presenta una gran cantidad de superposición en los espacios de transición de un cluster a otro. Las distribuciones de las diferentes metaheurísticas se encuentra en el anexo en la sección C.



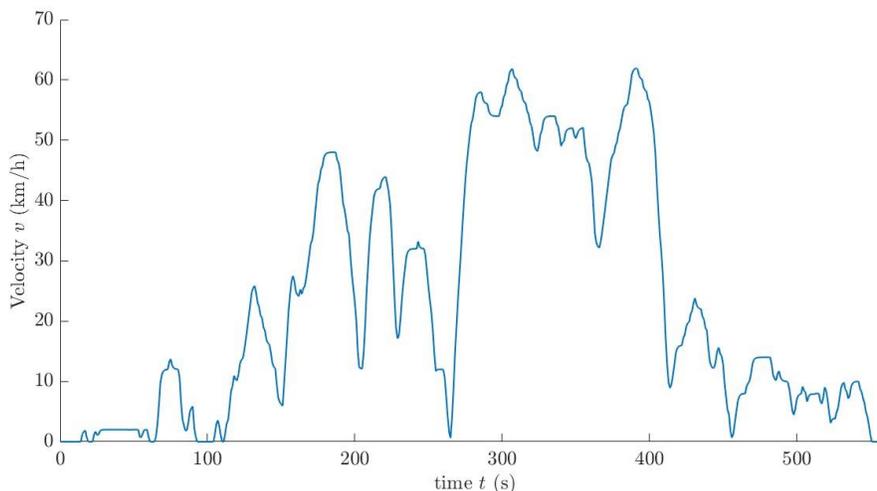
*Figura 5.17.* Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística ILS -RANDOM.

Los resultados obtenidos por la metaheurística ILS-RANDOM presentan pequeñas diferencias respecto a las demás metaheurísticas comparadas. Por esta razón, y teniendo en consideración que presentó un desempeño destacable en tiempo de ejecución de construcción del CTC de las bases de datos proporcionadas por NREL, esta metaheurística es seleccionada para la construcción del CTC para la zona Bogotá Región.

De esta gráfica es posible analizar al comprar las mismas gráficas desarrolladas para las bases de datos de El Paso, San Antonio, y Houston/Galveston, que es posible recolectar una mayor cantidad de datos para poder realizar un análisis aún más representativo del comportamiento de los vehículos.

### 5.3.2. Construcción, comparación y selección del CTC Bogotá

A partir de las bases de datos de la zona Bogotá región, tanto con la inclusión del parámetro “tipo de vehículo” como sin este, se construyeron en total seis CTCs teniendo en consideración que se cuenta con tres métodos para su construcción (MWD-CP, Error medio, Metodología combinada MWD-CP y error medio), así como las bases de datos con y sin el parámetro vehículo. A continuación, se presentan cada uno de estos CTCs desarrollados para la zona Bogotá región.



*Figura 5.18.* Ciclo seleccionado por la metodología MWD-CP.

En la Figura 5.18, se presenta el CTC seleccionado por la metodología MWD-CP, el cual tiene un error medio de  $\sim 18.72\%$  y  $\sim 18.73\%$ , sin considerar el parámetro tipo de vehículo y considerándolo respectivamente. Este CTC corresponde a la selección de los datos de un recorrido cuya diferencia ponderada sea la menor respecto a la base de datos, Sin embargo, este CTC presenta una representatividad menor en el análisis con los CTCs anteriormente mencionados. El mismo ciclo es seleccionado al considerar y no considerar el parámetro tipo de vehículo.

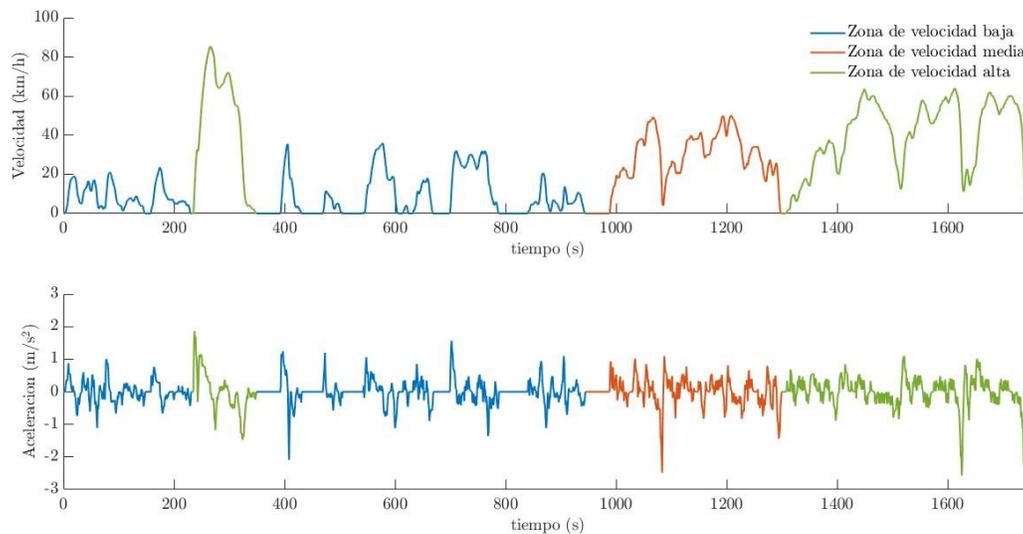
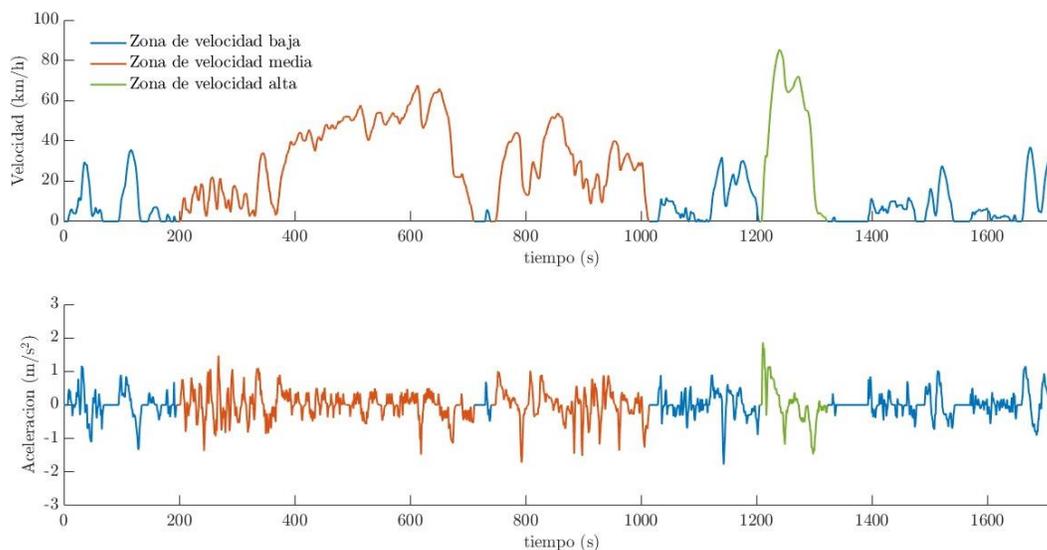


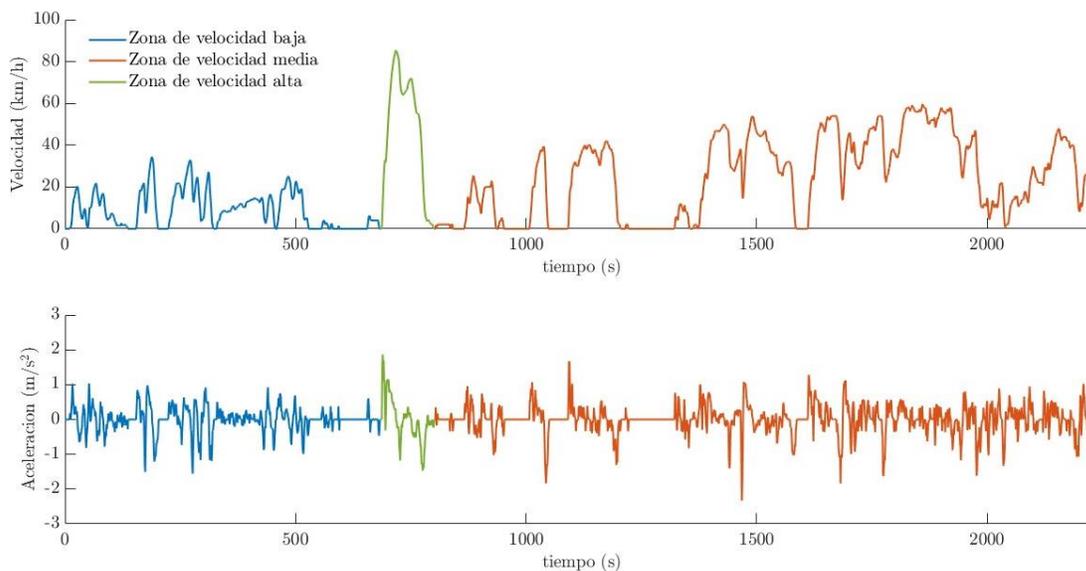
Figura 5.19. CTC desarrollado para la zona Bogotá-región con base al error medio general y sin considerar el tipo de vehículo.

Es ciclo desarrollado con base en la Ecuación 4.1, es decir la metodología del error medio con pesos iguales, el cual fue desarrollado en 37609 iteraciones presentó un error medio respecto a la base de datos de  $\sim 7.34\%$  y una duración de 1751 segundos ( $\sim 29.2$  minutos). Este CTC desarrollado presenta microciclos de velocidad promedio alta media y baja, permitiendo así representar todas las condiciones de tráfico de la zona analizada. Sin embargo, presenta su mayor porcentaje de rangos velocidad en los rangos de velocidad de 0-20 km/h, como es de esperarse en una ciudad con las condiciones de tráfico, con altos niveles de congestión y bajos límites de velocidad como lo es Bogotá.



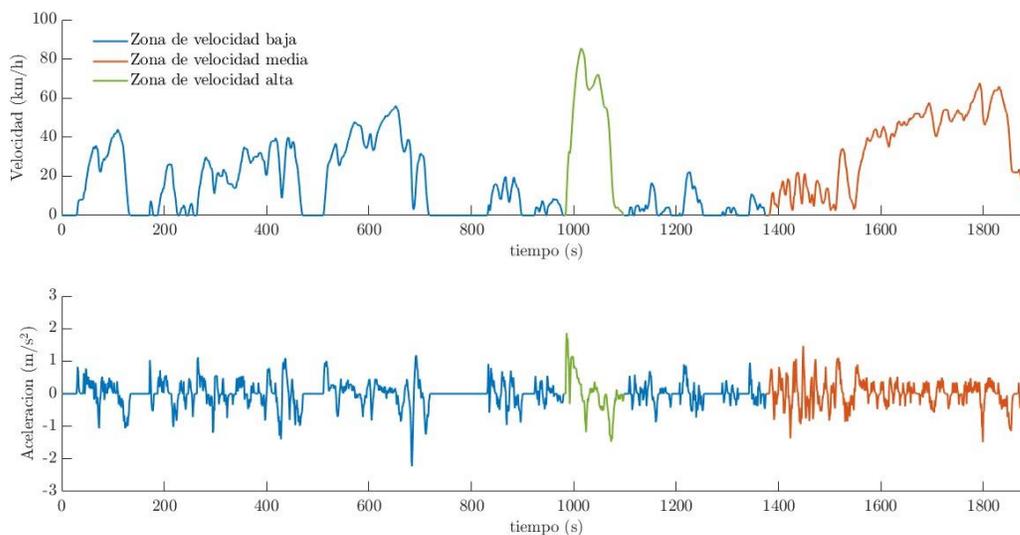
*Figura 5.20.* CTC desarrollado para la zona Bogotá-región con base al error medio general y considerando el tipo de vehículo.

El CTC desarrollado y presentado en la Figura 5.20 fue construido con base el error medio con pesos iguales, de la misma que el ciclo anteriormente presentado. Este CTC fue desarrollado en 29481 iteraciones, con un error del  $\sim 8.93\%$  y una duración de  $\sim 1724$  segundos ( $\sim 28.7$  minutos). Este ciclo, al igual que el anterior, contiene microciclos de todos los tipos en su construcción, permitiendo así representar de manera general el comportamiento de los vehículos en la zona Bogotá región. Sin embargo, este microciclo contiene una mayor cantidad de microciclos de velocidad media y un solo microciclo de velocidad alta, pero considera el comportamiento de los vehículos en todos los rangos de velocidad. Así mismo, al tener el parámetro “tipo de vehículo” en consideración, es posible analizar que el aporte en este ciclo por parte de ambos vehículos experimentales fue casi el mismo, teniendo un valor para este parámetro de 0.51.



*Figura 5.21.* CTC desarrollado para la zona Bogotá-región con base la metodología combinada y sin considerar el tipo de vehículo.

El CTC desarrollado y presentado en la Figura 5.21 fue construido con base en la metodología combinada con base en la Ecuación 4.12, teniendo en consideración los pesos relativos asignados a cada uno de los parámetros como parámetros predictores del consumo de combustible. Este CTC fue construido en 52131 iteraciones, con un error del 4.79% y una duración ~2240 segundos (~37.33 minutos). Este ciclo, al igual que los anteriores, representa la variabilidad del tráfico en la zona Bogotá región, debido al que contiene en su construcción todos los tipos microciclos. Sin embargo, presenta una mayor cantidad de ciclos de velocidad promedio intermedia en su construcción.



*Figura 5.22.* CTC desarrollado para la zona Bogotá-región con base la metodología combinada y considerando el tipo de vehículo.

El CTC presentado en la Figura 5.22 fue construido con base en la combinación de los métodos de aprendizaje automático y la metodología MWD-CP, teniendo en consideración el parámetro “tipo de vehículo”. Este CTC se desarrolló en 52474 iteraciones, con un error medio respecto a la base de datos del  $\sim 4.81\%$  y con una duración de  $\sim 1892$  segundos ( $\sim 31.5$  minutos). Al igual que, la mayoría de CTCs desarrollados hasta el momento, este CTC representa en general el comportamiento de los vehículos de la zona Bogotá región por la inclusión de todos los tipos de microciclos. Así mismo, este ciclo representa el consumo de combustible de los vehículos analizados, debido al uso de la metodología MWD-CP donde los pesos de cada parámetro son asignados en función del consumo de combustible. Para este CTC el vehículo que tuvo más influencia fue el vehículo Nissan Versa, con un valor medio de 0.67 para el parámetro tipo de vehículo, donde el valor medio de los microciclos utilizados tiende hacia el valor 1, el cual es asignado a este vehículo.

Es importante mencionar que la metodología MWD-CP, escoge el ciclo que mejor representa la base de datos en generar, por lo cual si se agregan recorridos con mayor duración, que mejor representen la base de datos, esta metodología seleccionara estos ciclos, presentando recorridos de una extensa duración, lo cual no es recomendable para el desarrollo de pruebas dinámicas. En comparación la metodología de aprendizaje automático permite desarrollar CTCs con duraciones

mínimas o máximas determinadas, construyendo el CTC que mejor represente la base de datos mediante la unión de los diferentes microciclos.

Teniendo en consideración el uso de la metodología MWD-CP y la metodología combinada entre los métodos de aprendizaje automático y los pesos calculados mediante la metodología RWA utilizada en el método MWD-CP, es necesario presentar los pesos asignados a cada uno de los parámetros de análisis en la Tabla 5.18.

Tabla 5.18. *Pesos relativos de cada uno de los parámetros característicos en cada metodología desarrollada. (CV – con el parámetro tipo de vehículo, SV - sin el parámetro tipo de vehículo)*

Parámetro	K-MEANS SV	K-MEANS CV	K-MEANS - MWD-CP - SV	K-MEANS - MWD-CP - CV	MWD-CP - SV	MWD-CP - CV
Velocidad máxima	6.11%	5.47%	6.11%	5.47%	6.11%	5.47%
Velocidad promedio	5.94%	5.44%	5.94%	5.44%	5.94%	5.44%
Velocidad promedio en movimiento	6.79%	6.48%	6.79%	6.48%	6.79%	6.48%
Aceleración máxima	2.84%	2.98%	2.84%	2.98%	2.84%	2.98%
Desaceleración mínima	6.29%	6.48%	6.29%	6.48%	6.29%	6.48%
Aceleración promedio	2.06%	2.57%	2.06%	2.57%	2.06%	2.57%
Desaceleración promedio	2.57%	3.41%	2.57%	3.41%	2.57%	3.41%
Proporción de aceleración	5.16%	4.91%	5.16%	4.91%	5.16%	4.91%
Proporción de desaceleración	4.64%	4.43%	4.64%	4.43%	4.64%	4.43%
Proporción de ralentí	11.24%	11.92%	11.24%	11.92%	11.24%	11.92%
Rango de velocidad (0-20 km/h)	4.55%	4.34%	4.55%	4.34%	4.55%	4.34%
Rango de velocidad (20-40 km/h)	3.13%	3.23%	3.13%	3.23%	3.13%	3.23%
Rango de velocidad (40-60 km/h)	3.06%	2.68%	3.06%	2.68%	3.06%	2.68%
Rango de velocidad (60-80 km/h)	1.21%	0.98%	1.21%	0.98%	1.21%	0.98%
Rango de velocidad (>80 km/h)	0.49%	0.46%	0.49%	0.46%	0.49%	0.46%
Fuerzas aerodinámicas	3.22%	2.78%	3.22%	2.78%	3.22%	2.78%
Aceleración mínima	9.78%	10.81%	9.78%	10.81%	9.78%	10.81%
Desaceleración máxima	3.51%	3.46%	3.51%	3.46%	3.51%	3.46%
Desviación estándar de la velocidad	6.91%	6.88%	6.91%	6.88%	6.91%	6.88%
Desviación estándar de la desaceleración	2.59%	2.74%	2.59%	2.74%	2.59%	2.74%
Desviación estándar de la aceleración	2.01%	2.31%	2.01%	2.31%	2.01%	2.31%
Carga del motor	5.92%	5.22%	5.92%	5.22%	5.92%	5.22%

Estos pesos relativos permiten conocer la influencia de cada uno de los parámetros en el consumo de combustible. En este caso, la variable que tiene un mayor peso es la proporción de ralentí, seguida por aceleración mínima. De acuerdo con esto, es posible argumentar que, debido al tráfico de la ciudad, los periodos largos de espera mientras el vehículo está en reposo y las aceleración requeridas para desplazar el vehículo después de estos periodos, son los parámetros que aumentan el consumo de combustible y, por ende, la cantidad de emisiones.

De igual manera, se calcularon cada uno de los parámetros de los CTCs construidos o seleccionados tal como se presentan en la Tabla 5.19.

Tabla 5.19. Comparación de los parámetros de los CTCs construidos para la zona Bogotá región.

Comparación de los parámetros de los CTCs construidos							
Parámetro	Unidad	K-MEANS SV	K-MEANS CV	K-MEANS – MWD-CP – SV	K-MEANS – MWD- CP-CV	MWD-CP- SV	MWD-CP- CV
Velocidad máxima	km/h	85.40	85.40	85.40	85.40	61.93	61.93
Velocidad promedio	km/h	22.70	21.40	21.01	21.71	22.98	22.98
Velocidad promedio en movimiento	km/h	27.63	25.52	26.41	27.51	25.22	25.22
Aceleración máxima	m/s <sup>2</sup>	1.87	1.87	1.87	1.87	1.57	1.57
Desaceleración mínima	m/s <sup>2</sup>	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15
Aceleración promedio	m/s <sup>2</sup>	0.46	0.47	0.46	0.45	0.48	0.48
Desaceleración promedio	m/s <sup>2</sup>	-0.49	-0.48	-0.53	-0.49	-0.52	-0.52
Proporción de aceleración	-	25.53%	27.96%	25.58%	26.43%	29.26%	29.26%
Proporción de desaceleración	-	24.61%	27.78%	22.41%	24.52%	27.13%	27.13%
Proporción de ralentí	-	17.02%	14.50%	19.60%	19.98%	7.80%	7.80%
Rango de velocidad (0-20 km/h)	-	53.06%	56.73%	55.45%	54.55%	55.67%	55.67%
Rango de velocidad (20-40 km/h)	-	24.79%	20.88%	24.51%	22.94%	17.55%	17.55%
Rango de velocidad (40-60 km/h)	-	17.19%	17.05%	17.54%	17.65%	24.29%	24.29%
Rango de velocidad (60-80 km/h)	-	4.28%	4.64%	1.96%	4.23%	2.48%	2.48%
Rango de velocidad (>80 km/h)	-	0.69%	0.70%	0.54%	0.63%	0.00%	0.00%
Fuerzas aerodinámicas	m <sup>3</sup> /s <sup>3</sup>	1029.91	960.25	842.80	960.42	943.96	943.96
Aceleración mínima	m/s <sup>2</sup>	0.15	0.15	0.15	0.15	0.15	0.15
Desaceleración máxima	m/s <sup>2</sup>	-2.58	-1.78	-2.33	-2.22	-1.59	-1.59
Desviación estándar de la velocidad	km/h	20.99	20.77	19.77	20.85	20.02	20.02
Desviación estándar de la desaceleración	m/s <sup>2</sup>	0.37	0.30	0.36	0.30	0.33	0.33
Desviación estándar de la aceleración	m/s <sup>2</sup>	0.28	0.27	0.26	0.27	0.31	0.31
Carga de motor	%	35.85	38.80	36.81	35.16	25.42	25.42
Tipo de vehículo (promedio)	-	-	0.51	-	0.67	-	1.00
Consumo de combustible	L/km	0.06	0.06	0.06	0.06	0.07	0.07

A partir del desarrollo de estos seis CTCs, se compararon entre sí y respecto a los CTCs estándar internacionales, como se presenta en la Tabla 5.20. Aquí se calculó el respectivo error medio y el valor de las diferencias ponderadas como fue presentado en la sección 3.3.1. para cada ciclo desarrollado.

Tabla 5.20. Comparación del error medio y las diferencias ponderadas para todos los ciclos construidos para la zona Bogotá región.

	<b>K-MEANS SV</b>	<b>K-MEANS CV</b>	<b>K-MEANS - MWD-CP - SV</b>	<b>K-MEANS - MWD-CP-CV</b>	<b>MWD-CP - SV</b>	<b>MWD-CP - CV</b>
<b>Ciclo Típico de conducción</b>	<b>Error (%)</b>	<b>Error (%)</b>	<b>Error (%)</b>	<b>Error (%)</b>	<b>Error (%)</b>	<b>Error (%)</b>
CTC desarrollado	7.34	8.93	4.79	4.81	18.72	18.73
FTP-75	97.72	98.23	97.72	98.23	-	-
WLTC Clase 3	179.19	179.98	179.19	179.98	-	-
WLTC Clase 2	57.22	57.65	57.22	57.65	-	-
WLTC Clase 1	30.62	30.81	30.62	30.81	-	-
LA 92	158.91	159.59	158.91	159.59	-	-
<b>MWD-CP</b>	3.26	1.20	5.02	1.16	28.81	13.25

Los resultados presentados, muestran que los ciclos construidos mediante métodos de aprendizaje automático y métodos estadísticos tienen un menor error medio que los ciclos estándar. En el caso más extremo del ciclo WLTC clase 3, se alcanzan errores del ~179.98%, desviándose por completo de la realidad de los vehículos en la zona de análisis.

Así mismo, los métodos de aprendizaje automático presentan mejores resultados que los métodos estadísticos, incluso utilizando la misma métrica de la metodología MWD-CP para la evaluación de los ciclos construidos con el algoritmo k-means respecto a los construidos con la metodología MWD-CP. Los CTCs construidos en combinación de las metodologías de aprendizaje automático y los métodos estadísticos tienen un mejor comportamiento con base en el error medio, esto debido a la aplicación de los pesos relativos de cada parámetro.

Finalmente, con el objetivo de la validación de la metodología para la evaluación ambiental, se escoge el CTC con el menor error medio, el cual representa de una mejor manera el comportamiento de los vehículos en la zona Bogotá Región. Para estos efectos, el CTC Bogotá es el calculado mediante el uso de las metodologías combinadas, el cual no considera el parámetro vehículo y el cual es presentado en la

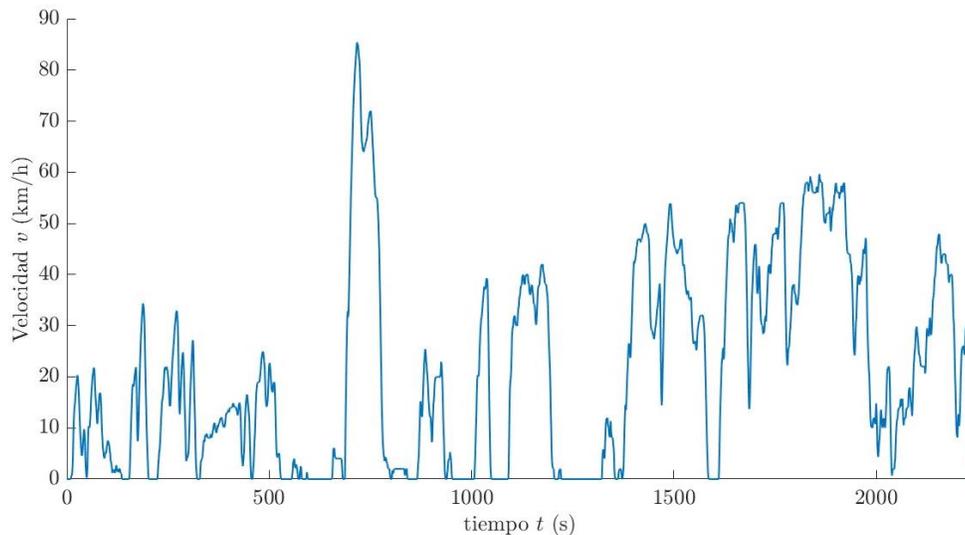
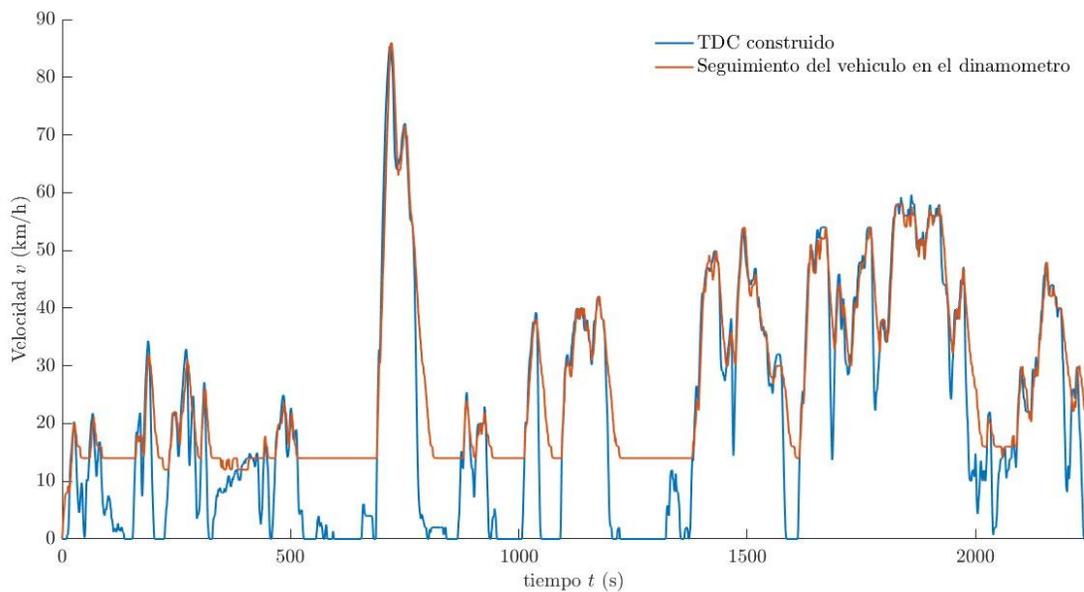


Figura 5.23. Ciclo seleccionado – Ciclo Bogotá

#### 5.4. Pruebas dinámicas preliminares del CTC Bogotá y CTC estándar.

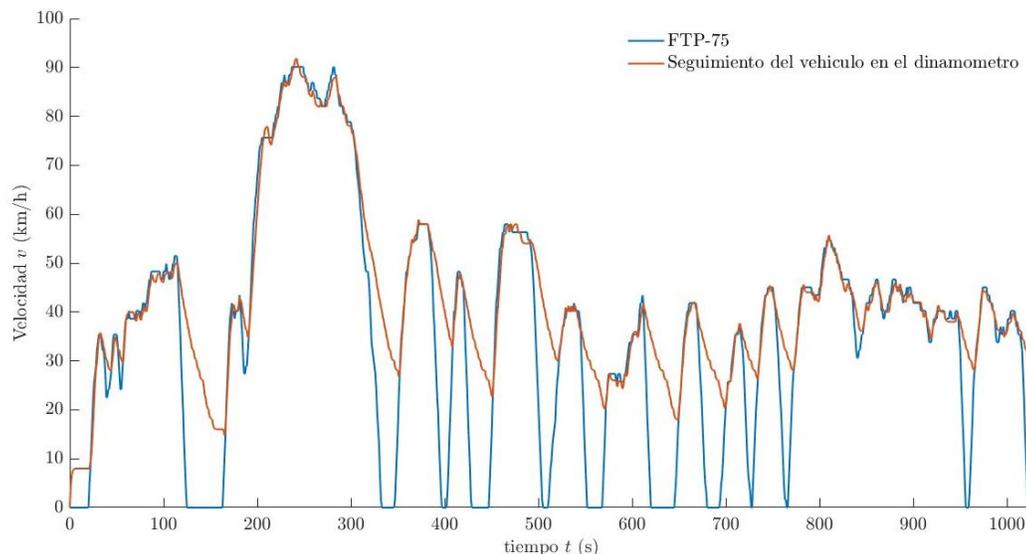
Para la validación operacional del ciclo construido, este debe ser puesto a prueba en el dinamómetro de chasis, donde el vehículo experimental siga la línea de velocidad mientras se toman las muestras de los gases de escape. En la siguiente figura se muestra el seguimiento del vehículo experimental al CTC Bogotá.



*Figura 5.24.* Seguimiento del vehículo experimental al CTC construido para la zona Bogotá región.

En la Figura 5.24 se evidencia que gráfica de velocidad seguida por el vehículo experimental no coincide en todos los puntos con el ciclo construido, especialmente para las bajas velocidades. Esto se debe a que no se presionó el freno durante las pruebas, esto con el objetivo de salvaguardar la transmisión automática del vehículo experimental debido que los rodillos del dinamómetro de chasis pueden tener una gran cantidad de inercia la cual puede dañar el vehículo. Así mismo, el vehículo tampoco se llevó a velocidad de reposo, sino que se mantuvo conectado la caja de transmisión a los ejes en la velocidad mínima del estado de “Driving” del vehículo. Esta velocidad es la velocidad de transferida a las ruedas por el vehículo en su estado de ralentí, por lo cual no se supone una mayor carga al motor y se establecen la misma cantidad de emisiones. Sin embargo, es importante mencionar que parte emisiones presentadas en los procesos de aceleración desde reposo a una velocidad determinada, no se tienen en consideración, debido a que no se llevo el vehículo a reposo, con excepción del inicio del CTC.

En los tramos de altas velocidades el vehículo coincide por completo con la velocidad establecida por el ciclo. Esto mismo sucede con los ciclos estándar puestos probados en el mismo vehículo, los cuales se muestran en la Figura 5.25 y Figura 5.26.



*Figura 5.25.* Seguimiento del vehículo experimental al CTC FTP-75.

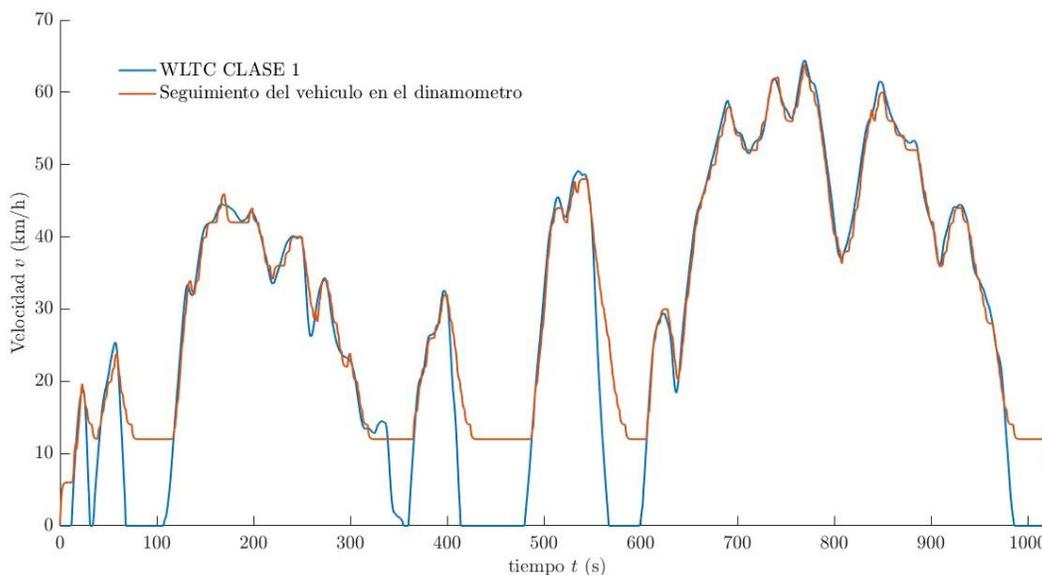


Figura 5.26. Seguimiento del vehículo experimental al CTC WLTC-clase 1.

### 5.5. Demostración de determinación de factores de emisión

Los factores de emisión fueron calculados únicamente para el CTC Bogotá de manera demostrativa de la metodología debido a la disponibilidad del analizador de gases. Sin embargo, estos factores de emisión fueron calculados para gases como  $\text{NO}_2$ ,  $\text{NO}$ ,  $\text{CO}_2$ ,  $\text{CO}$ ,  $\text{HC}$  y  $\text{O}_2$ . Para esto, en primer lugar, se calcularon todos valores de los flujos máxicos de cada uno de los contaminantes tomados por la sonda de la muestra instalada en el tubo de escape. Al integrarlos en el tiempo prueba del CTC Bogotá, y dividirlo por la distancia recorrida del este mismo, se evalúan los resultados de factores de emisión para cada contaminante del aire medido. En la Tabla 5.21 se presentan estos factores de emisión y se comparan con los límites de emisiones propuestos por la Secretaría Distrital de Ambiente (Secretaría de Ambiente de Bogotá, 2022) y por el Ministerio de ambiente y desarrollo sostenible (Ministerio Ambiente Y Desarrollo Sostenible, 2022).

Tabla 5.21. *Comparación de los factores de emisión para el vehículo de pruebas dinámica Nissan Versa 2013.*

Gas de escape	Factor de emisión (gr/km)	Límites de emisión (gr/km)	
	CTC-Bogotá	Secretaría Distrital de Ambiente	Ministerio de Ambiente
NO <sub>2</sub>	1.25x10 <sup>-6</sup>	6.2x10 <sup>-2</sup>	0.25
NO	1.54x10 <sup>-4</sup>	6.2x10 <sup>-2</sup>	0.25
CO <sub>2</sub>	0.60	-	-
CO	0.35	9.19	2.11
HC	9.56x10 <sup>-5</sup>	-	0.25
O <sub>2</sub>	0.01	-	-

Al realizar la comparación de los factores de emisión calculados con el CTC Bogotá respecto a los límites establecidos por la Secretaría Distrital de Ambiente y el Ministerio de Ambiente, se evidencia que existen brechas significativas de hasta 4 órdenes de magnitud para NO<sub>2</sub> entre la regulación distrital vigente y las emisiones bajo condiciones de CTC local, incluso para vehículos ligeros con más de 10 años de operación. Para las pruebas dinámicas y mediciones realizadas con el CTC Bogotá bajo condiciones locales de comportamiento de vehículos, se puede esperar con cierta seguridad que los límites de emisiones a nivel distrital sean demasiado laxos para los vehículos ligeros a gasolina que circulan en la ciudad y sus alrededores.

Esto mismo sucede con los factores establecidos como límites por el Ministerio de Ambiente, los cuales se deben soportar con pruebas dinámicas, específicamente con el CTC FTP-75. Los límites de emisiones son demasiado laxos respecto a los valores hallados con el CTC Bogotá, el cual permitiría el desarrollo de normas más efectivas para la disminución de contaminantes y mejora de tecnologías de combustión interna que ingresan al país, respecto a la normativa del desarrollo de pruebas dinámicas para fines de homologación.

Los valores de eficiencia energética, asociadas al consumo de combustible, obtenidos en la prueba son los siguientes: 10.9x10<sup>-3</sup> gal/km para el CTC Bogotá y 8.8x10<sup>-3</sup> gal/km para los ciclos FTP-75 y ciclo WLTC clase 1. Si bien el valor del consumo de combustible calculado mediante la prueba dinámica para el CTC Bogotá concuerda con los datos calculados para su construcción, se evidencia una diferencia de 2.2x10<sup>-3</sup> L/km para Bogotá, lo que puede ser debido al tráfico al cual se encuentran expuestos los vehículos en la ciudad.

## 6. Conclusiones

En este proyecto se llevó a cabo el desarrollo y validación de una metodología para la evaluación medioambiental del parque automotor ligero de la zona Bogotá – Región, a partir de la

construcción de ciclos típicos de conducción utilizando métodos estadísticos y de aprendizaje automático. Para esto, se desarrolló un dispositivo IVDR, el cual permitió recolectar datos de operación de la ECU de diferentes vehículos en vías principales representativas de la ciudad de Bogotá y sus alrededores. Estos datos fueron procesados utilizando diferentes metodologías y algoritmos, estadísticos y de aprendizaje autónomo para obtener un ciclo típico de conducción representativo de la zona Bogotá Región, CTC Bogotá. Este ciclo fue utilizado para el desarrollo de pruebas dinámicas demostrativas, de donde se presentó la metodología para la determinación de los factores de emisión de contaminantes de aire característicos para la zona. A Partir de esta investigación se concluye:

- Se desarrollaron dispositivos de adquisición de datos programables para la adquisición de datos para el desarrollo de CTCs, que también pueden ser utilizados como herramienta de diseño vehicular, dada su alta aplicabilidad.
- Se desarrolló una metodología de construcción de ciclos típicos de conducción robusta a través de métodos de aprendizaje de automático, obteniendo resultados similares sin importar la metaheurística utilizada, respecto al error medio del CTC construido. La combinación de los métodos de aprendizaje automático y métodos estadísticos para la disminución del error medio de los CTCs construidos permitió aumentar la representatividad de los CTCs respecto a las bases de datos.
- Se construyeron CTCs a partir de datos locales, dado que los CTC estándar no representan de una manera correcta la realidad al utilizarse en regiones diferentes. Esto se validó por medio de la comparación estadística entre los ciclos construidos con información del lugar de análisis, donde el error medio fue considerablemente menor que los CTC estándar.
- Los factores de emisión obtenidos mediante pruebas dinámicas con el CTC de mejor desempeño construido para la zona Bogotá región, CTC Bogotá, demuestran la laxitud de los límites de emisiones de contaminantes del aire establecidos por las autoridades regulatoria de Bogotá y Colombia para el desarrollo y validación de vehículos mediante pruebas dinámicas.
- La metodología desarrollada y validada es aplicable para la estimación de factores de emisión del parque automotor liviano de la zona Bogotá Región, con el objetivo de

establecer los fundamentos técnicos para informar políticas públicas y regulación que busque la disminución de contaminantes y desarrollo de tecnologías de transporte.

- La metodología presentada puede ser aplicada a diferentes contextos tanto de desarrollo de fundamentos técnicos de políticas públicas en general, como su aplicación en el ámbito académico y como en este caso de investigación.

## 7. Trabajo futuro

Teniendo en consideración los resultados obtenidos, se propone como trabajo futuro:

- Recolectar una mayor cantidad de información mediante los dispositivos IVDR, a través de instalación en diferentes vehículos los cuales realicen recorridos aleatorios en la zona de estudio.
- Investigar en los algoritmos de aprendizaje automático para la construcción de CTCs, específicamente redes neuronales, y algoritmos de aprendizaje profundo, con el objetivo de aumentar la representatividad de los CTCs construidos
- Realizar mediciones de contaminantes a una mayor cantidad de vehículos representativos del parque automotor de la ciudad con el objetivo de determinar sus factores de emisión para propósitos de inventarios de emisiones, homologación de vehículos importados o desarrollo de regulación de emisiones.

## Referencias

- 3DATX. (2023). *par SYNC FLEX PNC Brochure*. [https://3datx.com/wp-content/uploads/3DATX\\_parSYNC\\_FLEX-PNC\\_brochure\\_EU\\_06-02-23\\_v2.pdf](https://3datx.com/wp-content/uploads/3DATX_parSYNC_FLEX-PNC_brochure_EU_06-02-23_v2.pdf)
- Alves, C. A., Lopes, D. J., Calvo, A. I., Evtuyugina, M., Rocha, S., & Nunes, T. (2015). Emissions from light-duty diesel and gasoline in-use vehicles measured on chassis dynamometer test cycles. *Aerosol and Air Quality Research*, 15(1), 99–116. <https://doi.org/10.4209/aaqr.2014.01.0006>
- Berthold, M. R., Lenz, H.-J., Bradley, E., Kruse, R., & Borgelt, C. (2003). *LNCS 2810 - Advances in Intelligent Data Analysis V*.
- Cano, J. R., Herrera, F., Sanchez, L., Cano, J. R., Cordón, O., Herrera, F., & Sánchez, L. (2002). A greedy randomized adaptive search procedure applied to the clustering problem as an initialization process using K-Means as a local search procedure. *Journal of Intelligent & Fuzzy Systems*, 12, 235–242. <https://www.researchgate.net/publication/220256785>
- Chauhan, B., Pithawala, C. K., Joshi, G., Vallabhbai, S., Chauhan, B. P., Joshi, G. J., & Parida, P. (2020). *Development of Candidate Driving Cycles for an Urban Arterial Corridor of Vadodara City*. <https://www.researchgate.net/publication/348252664>

- Clarkson, D., & Middleton, J. T. (1962). The California control program for motor vehicle created air pollution. *Journal of the Air Pollution Control Association*, 12(1), 22–28. <https://doi.org/10.1080/00022470.1962.10468042>
- DANE. (2019). *Estructura de población : Bogotá D. C., CG 2005 y CNPV 2018 Hombres Mujeres*. <https://sitios.dane.gov.co/cnpv/app/views/informacion/fichas/11.pdf>
- Desinedi, R. M., Mahesh, S., & Ramadurai, G. (2020). Developing driving cycles using k-means clustering and determining their optimal duration. *Transportation Research Procedia*, 48(2018), 2083–2095. <https://doi.org/10.1016/j.trpro.2020.08.268>
- Division, C., & Agency, U. S. E. P. (1993). Federal Test Procedure Review Project : Preliminary Technical Report May 1993. *Environmental Protection*, May.
- Doenhoff, von, Marks II, R. J., Choi, J. J., Healy, M., Tsao, C. K., Bezdek, J. C., Pal, N. R., Kohonen, F., Krishna, K., & Narasimha Murty, M. (1999). Dynamic parameter encoding for genetic algorithms. In *Pattern Recognit* (Vol. 29, Issue 3). Morgan Kaufman.
- Duran, A., & Earleywine, M. (2012). GPS data filtration method for drive cycle analysis applications. *SAE Technical Papers*. <https://doi.org/10.4271/2012-01-0743>
- Echeverry Mejía, J. M. (2018). *Metodología para reducir el gasto de combustible en rutas fijas mediante el uso de hábitos de conducción eficiente, empleando un sistema IVDR y ciclos de conducción*.
- Echeverry-Mejía, J., Arenas-Urbe, F., Contreras, D., & Vásquez, V. (2022). Design and Validation of an In-Vehicle Data Recorder System for Testing Purposes. *IEEE Latin America Transactions*, 100(XXX). <https://latam.ieceer9.org/index.php/transactions/article/view/6489/1654>
- Fotouhi, A., & Montazeri-Gh, M. (2013). Tehran driving cycle development using the k-means clustering method. *Scientia Iranica*, 20(2), 286–293. <https://doi.org/10.1016/j.scient.2013.04.001>
- Freescale. (2014). *KEA128 Sub-Family Reference Manual*.
- Frey, C., & Eichenberger, D. (1997). *Remote Sensing of Mobile Source Air Pollutant Emissions : Variability and Uncertainty in On-Road Emissions Estimates of Carbon Monoxide and Hydrocarbons for School and Transit Buses*. June, 168.
- Giakoumis, E. G. (2016). Driving and engine cycles. In *Driving and Engine Cycles*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-49034-2>
- Gitelman, V., Bekhor, S., Doveh, E., Pesahov, F., Carmel, R., & Morik, S. (2018). Exploring relationships between driving events identified by in-vehicle data recorders, infrastructure characteristics and road crashes. *Transportation Research Part C: Emerging Technologies*, 91(April), 156–175. <https://doi.org/10.1016/j.trc.2018.04.003>
- Hass, G. C., & Brubacher, M. L. (1962). A test procedure for motor vehicle exhaust emissions. *Journal of the Air Pollution Control Association*, 12(11), 505–543. <https://doi.org/10.1080/00022470.1962.10468120>
- Ho, S. H., Wong, Y. D., & Chang, V. W. C. (2014). Developing Singapore Driving Cycle for passenger cars to estimate fuel consumption and vehicular emissions. *Atmospheric Environment*, 97, 353–362. <https://doi.org/10.1016/j.atmosenv.2014.08.042>

- Huertas, J. I. (2017). *A new methodology to determine typical driving cycles for the design of vehicles power trains*. 12008. <https://doi.org/10.1007/s12008-017-0379-y>
- Huertas, J. I., Díaz, J., Cordero, D., & Cedillo, K. (2018). A new methodology to determine typical driving cycles for the design of vehicles power trains. *International Journal on Interactive Design and Manufacturing*, 12(1), 319–326. <https://doi.org/10.1007/s12008-017-0379-y>
- Huertas, J. I., Quirama, L. F., Giraldo, M., & Díaz, J. (2019). Comparison of three methods for constructing real driving cycles. *Energies*, 12(4). <https://doi.org/10.3390/en12040665>
- ICONTEC. (2012). Ntc 4983. *Icontec*, 571, 31.
- International Agency for Research on Cancer. (2013). *IARC: Outdoor air pollution a leading environmental cause of cancer deaths*. 221, 1–4. <https://doi.org/10.1002/em>
- International Organization for Standardization (ISO). (2006). *Road vehicles — Controller area network (CAN) — Part 1 : Data link layer and physical signalling*. 2006, 1–6.
- Jia, X., Wang, H., Xu, L., Wang, Q., Li, H., Hu, Z., Li, J., & Ouyang, M. (2021). Constructing representative driving cycle for heavy duty vehicle based on Markov chain method considering road slope. *Energy and AI*, 6, 100115. <https://doi.org/10.1016/J.EGYAI.2021.100115>
- Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, 35(1), 1–19. [https://doi.org/10.1207/S15327906MBR3501\\_1](https://doi.org/10.1207/S15327906MBR3501_1)
- Jun, G., Guensler, R., & Ogle, J. H. (2006). Smoothing methods to minimize impact of global positioning system random error on travel distance, speed, and acceleration profile estimates. *Transportation Research Record*, 1972, 141–150. <https://doi.org/10.3141/1972-19>
- Khadse, V. M., Mahalle, P. N., & Shinde, G. R. (2020). Statistical study of machine learning algorithms using parametric and non-parametric tests: A comparative analysis and recommendations. *International Journal of Ambient Computing and Intelligence*, 11(3), 80–105. <https://doi.org/10.4018/IJACI.2020070105>
- Kharrazi, S., Almen, M., Frisk, E., Nielsen, L., Kharrazi, S., Almén, M., Frisk, E., & Nielsen, L. (2019). *Extending Behavioral Models to Generate Mission-Based Driving Cycles for Data-Driven Vehicle Development* Extending behavioral models to generate mission-based driving cycles for data-driven vehicle development.
- Kim, Y., & Bang, H. (n.d.). *Introduction to Kalman Filter and Its Applications*. [www.intechopen.com](http://www.intechopen.com)
- Lin, J., & Niemeier, D. A. (2002). An exploratory analysis comparing a stochastic driving cycle to California's regulatory cycle. *Atmospheric Environment*, 36(38), 5759–5770. [https://doi.org/10.1016/S1352-2310\(02\)00695-7](https://doi.org/10.1016/S1352-2310(02)00695-7)
- Liu, J., Wang, X., & Khattak, A. (2016). Customizing driving cycles to support vehicle purchase and use decisions: Fuel economy estimation for alternative fuel vehicle users. *Transportation Research Part C: Emerging Technologies*, 67, 280–298. <https://doi.org/10.1016/j.trc.2016.02.016>
- Liu, Y., Xin Wu, Z., Zhou, H., Zheng Nan Yu, H., Pan An, X., Yuan Li, J., & Liang Li, M. (2020). DEVELOPMENT OF CHINA LIGHT-DUTY VEHICLE TEST CYCLE. *International Journal of Automotive Technology*, 21(5), 1233–1246. <https://doi.org/10.1007/s12239-020-0117-5>

- Lundby, K. M., & Johnson, J. W. (2006). *Relative Weights of Predictors What Is Important When Many Forces Are Operating*.
- Ministerio de Ambiente y Desarrollo Sostenible. (2017). *Guía para la elaboración de Inventarios de Emisiones Atmosféricas*.  
[http://www.minambiente.gov.co/images/AsuntosambientalesySectorialyUrbana/pdf/emisiones\\_atmosfericas\\_contaminantes/documentos\\_relacionados/GUIA\\_PARA\\_LA\\_ELABORACION\\_DE\\_INVENTARIOS\\_DE\\_EMISIONES\\_ATMOSFERICAS.pdf](http://www.minambiente.gov.co/images/AsuntosambientalesySectorialyUrbana/pdf/emisiones_atmosfericas_contaminantes/documentos_relacionados/GUIA_PARA_LA_ELABORACION_DE_INVENTARIOS_DE_EMISIONES_ATMOSFERICAS.pdf)
- Ministerio de ambiente y desarrollo sostenible. (2018). *Resolución 2254 de 2017*.  
[https://www.icbf.gov.co/cargues/avance/docs/resolucion\\_minambientes\\_2254\\_2017.htm](https://www.icbf.gov.co/cargues/avance/docs/resolucion_minambientes_2254_2017.htm)
- Ministro Ambiente Y Desarrollo Sostenible. (2022). *Resolución 0762 del 18 de Julio de 2022*.  
<https://www.minambiente.gov.co/wp-content/uploads/2022/09/Resolucion-762-de-2022.pdf>
- Moradi, E., & Miranda-Moreno, L. (2020). Vehicular fuel consumption estimation using real-world measures through cascaded machine learning modeling. *Transportation Research Part D: Transport and Environment*, 88. <https://doi.org/10.1016/j.trd.2020.102576>
- NEO-6 u-blox 6 GPS Modules Data Sheet. (2011). [www.u-blox.com](http://www.u-blox.com)
- Nguyen, Y.-L. T., Bui, N. D., Nghiem, T.-D., & Le, A.-T. (2020). GPS DATA PROCESSING FOR DRIVING CYCLE DEVELOPMENT IN HANOI, VIETNAM. In *Journal of Engineering Science and Technology* (Vol. 15, Issue 2).
- Nouri, P., & Morency, C. (2017). Evaluating microtrip definitions for developing driving cycles. *Transportation Research Record*, 2627, 86–92. <https://doi.org/10.3141/2627-10>
- Nyberg, P., Frisk, E., & Nielsen, L. (2014). Generation of equivalent driving cycles using Markov chains and mean tractive force components. In *IFAC Proceedings Volumes (IFAC-PapersOnline)* (Vol. 19, Issue 3). IFAC. <https://doi.org/10.3182/20140824-6-za-1003.02239>
- Observatorio Ambiental de Bogotá. (2022). *Cifras e Indicadores de Medio Ambiente en Bogotá Concentración de Material Particulado Inferior a 10 Micrómetros {PM10} Promedio Anual por Estación - PM10PAE*. <https://oab.ambientebogota.gov.co/indicadores/?id=24341000-a6e7-11eb-b6f5-a915ff441b6d>
- Observatorio ambiental de Bogotá. (2022). *Cifras e Indicadores de Medio Ambiente en Bogotá Material Particulado Inferior a 10 Micras {μ} Promedio Anual - PM10*.  
<https://oab.ambientebogota.gov.co/indicadores/?id=d2ccd170-0178-11ea-8cc7-8197075aabad>
- Onnegren, P. (2013). Driving cycle generation using statistical analysis and markov chains. *Department of Electrical Engineering, LiTH-ISY-EX-13/4670--SE*, 89.
- Pavlovic, J., Ciuffo, B., Fontaras, G., Valverde, V., & Marotta, A. (2018). How much difference in type-approval CO2 emissions from passenger cars in Europe can be expected from changing to the new test procedure (NEDC vs. WLTP)? *Transportation Research Part A: Policy and Practice*, 111(February), 136–147. <https://doi.org/10.1016/j.tra.2018.02.002>
- Peña, J. M., Lozano, J. A., & Larrañaga, P. (1999). *An empirical comparison of four initialization methods for the K-Means algorithm*. [www.elsevier.nl/locate/patrec](http://www.elsevier.nl/locate/patrec)

- Peng, Y., Zhuang, Y., & Yang, Y. (2020). A driving cycle construction methodology combining k-means clustering and Markov model for urban mixed roads. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 234(2–3), 714–724. <https://doi.org/10.1177/0954407019848873>
- Qiu, D., Li, Y., & Qiao, D. (2018). Recurrent Neural Network Based Driving Cycle Development for Light Duty Vehicles in Beijing. *Transportation Research Procedia*, 34, 147–154. <https://doi.org/10.1016/j.trpro.2018.11.026>
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Roy, F., & Morency, C. (2020). Comparing Driving Cycle Development Methods Based on Markov Chains. *Transportation Research Record*, 2675(3), 212–221. <https://doi.org/10.1177/0361198120968829>
- Secretaría de Ambiente de Bogotá. (2022). *Inventario de emisiones contaminantes atmosféricos 2020*. <https://drive.google.com/file/d/1a8gyjx0h0OPa6Apx8J9h-9lgfiKN-Df-/view>
- Selesnick, I. (2022). *Least Squares with Examples in Signal Processing*. [http://eeweb.poly.edu/iselesni/lecture\\_notes/](http://eeweb.poly.edu/iselesni/lecture_notes/)
- Shi, S., Lin, N., Zhang, Y., Cheng, J., Huang, C., Liu, L., & Lu, B. (2016). Research on Markov property analysis of driving cycles and its application. *Transportation Research Part D: Transport and Environment*, 47, 171–181. <https://doi.org/10.1016/j.trd.2016.05.013>
- Society of Automotive Engineers. (2007). *SAE J1979*.
- Society of Automotive Engineers. (2016). SAE J1962 - Diagnostic Connector. In *SAE Standards*.
- Society of Automotive Engineers (SAE). (2010). SAE J1263 - Road load measurement and dynamometer simulation using coastdown techniques. In *SAE Technical Papers*. <https://doi.org/10.4271/810828>
- Society of Automotive Engineers (SAE). (2011). SAE J254 - Instrumentation and techniques for exhaust gas emissions measurement. In *SAE Standards*. [https://www.sae.org/standards/content/j254\\_201106/](https://www.sae.org/standards/content/j254_201106/)
- Society of Automotive Engineers (SAE). (2014). SAE J2951 - Drive Quality Evaluation for Chassis Dynamometer Testing: Surface Vehicle Recommended Practice. In *SURFACE VEHICLE RECOMMENDED PRACTICE: Vol. SAE J2951*. [https://www.sae.org/standards/content/j2951\\_201401/](https://www.sae.org/standards/content/j2951_201401/)
- The Mathworks Inc. (2022a). *MATLAB* (2022a). The Mathworks Inc.
- The Mathworks Inc. (2022b). *Statistics and Machine Learning Toolbox* (2022a). Mathworks Inc.
- The National Renewable Energy Laboratory. (n.d.). *DriveCAT: Drive Cycle Analysis Tool*. DriveCAT: Drive Cycle Analysis Tool. Retrieved October 12, 2022, from <https://www.nrel.gov/transportation/drive-cycle-tool/>
- The National Renewable Energy Laboratory (NREL). (2015). *Transportation Secure Data Center - NREL*. Drive Cycle Data. [www.nrel.gov/tsdc](http://www.nrel.gov/tsdc).

- Toledo, T., Musicant, O., & Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. *Transportation Research Part C: Emerging Technologies*, 16(3), 320–331. <https://doi.org/10.1016/j.trc.2008.01.001>
- Tutuianu, M., Bonnel, P., Ciuffo, B., Haniu, T., Ichikawa, N., Marotta, A., Pavlovic, J., & Steven, H. (2015). Development of the World-wide harmonized Light duty Test Cycle (WLTC) and a possible pathway for its introduction in the European legislation. *Transportation Research Part D: Transport and Environment*, 40, 61–75. <https://doi.org/10.1016/j.trd.2015.07.011>
- USEPA AP 42. (1995). AP 42, Fifth Edition Compilation of Air Pollutant Emission Factors, Volume 1: Stationary and Point Sources. *AP 42, Fifth Edition Compilation of Air Pollutant Emission Factors, Volume 1: Stationary and Point Sources*, 1–10. <https://www3.epa.gov/ttnchie1/ap42/c00s00.pdf%0Ahttps://www3.epa.gov/ttn/chie/ap42/c00s00.pdf>
- World Health Organization. (2018). *Ambient (outdoor) air pollution*. [https://www.who.int/en/news-room/fact-sheets/detail/ambient-\(outdoor\)-air-quality-and-health](https://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health)
- World Health Organization. (2021a). *WHO global air quality guidelines*.
- World Health Organization. (2021b). *WHO Global Air Quality Guidelines*.
- Wu, Y., Zhang, W., Zhang, L., Qiao, Y., Yang, J., & Cheng, C. (2020). A multi-clustering algorithm to solve driving cycle prediction problems based on unbalanced data sets: A Chinese case study. *Sensors (Switzerland)*, 20(9). <https://doi.org/10.3390/s20092448>
- Yang, Y., Li, T., Hu, H., Zhang, T., Cai, X., Chen, S., & Qiao, F. (2019). Development and emissions performance analysis of local driving cycle for small-sized passenger cars in Nanjing, China. *Atmospheric Pollution Research*, 10(5), 1514–1523. <https://doi.org/10.1016/j.apr.2019.04.009>
- Yuan, M., Kan, X., Chi, C. H., Cao, L., Shu, H., Fan, Y., & Yao, W. (2021). Study of Driving Cycle of City Tour Bus Based on Coupled GA-K-Means and HMM Algorithms: A Case Study in Beijing. *IEEE Access*, 9, 20331–20345. <https://doi.org/10.1109/ACCESS.2021.3054118>
- Yuhui, P., Yuan, Z., & Huibao, Y. (2019). Development of a representative driving cycle for urban buses based on the K-means cluster method. *Cluster Computing*, 22(s3), 6871–6880. <https://doi.org/10.1007/s10586-017-1673-y>
- Zhang, C., Kotz, A., Kelly, K., & Rippelmeyer, L. (2021). Development of heavy-duty vehicle representative driving cycles via decision tree regression. *Transportation Research Part D: Transport and Environment*, 95(May), 102843. <https://doi.org/10.1016/j.trd.2021.102843>

## Anexos

### A. Desarrollo del software del dispositivo IVDR

En este apéndice se realiza un análisis detallado de cada uno de los algoritmos utilizados en las rutinas presentadas en la sección 4.1.1, específicamente en la máquina de estados presentada en la Figura 4.6, es importante mencionar que estos algoritmos son muy parecidos a los desarrollados en trabajos previos (Echeverry Mejía, 2018; Echeverry-Mejía et al., 2022). Sin embargo, presentan diferentes modificaciones. Es importante tener en consideración:

- Los comandos necesarios para la inicialización del microcontrolador y periféricos (Módulo GPS, Módulo de almacenamiento, y LEDs indicadores), se encuentran descritos en el programa principal del IVDR. De igual manera, este programa es el encargado de enviar a través del canal UART asignado al módulo de almacenamiento la información del buffer con el cual se encuentre en el momento.
- Las variables de estado se resaltan en negrilla.

#### A.1. Desarrollo del programa principal

El programa principal del dispositivo IVDR, contiene principalmente la inicialización de los comandos para el correcto funcionamiento del microcontrolador y periféricos, así como la declaración de variables necesarias para el funcionamiento del IVDR. El microcontrolador y los periféricos son inicializados de la siguiente manera.

- MCU: Reloj del microcontrolador inicializado a 48 MHz
- GPIO: Inicialización de los LEDs indicadores
- PIT: Configuración de interrupciones
  - PIT0: Interrupción periódica, esta se activa cuando la variable **velobd2msm**, es diferente del valor máximo de transferencia de datos, permitiendo al usuario leer datos cada 10 mseg, 100 mseg, o 1 seg.
- UART: Establece la configuración de cada uno de los puertos seriales:
  - UART0: Puerto serial asignado al módulo GPS, configurado a 9600 baudios, activando su interrupción de recepción.

- UART1: Puerto serial asignado al módulo de almacenamiento, configurado a 115200 baudios.
- CAN: Establece la configuración del protocolo CAN, realizando su activación a una velocidad de 0.5Mbps, velocidad la cual puede ser modificada acorde a cada vehículo, debido que algunos vehículos manejan velocidades de transferencia de datos diferentes. De igual manera se configura para que el timestamp sea establecido a un timestep de 2μseg. Se activa la respectiva interrupción, se envían las variables acorde a una lista establecida, y se espera hasta que se tenga un mensaje de OBD2.
- IRQ: Se activan cada una de las solicitudes de interrupción par la interrupción PIT, los puertos seriales y el protocolo de comunicación.

Los principales periféricos inicializados en el programa principal, utilizando el comando GPIO, consiste en establecer las salidas de los pines correspondientes a los LEDs indicadores. Los periféricos Módulo GPS y módulo de almacenamiento se activan una vez el dispositivo esta encendido, a enviando información y a la espera de recibir información respectivamente.

Se inicializan las variables **controlGPS**, y los respectivos buffers NMEA\_GPGGA, y CANbuff, y los contadores para la organización y envío de datos al módulo SD, i\_canbuff (tamaño del mensaje de OBD2), i\_canbuff\_tx (índice para el buffer CAN, mensaje de OBD2), err\_overbuff (Variable para validación del mensaje OBD2), y k\_gps\_tx (índice para el buffer GPS).

Una vez inicializados el microcontrolador, los periféricos y establecidas las variables para el funcionamiento de cada uno de los módulos, el programa entra en in ciclo infinito, donde verifica si **controlGPS** es igual a uno, lo cual indica que hay un mensaje del módulo GPS a la espera, se organiza este mensaje acorde a la estructura del mismo la cual es analizada en la sección 4.2.2 y se envía al módulo de almacenamiento, si **controlGPS** es igual a 0, indica que no hay un mensaje del módulo GPS, así que se revisa si se tiene un mensaje de CAN a la espera de ser almacenado, de ser así, se organiza el mensaje en el buffer asignado a CAN y se envía al módulo de almacenamiento. El programa solo sale del ciclo infinito hasta que se apague el dispositivo.

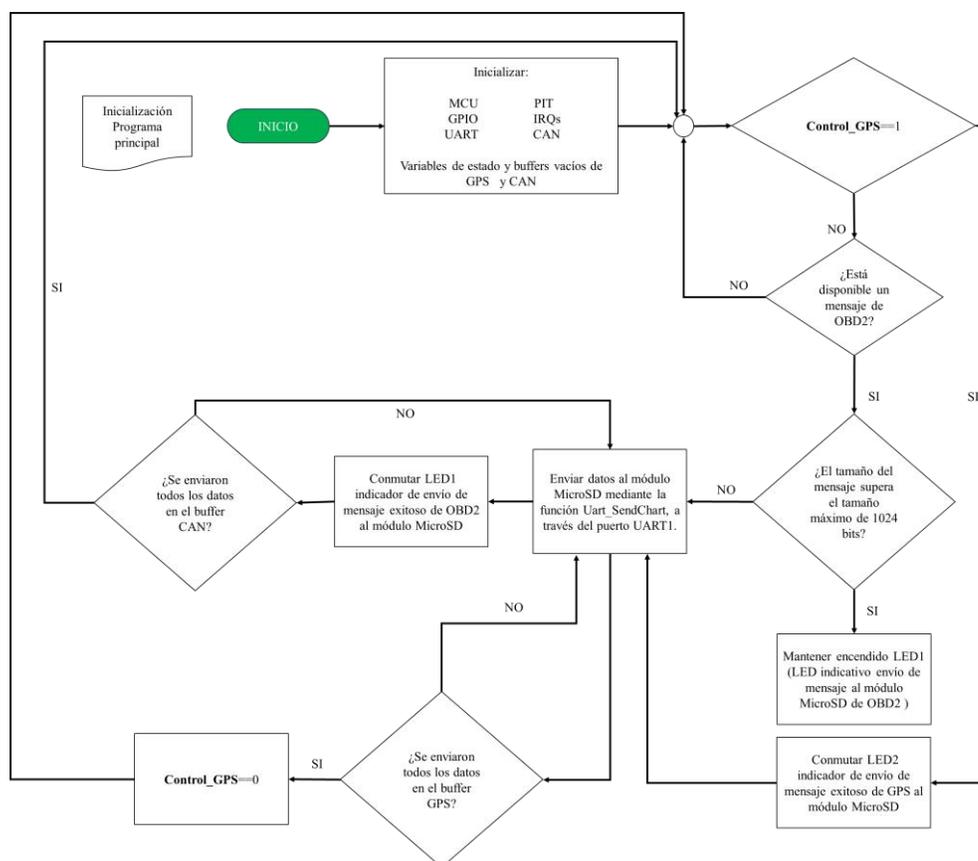


Figura A. 1. Diagrama de flujo de la rutina del programa principal

## A.2 Interrupción RX UART

La interrupción UART, es asignada principalmente a la lectura de los datos del módulo GPS, mediante puerto UART0 del microcontrolador, el puerto UART1 es asignado al módulo de almacenamiento, el cual no requiere interrupción debido a que el programa de inicialización UART, inicializa cada uno de los puertos, así como una función llamada Uart\_SendChar, la cual es llamada en la función principal, y envía a través del puerto UART1 conectado al módulo de almacenamiento los datos, carácter por carácter, así mismo establece el puerto UART0 con interrupción.

De esta manera el puerto UART0 es el único que tiene una interrupción con el fin de leer los datos del módulo GPS NEO-6M UBLOX, el cual solamente requiere el canal de recepción, y no de transmisión, ya que solamente se leen los mensajes de este módulo. El módulo GPS NEO-6M UBLOX, una vez está encendido y valida su posición geográfica con los satélites empieza a enviar tramas de datos acorde a la codificación de datos del sistema NMEA (*National Marine Electronics*

Association, por sus siglas en inglés) (NEO-6 u-Blox 6 GPS Modules Data Sheet, 2011). La trama de datos de es la que empieza con las serie de datos “GPGGA” (Global Positioning system), cuyo primer carácter es “\$”, entregando los datos de: tiempo, longitud, latitud, y altitud. Se inicializan los vectores o buffers NMEA\_mss para el análisis del mensaje recibido, y NMEA\_GPGGA para almacenar el mensaje a enviar, así como el índice `i_gpsbuff`, y la variable de control **controlGPS**. En la Figura A. 2 se presenta el funcionamiento general de esta rutina.

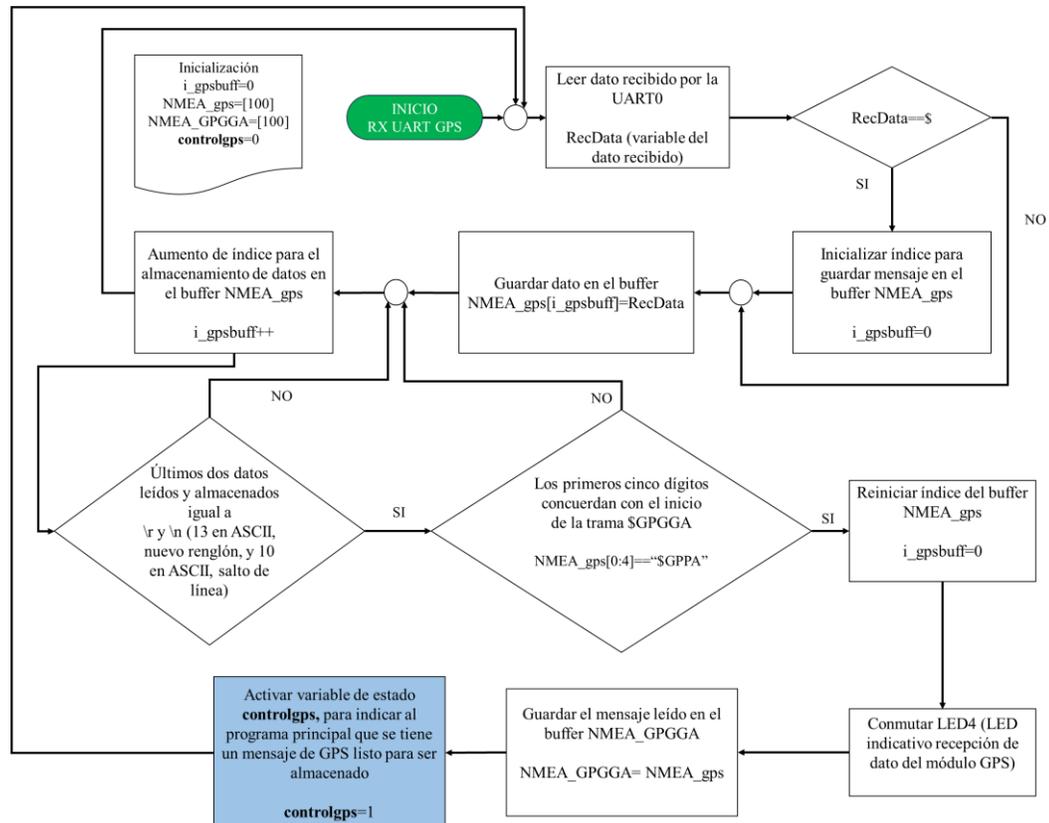


Figura A. 2. Diagrama de flujo de la rutina RX UART0

### A.3. Interrupción RX CAN

Como se mencionó en la sección 3.2.2, en trabajos previos el dispositivo IVDR fue desarrollado con el fin de poder leer diferentes tipos de datos dentro del mismo protocolo CAN, basado en los diferentes modos, como también se mencionó para el desarrollo del proyecto solamente es necesario uno de los modos y este es el modo OBDII, el cual permite la solicitud y lectura de parámetros específicos del vehículo mediante el protocolo de diagnóstico OBDII.

Una de las principales funciones de esta rutina de interrupción, corresponde a la lectura de datos transmitidos con base en el protocolo CAN, mediante el protocolo de diagnóstico OBDII, para que estos datos sean almacenados en el buffer asignado al mensaje de OBD2, y posteriormente enviados al respectivo periférico en este caso el módulo de almacenamiento. Esta interrupción se activa con alta periodicidad, debido a la alta velocidad de transmisión de datos manejada por el protocolo CAN.

Una vez, se activa la interrupción y se cuenta con el mensaje de OBDII, se procede a realizar los siguientes pasos

- Lectura de los datos más relevantes, ID del mensajes, trama de datos del mensaje y timesatamp, este último es un dato de suma importancia ya que permite localizar en el tiempo en el mensaje almacenado, la utilización de este dato se presenta en la sección 5.1.3.
- Teniendo el mensaje de OBDII, este se almacena en el buffer CANbuff\_tx, para ser enviado al módulo de almacenamiento, basado en los trabajos previos este mensaje es almacenado en formato binario y no en ASCII, con el objetivo de ahorrar tiempo de procesamiento.

Dentro de la misma interrupción se encuentra el procedimiento para la solicitud de mensajes al protocolo de comunicación OBII (ID CAN 0x7E8 (2024), para lectura de datos de OBDII, ID CAN 0x7DF (2015), para la solicitud de datos de OBDII), esto mediante la escritura del mensaje en el puerto CAN para solicitud de dato, similar al presentado en la Tabla 3.6. Estos mensajes deben ser solicitados antes de leer un mensaje y almacenar un mensaje. Cada uno de los mensajes es solicitado mediante el ID 0x7DF, y así mismo los PIDs, necesarios son solicitados mediante un vector de PIDs, el cual se está recorriendo continuamente durante el desarrollo del programa, así mismo la rutina asegura que se solicitará un mensaje una vez el mensaje que haya solicitado con anterioridad, haya sido leído y almacenado correctamente . En caso tal que el mensaje de solicitud de datos no haya sido enviado correctamente se enciende un LED indicando un error y el programa para.

Para inicializar esta rutina, es necesario solicitar un mensaje inicial, este mensaje inicial se solicita mediante la rutina principal en caso tal de que se haya solicitado la velocidad máxima de transferencia de datos, en el caso contrario se realiza mediante la interrupción PITO, donde se

solicita el primer PID del vector de parámetros requeridos. El diagrama general de esta rutina se presenta a continuación.

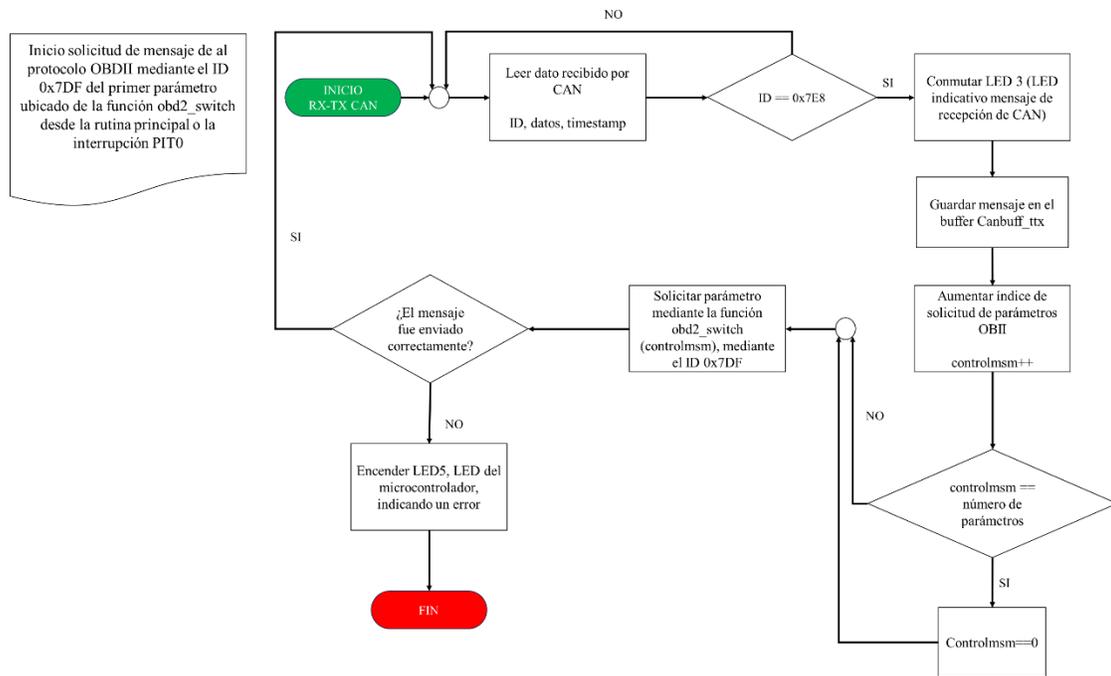


Figura A. 3. Diagrama de flujo de la rutina RX CAN

#### A.4. Interrupción PIT

Para el correcto funcionamiento del dispositivo IVDR, es necesario establecer una interrupción PIT, mediante el puerto establecido para esta interrupción, el cual es el puerto PIT0, con el objetivo de sincronizar los mensajes de solicitud de OBDII, esta interrupción cobra una gran relevancia al establecer una velocidad diferente a la velocidad máxima permitida por el dispositivo. Permitiendo la inicialización de la solicitud de mensajes de OBDII. Esta interrupción debe ser activada acorde a la velocidad de solicitud de datos de OBDII.

La función principal consiste en la inicialización de la solicitud de mensajes de OBDII, mediante la función obd2\_switch analizada en la **¡Error! No se encuentra el origen de la referencia.**, para la solicitud de datos. El diagrama de flujo de esta rutina se presenta a continuación.

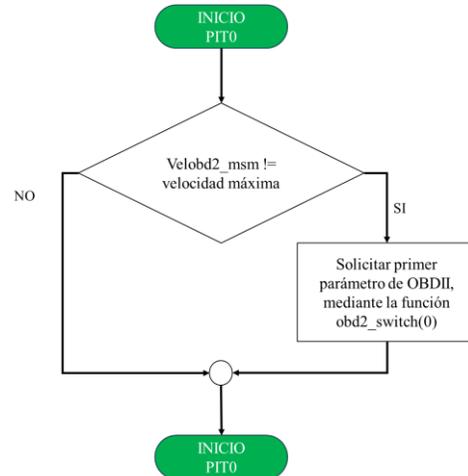


Figura A. 4. Diagrama de flujo de la rutina PIT0

### B. Distribución de velocidades para todas las metaheurísticas sin refinamiento de parámetros.

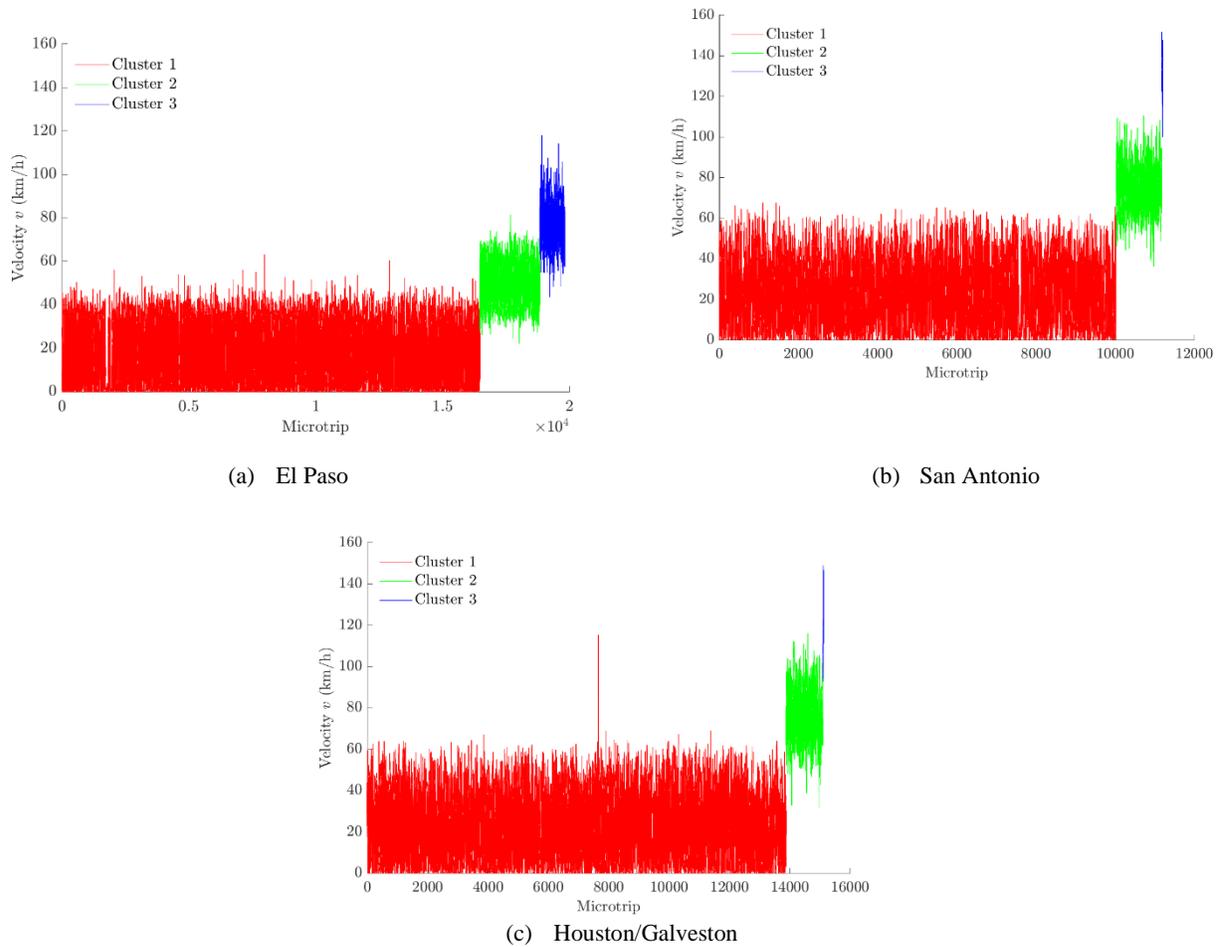
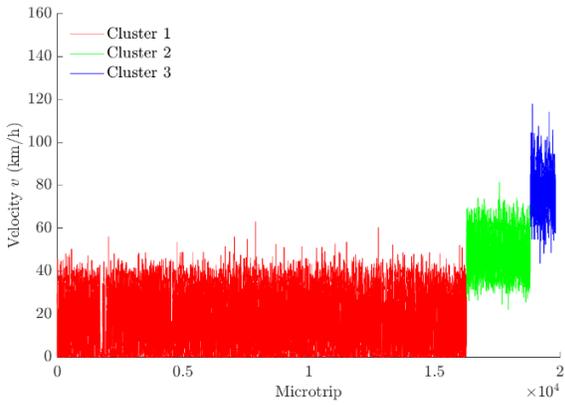
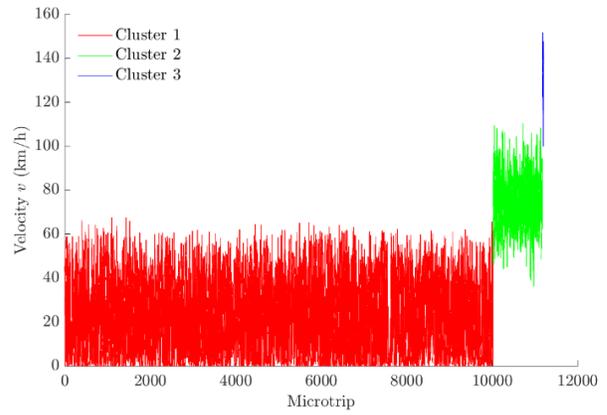


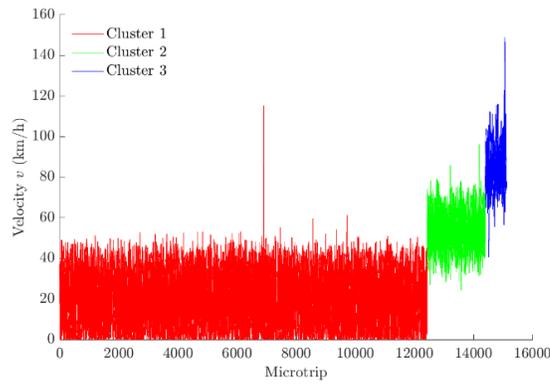
Figura A. 5 Distribución de los microciclos acorde a la velocidad media para la metaheurística GRASP-RANDOM



(a) El Paso

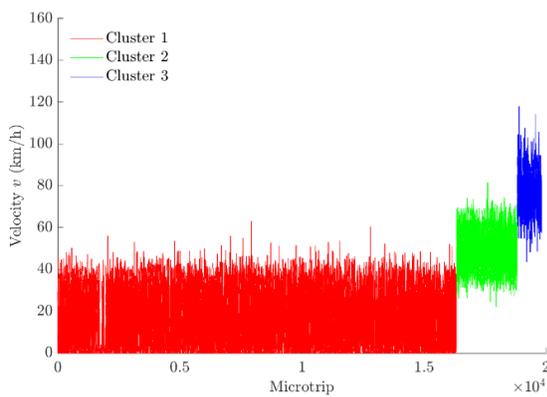


(b) San Antonio

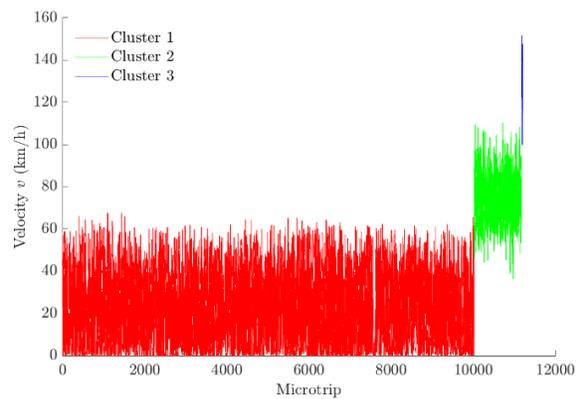


(c) Houston /Galveston

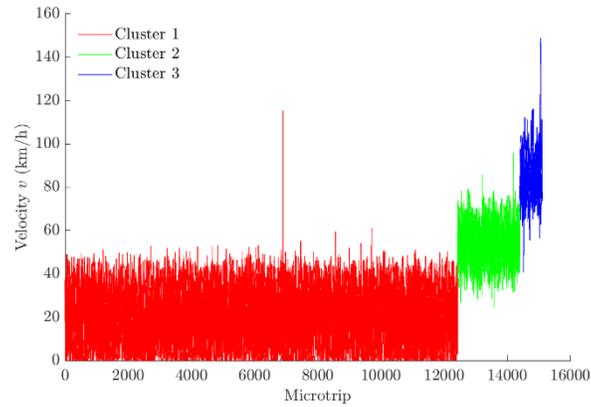
Figura A. 6. Distribución de los microciclos acorde a la velocidad media para la metaheurística GRASP-FORGY



(a) El Paso

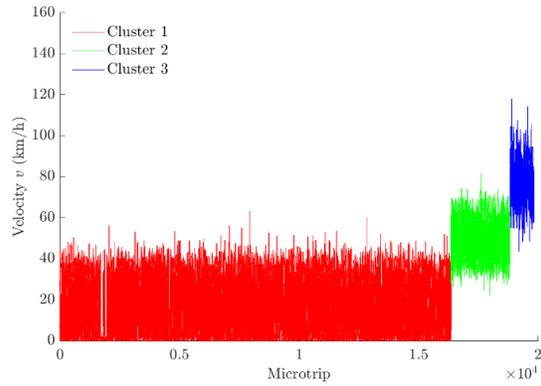


(b) San Antonio

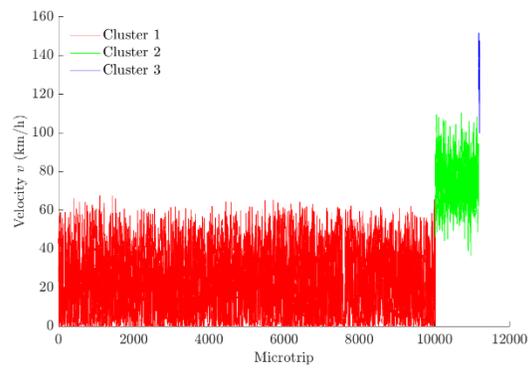


(c) Houston/Galveston

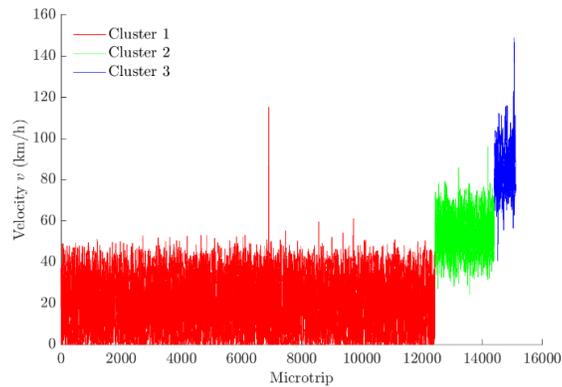
Figura A. 7. Distribución de los microciclos acorde a la velocidad media para la metaheurística ILS-FORGY



(a) El Paso



(b) San Antonio



(c) Houston/Galveston

Figura A. 8. Distribución de los microciclos acorde a la velocidad media para la metaheurística GKA

### C. Resultados de la evaluación comparativa de algoritmos con el refinamiento de parámetros

La variabilidad es expresada en términos de la desviación estándar, durante las 30 corridas . Cuando los resultados de las 30 corridas llegan al mismo resultado la variabilidad se expresa como cero.

Tabla A. 1. *Resultados generales para cada algoritmo con el refinamiento de parámetros.*

Algoritmo	Promedio de la función de costo	Promedio de iteraciones
<b>EL PASO</b>		
GRASP - RANDOM	$1.76774 \pm 1.6 \times 10^{-3}$	$179 \pm 61$
GRASP - FORGY	$1.76702 \pm 7.7 \times 10^{-4}$	$169 \pm 46$
ILS - RANDOM	$1.77242 \pm 2.0 \times 10^{-5}$	$129 \pm 25$
ILS - FORGY	$1.76639 \pm 5.0 \times 10^{-5}$	$182 \pm 61$
GENETIC K-MEANS	$1.76638 \pm 3.6 \times 10^{-5}$	$11852 \pm 2214$
<b>SAN ANTONIO</b>		
GRASP - RANDOM	$2.34218 \pm 2.4 \times 10^{-5}$	$151 \pm 43$
GRASP - FORGY	$2.34218 \pm 0.0$	$102 \pm 1$
ILS - RANDOM	$2.34213 \pm 0.0$	$122 \pm 16$
ILS - FORGY	$2.34213 \pm 1.3 \times 10^{-5}$	$162 \pm 48$
GENETIC K-MEANS	$2.34217 \pm 1.9 \times 10^{-5}$	$15511 \pm 3613$
<b>HOUSTON</b>		
GRASP - RANDOM	$2.17140 \pm 3.1 \times 10^{-4}$	$135 \pm 42$
GRASP - FORGY	$2.17081 \pm 3.4 \times 10^{-4}$	$174 \pm 57$
ILS - RANDOM	$2.17121 \pm 0.0$	$106 \pm 4$
ILS - FORGY	$2.17042 \pm 3.9 \times 10^{-5}$	$161 \pm 55$
GENETIC K-MEANS	$2.17045 \pm 1.5 \times 10^{-4}$	$16010 \pm 3346$

Tabla A. 2. *Distribución de porcentajes de microciclos para todas las bases de datos.*

Algoritmo	Cluster 1 (%)	Cluster 2 (%)	Cluster 3 (%)
<b>EL PASO</b>			
GRASP - RANDOM	$78.36 \pm 1.61$	$15.99 \pm 1.40$	$5.64 \pm 0.21$
GRASP - FORGY	$78.11 \pm 1.03$	$16.18 \pm 0.90$	$5.70 \pm 0.12$
ILS - RANDOM	$81.90 \pm 5.50 \times 10^{-3}$	$12.93 \pm 4.60 \times 10^{-3}$	$5.16 \pm 9.21 \times 10^{-4}$
ILS - FORGY	$78.32 \pm 0.11$	$15.98 \pm 0.09$	$5.69 \pm 0.01$
GENETIC K-MEANS	$78.63 \pm 0.09$	$16.00 \pm 0.08$	$5.69 \pm 0.01$
<b>SAN ANTONIO</b>			
GRASP - RANDOM	$89.42 \pm 0.02$	$10.33 \pm 1.40$	$0.24 \pm 0.0$
GRASP - FORGY	$89.41 \pm 0.0$	$10.35 \pm 0.0$	$0.24 \pm 0.0$
ILS - RANDOM	$89.43 \pm 0.0$	$10.33 \pm 0.0$	$0.24 \pm 0.0$
ILS - FORGY	$89.43 \pm 0.01$	$10.32 \pm 0.01$	$0.24 \pm 0.0$
GENETIC K-MEANS	$89.42 \pm 0.01$	$10.34 \pm 0.01$	$0.24 \pm 0.0$
<b>HOUSTON</b>			
GRASP - RANDOM	$83.31 \pm 0.10$	$12.08 \pm 0.08$	$4.60 \pm 0.01$
GRASP - FORGY	$82.85 \pm 0.31$	$12.46 \pm 0.25$	$4.68 \pm 0.06$
ILS - RANDOM	$83.27 \pm 0.0$	$12.12 \pm 0.0$	$4.61 \pm 0.0$
ILS - FORGY	$82.89 \pm 0.04$	$12.45 \pm 0.04$	$4.66 \pm 0.01$
GENETIC K-MEANS	$82.86 \pm 0.10$	$12.47 \pm 0.08$	$4.66 \pm 0.02$

Tabla A. 3. *Promedio de la velocidad media en cada cluster para cada base de datos.*

Algoritmo	Cluster 1 (km/h)	Cluster 2 (km/h)	Cluster 3 (km/h)
<b>EL PASO</b>			
GRASP - RANDOM	$15.51 \pm 0.45$	$45.58 \pm 1.10$	$75.72 \pm 0.47$
GRASP - FORGY	$15.44 \pm 0.27$	$45.34 \pm 0.66$	$75.57 \pm 0.27$
ILS - RANDOM	$16.50 \pm 1.8 \times 10^{-3}$	$48.16 \pm 3.8 \times 10^{-3}$	$76.84 \pm 1.9 \times 10^{-3}$
ILS - FORGY	$15.50 \pm 0.03$	$45.44 \pm 0.07$	$75.61 \pm 0.03$
GENETIC K-MEANS	$15.49 \pm 0.02$	$45.43 \pm 0.06$	$75.61 \pm 0.03$
<b>SAN ANTONIO</b>			
GRASP - RANDOM	$23.61 \pm 0.01$	$74.12 \pm 0.04$	$135.18 \pm 0.0$
GRASP - FORGY	$23.60 \pm 0.0$	$74.10 \pm 0.0$	$135.18 \pm 0.0$
ILS - RANDOM	$23.61 \pm 0.0$	$74.12 \pm 0.0$	$135.18 \pm 0.0$
ILS - FORGY	$23.61 \pm 4.9 \times 10^{-3}$	$74.13 \pm 0.02$	$135.18 \pm 0.0$
GENETIC K-MEANS	$23.60 \pm 4.5 \times 10^{-3}$	$74.11 \pm 0.02$	$135.18 \pm 0.0$
<b>HOUSTON</b>			
GRASP - RANDOM	$20.59 \pm 0.02$	$54.35 \pm 0.09$	$84.56 \pm 0.04$
GRASP - FORGY	$20.46 \pm 0.09$	$53.88 \pm 0.32$	$84.33 \pm 0.16$
ILS - RANDOM	$20.59 \pm 0.00$	$54.29 \pm 0.0$	$84.53 \pm 0.0$
ILS - FORGY	$20.47 \pm 0.01$	$53.94 \pm 0.04$	$84.40 \pm 0.03$
GENETIC K-MEANS	$20.47 \pm 0.03$	$53.91 \pm 0.10$	$84.39 \pm 0.06$

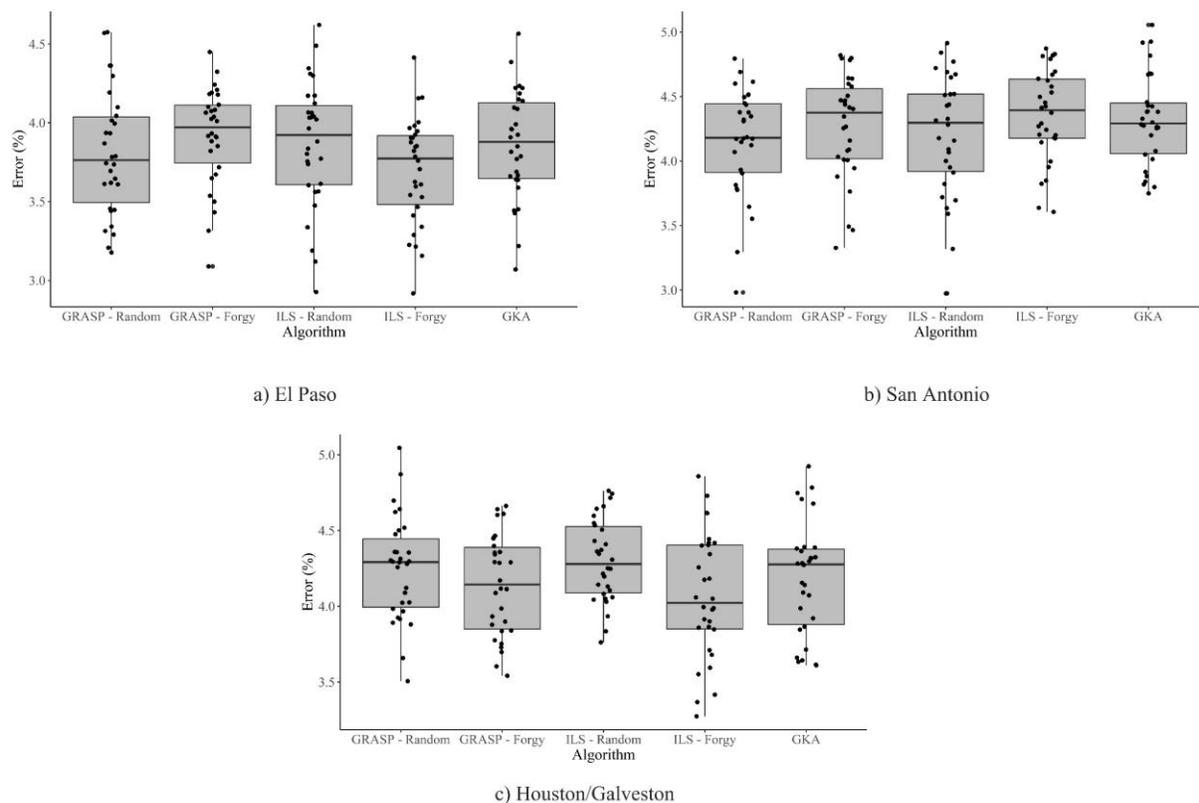
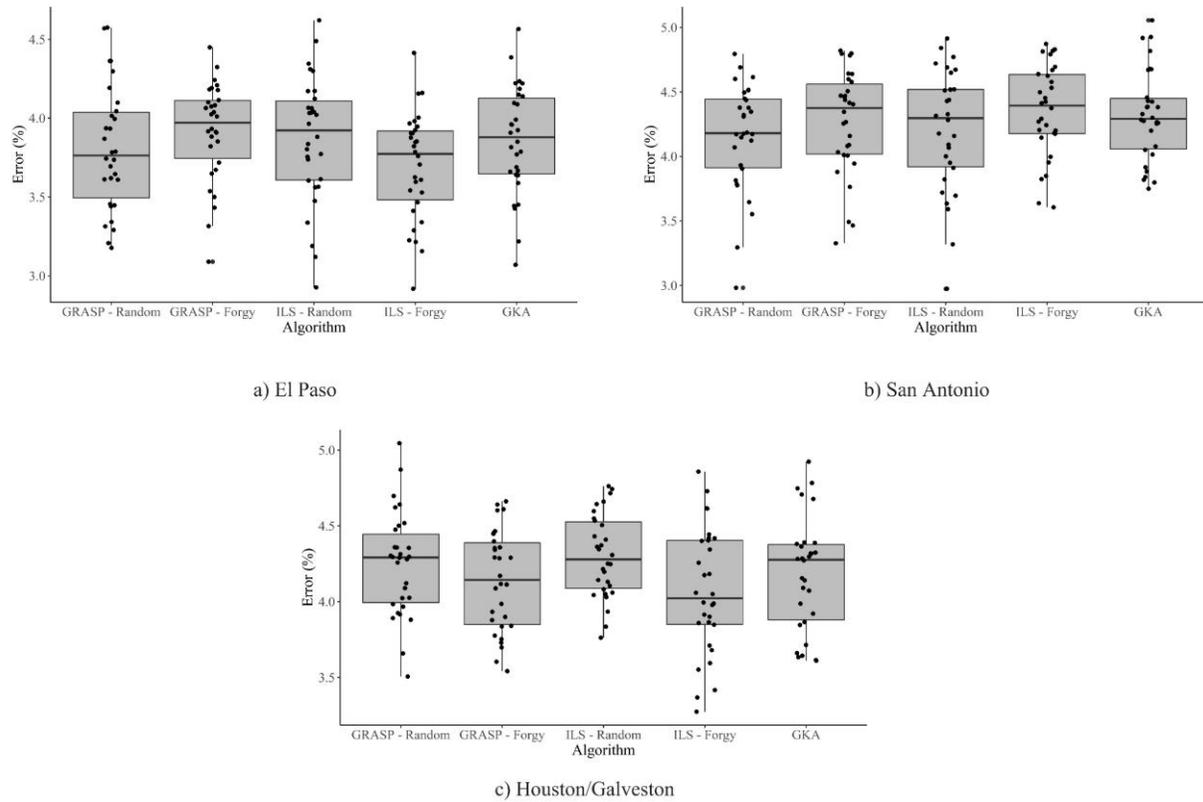
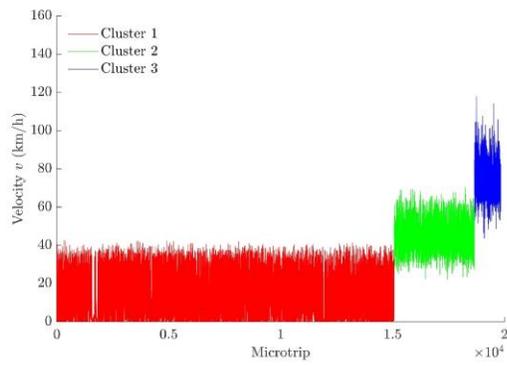


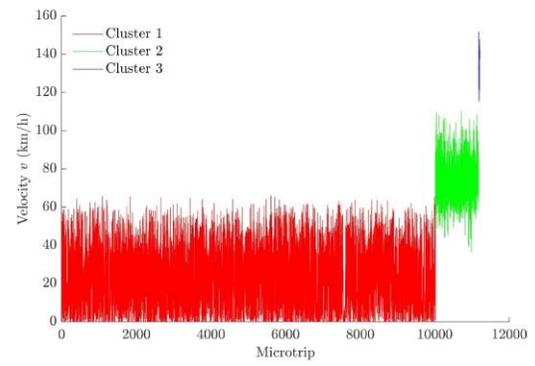
Figura A. 9. Distribución del error medio para todas las bases de datos con refinamiento de parámetros



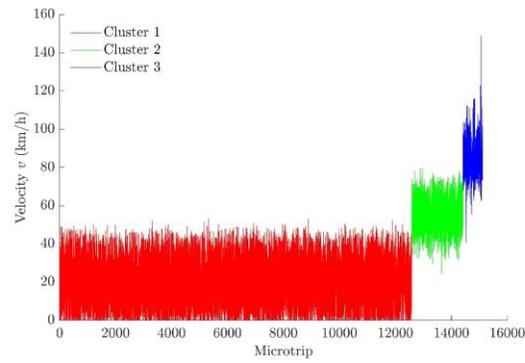
*Figura A. 10.* Distribución del error medio para todas las bases de datos y cada algoritmo con refinamiento de parámetros.



a) El Paso

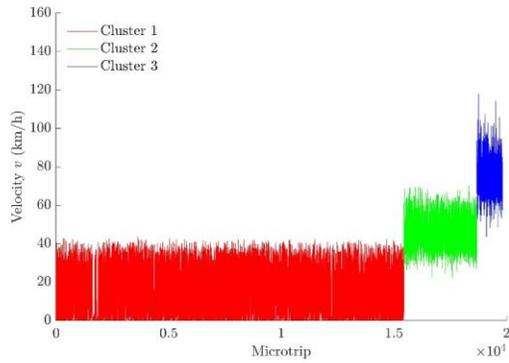


b) San Antonio

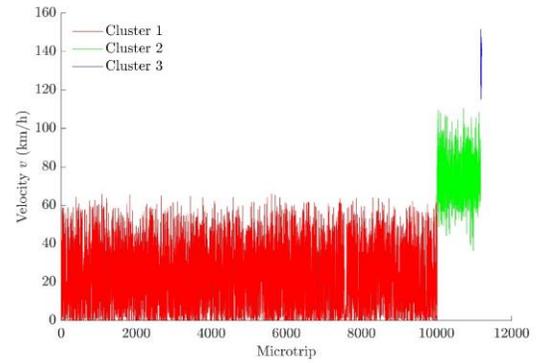


c) Houston/Galveston

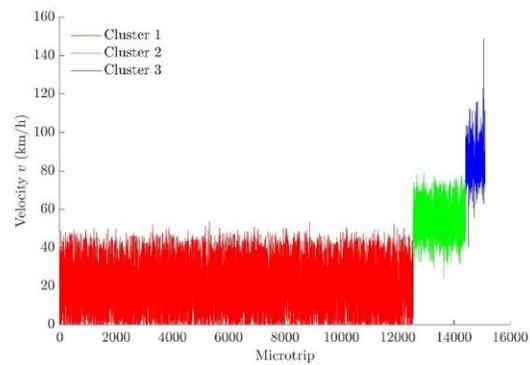
Figura A. 11. Distribución de microciclos en cada *cluster* acorde a la velocidad media para la metaheurística GRASP -RANDOM con refinamiento de parámetros



a) El Paso

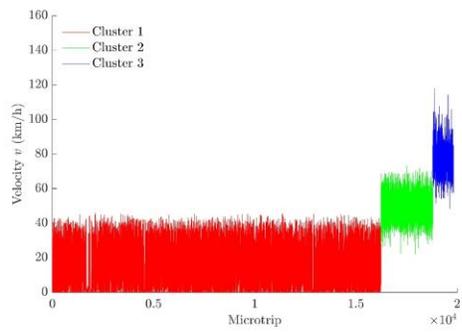


b) San Antonio

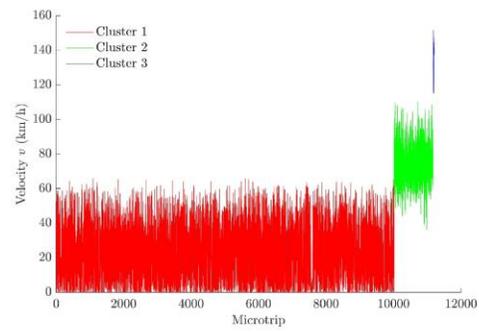


c) Houston/Galveston

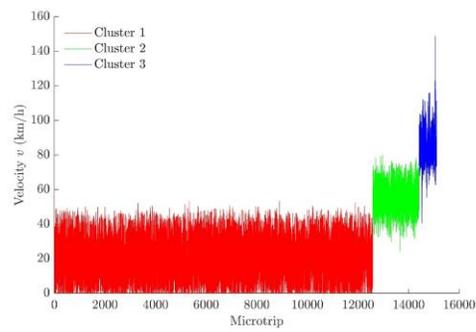
Figura A. 12. Distribución de microciclos en cada *cluster* acorde a la velocidad media para la metaheurística GRASP -FORGY con refinamiento de parámetros



a) El Paso

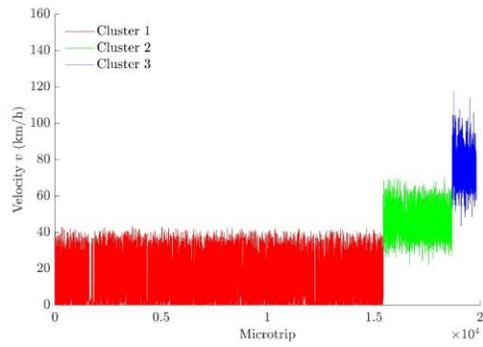


b) San Antonio

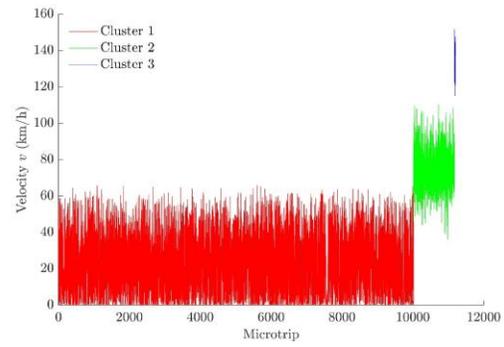


c) Houston/Galveston

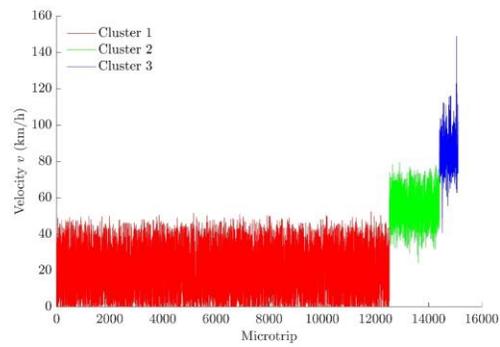
Figura A. 13. Distribución de microciclos en cada *cluster* acorde a la velocidad media para la metaheurística ILS-RANDOM con refinamiento de parámetros



a) El Paso

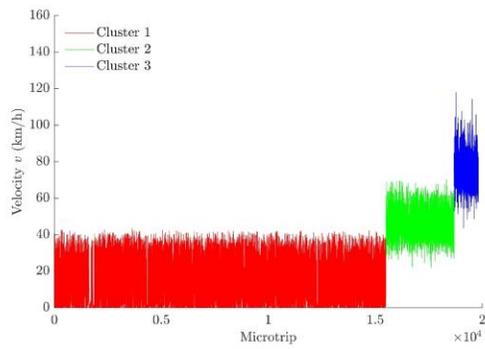


b) San Antonio

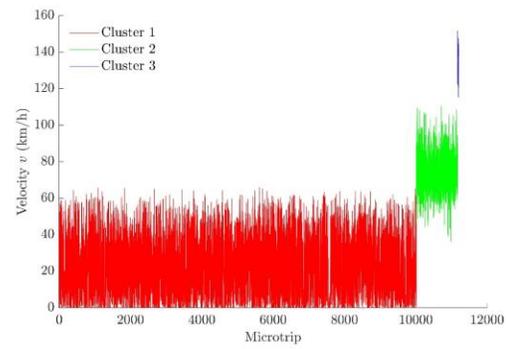


c) Houston/Galveston

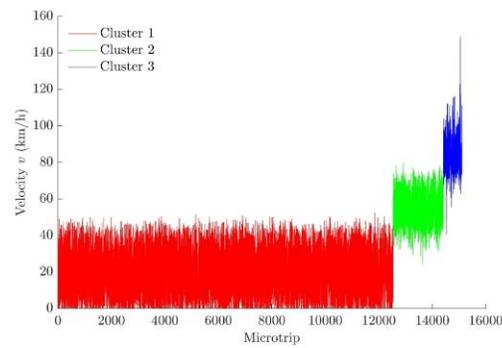
*Figura A. 14.* Distribución de microciclos en cada *cluster* acorde a la velocidad media para la metaheurística ILS-FORGY con refinamiento de parámetros



a) El Paso



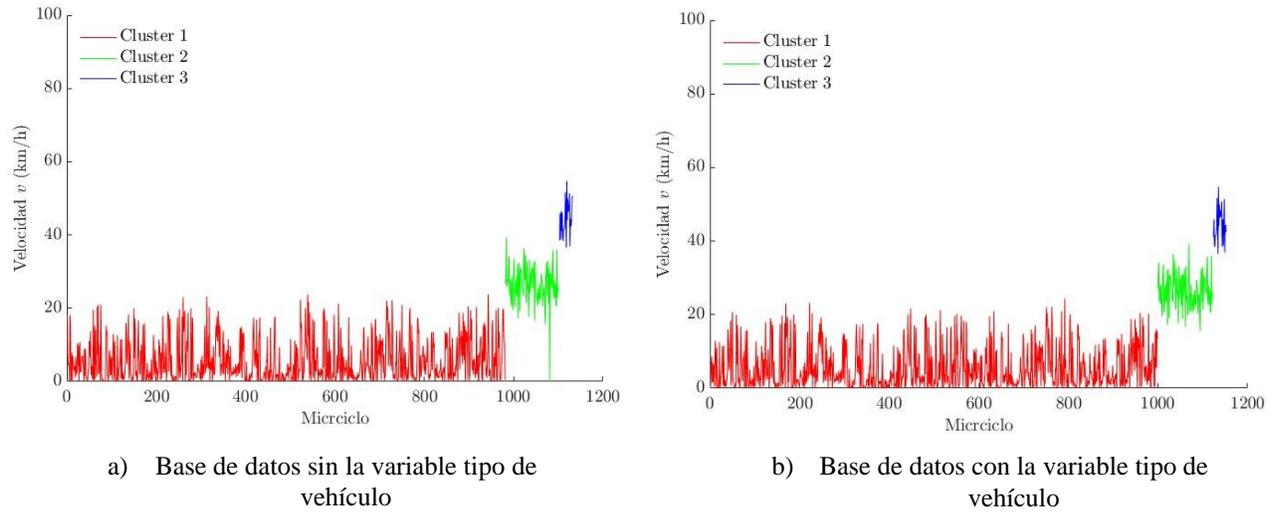
b) San Antonio



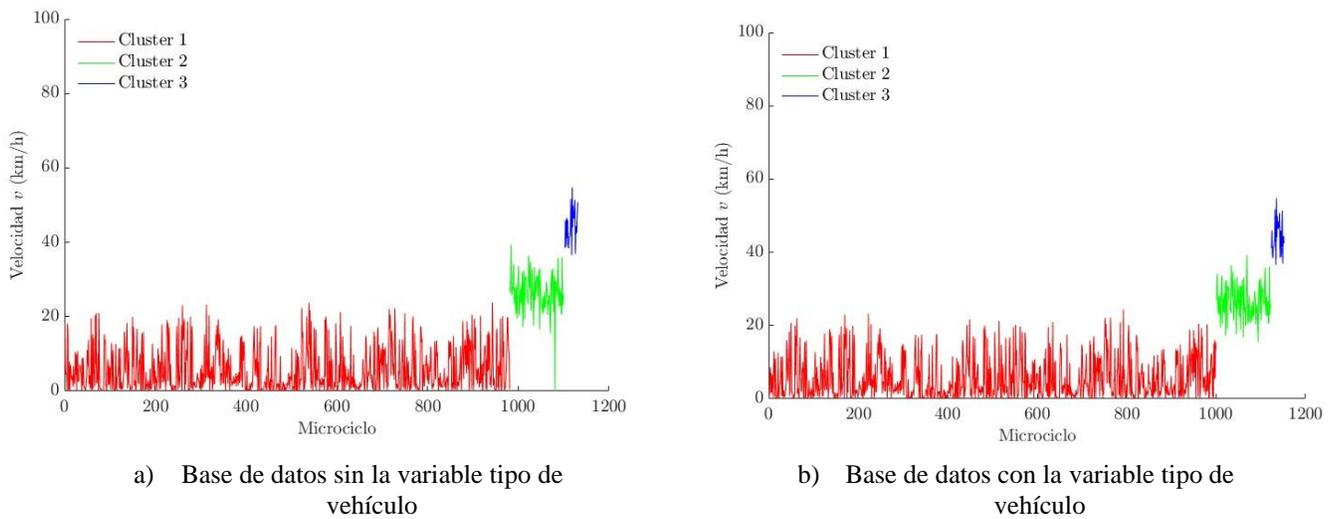
c) Houston/Galveston

Figura A. 15. Distribución de microciclos en cada *cluster* acorde a la velocidad media para la metaheurística GKA con refinamiento de parámetros

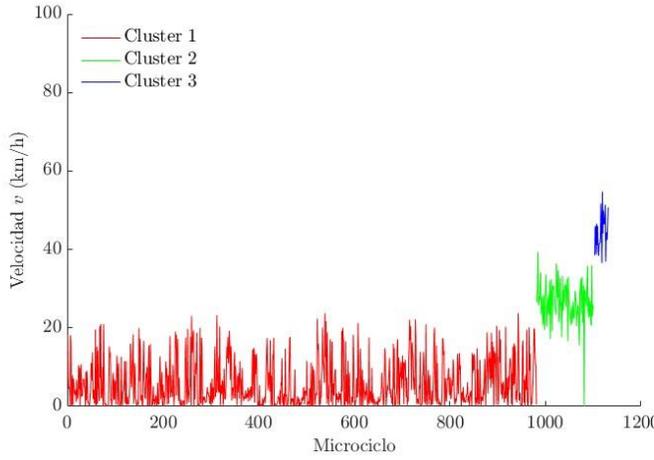
## D. Distribución de velocidades para todas las bases de datos de Bogotá región



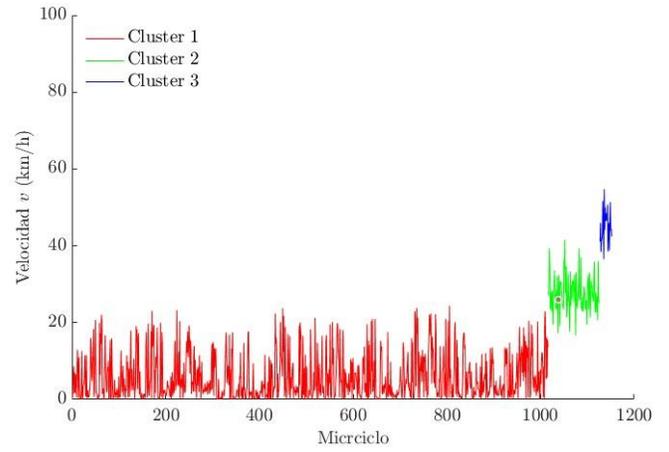
*Figura A. 16.* Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GRASP -FORGY



*Figura A. 17.* Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GRASP -FORGY

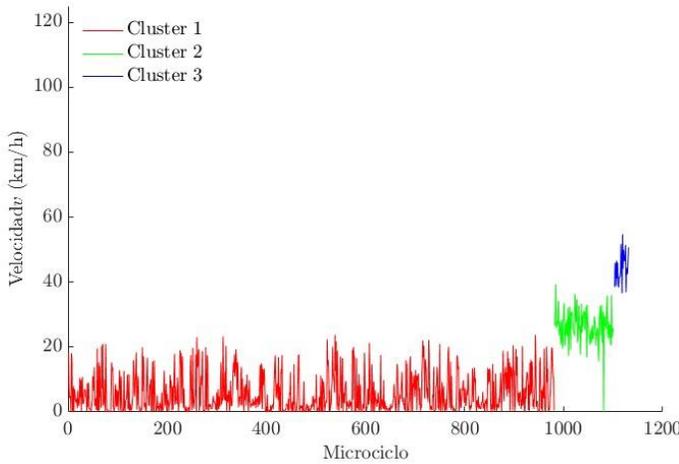


a) Base de datos sin la variable tipo de vehículo

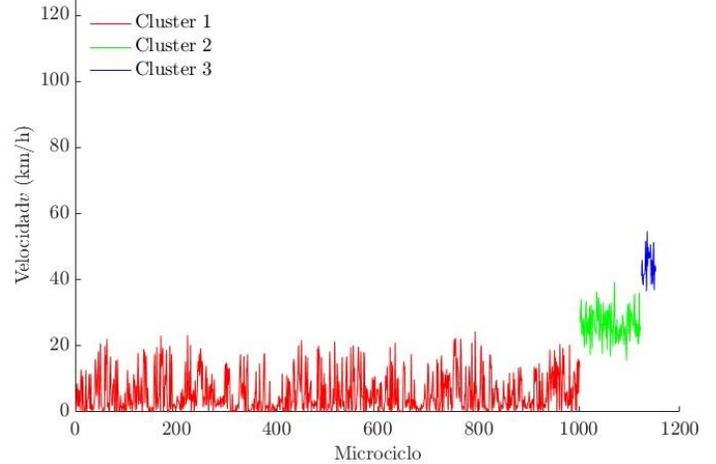


b) Base de datos con la variable tipo de vehículo

Figura A. 18. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística ILS -FORGY



a) Base de datos sin la variable tipo de vehículo



b) Base de datos con la variable tipo de vehículo

Figura A. 19. Distribución de microciclos en cada cluster acorde a la velocidad media para la metaheurística GKA