**Modelo de bajo costo para la estimación de emisiones contaminantes basado en GPS y aprendizaje automático**

**Néstor Rivera-Campoverde1,2, José Muñoz-Sanz 1, Blanca Arenas-Ramírez 3**

1División de Ingeniería de Maquinas, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, España. Email: joseluis.munozs@.upm.es

2 Grupo de investigación en Ingeniería del Transporte, Universidad Politécnica Salesiana, Ecuador. Email:

3Instituto Universitario de Investigación del Automóvil Francisco Aparicio, Universidad Politécnica de Madrid, España. Email: blanca.arenas@upm.es

**Resumen**

El presente trabajo presenta un método novedoso para la estimación de los contaminantes emitidos por vehículos propulsados con motores de combustión interna en conducciones reales de conducción, sin la necesidad de campañas extensas de medición ni el uso de instrumentación en el vehículo por períodos prolongados de tiempo; para lo cual se basa en las señales de posicionamiento y velocidad generadas por el GPS (Global Positioning System) y la aplicación de aprendizaje automático.

Para la obtención de los datos de entrenamiento y validación del modelo se realizan dos pruebas en ruta mediante las directivas de Euro 6 para la estimación de contaminantes mediante RDE (Real Driving Emissions), en las que se utiliza un sistema de medición de emisiones portátil, y un registrador que almacena los datos provenientes de OBD (On Board Diagnostics) y del GPS. A partir de los datos obtenidos en la primera ruta se determinan las prestaciones del vehículo y mediante aprendizaje automático se genera el modelo que estima las emisiones contaminantes, el cual es validado con los datos de la segunda ruta. Al comparar los resultados genera-dos por el modelo frente a los medidos en el RDE se obtienen errores relativos (%) de 0.0976, -0.2187, 0.2249 y -0.1379 en los factores de emisión de CO2, CO, HC y NOx respectivamente. Finalmente se alimenta al modelo con datos obtenidos en 1218 km de conducción aleatoria, obteniendo similares resultados a modelos basados en OBD y más próximos a las condiciones reales de circulación que generan modelos como el IVE (International Vehicle Emissions). La obtención de datos mediante el OBD y GPS presentes en los vehículos actuales es económica y la aplicación de modelos de aprendizaje automático para la estimación de emisiones contaminan-tes es una opción que abre un campo de trabajo sin la necesidad de contar con equipos a bordo de adquisición más costosos para la realización de campañas experimentales en muestras limitadas de vehículos por los costes asociados.

**Palabras clave:** modelo bajo costo, estimación de contaminantes, aprendizaje automático, emisiones basadas en GPS.

**Abstract**

This paper presents a novel method for estimating the pollutants emitted by vehicles powered by internal combustion engines in real driving conditions, without the need for extensive measurement campaigns or the use of instrumentation in the vehicle for prolonged periods of time; for which it is based on the positioning and speed signals generated by the GPS (Global Positioning System) and the machine learning application.

To obtain the data for training and validation of the model, two road tests are carried out using the Euro 6 directives for the estimation of pollutants through RDE (Real Driving Emissions), in which a portable emissions measurement system is used. and a logger that stores data from OBD (On Board Diagnostics) and GPS. From the data obtained in the first route, the performance of the vehicle is determined and through automatic learning, the model that estimates the polluting emissions is generated, which is validated with the data of the second route. When comparing the results generated by the model against those measured in the RDE, relative errors (%) of 0.0976, -0.2187, 0.2249 and -0.1379 are obtained in the emission factors of CO2, CO, HC and NOx, respectively. Finally, the model is fed with data obtained in 1218 km of random driving, obtaining similar results to OBD-based models and closer to the real traffic conditions generated by models such as IVE (International Vehicle Emissions). Obtaining data through the OBD and GPS present in current vehicles is cheap and the application of automatic learning models for the estimation of polluting emissions is an option that opens up a field of work without the need for on-board equipment. more expensive acquisition costs to carry out experimental campaigns on limited samples of vehicles due to the associated costs.

**Keywords:** low cost model, pollutant estimation, machine learning, GPS-based emissions.

# Introducción

Los vehículos propulsados por motores de combustión interna forman parte de las principales causas de contaminación en zonas urbanas [1], por lo que son motivo de estudio con el objeto de determinar la cantidad de contaminantes que emiten al ambiente [2][3]. La normativa vigente estipula que las emisiones contaminantes deben ser medidas durante conducción real [4], ya que estas son sustancialmente mayores a las obtenidas en pruebas de laboratorio [5], debido a que consideran factores como el tráfico, selección de rutas y el tipo de conducción que influyen directamente en el consumo y las emisiones generadas[6]. Ortenzi y Castermans et al determinan que parámetros como la excesiva aceleración y deceleración, marcha seleccionada errónea y velocidad elevada del motor y del vehículo, incrementan la emisión de contaminantes [7] [8], pero para determinar su real influencia se necesitan de extensas campañas de medición con vehículos ampliamente instrumentados [9]. [10] determina que las emisiones contaminantes dependen de factores propios del vehículo como el modelo, peso, tipo de combustible, nivel tecnológico y recorrido, y factores operacionales como la velocidad, aceleración, selección de marcha, gradiente del camino y la temperatura ambiente, por lo que todos los modelos de emisión deben considerar estos factores.

En este artículo se presenta un método novedoso para estimar las emisiones contaminantes utilizando como datos de entrada variables de conducción del vehículo tales como la velocidad y la pendiente, obtenidas mediante GPS a través de la aplicación de técnicas como Kmeans, árboles de clasificación y redes neuronales. Con el fin de crear un modelo de estimación de emisiones contaminantes se realizó un Real Driving Emissions (RDE) test en una ruta en la que se obtuvieron datos tantos de emisiones como de OBD y GPS, con estos datos se entrena una ANN (Artificial Neural Network) que fue validada con los datos obtenidos en un segundo RDE test, confirmando la validez del estimador de emisiones. Finalmente, este estimador fue aplicado a un data set de 1218 km de conducción real. Los resultados obtenidos fueron comparados con los obtenidos en el modelo IVE y OBD test evidenciando resultados similares.

# Materiales y Métodos

## Metodología Propuesta

Las emisiones contaminantes deben medirse en condiciones reales de circulación [4], con ello se consideran dentro de los resultados diferentes factores como el estilo de conducción, el tipo de combustible, la ubicación geográfica y condiciones ambientales en las que circula el vehículo [11], que actualmente no se consideran en los modelos utilizados por la Empresa de movilidad de la ciudad de Cuenca (EMOV-EP).

Para la estimación de contaminantes basados en los datos del GPS y aprendizaje automático se proponen los siguientes pasos:

* Recolección de datos de GPS y emisiones contaminantes en condiciones reales de conducción en dos rutas basadas en [4].
* Entrenamiento del modelo basado en aprendizaje automático con los datos de la ruta 1
* Validación del modelo obtenido con los datos de la ruta 2
* Aplicación del modelo generado al conjunto de datos de 1218.9 km
* Procesamiento y presentación de resultados

## Recolección de datos en condiciones reales de conducción

Para la recolección de datos el vehículo utilizado es un Kia Sportage modelo 2018, según [12] es el vehículo más vendido en el Ecuador en el segmento SUV. Posee un motor DOHC 2.0L, transmisión manual de 6 marchas y al momento de la experimentación contaba con 18720 km de recorrido y todas las operaciones de mantenimiento recomendadas por el fabricante. El sistema portátil de medición de emisiones (PEMS) utilizado es el analizador de gases Brain Bee AGS-688 que trabaja mediante el método no dispersivo de absorción de infrarrojos (NDIR) para la medición de CO2 [%], CO [%] e HC [ppm] y celda electroquímica para la medición de O2 [%] y NOX [ppm]. El equipo es alimentado mediante una batería independiente del vehículo de pruebas según se establece en [4]. El consumo de combustible se mide utilizando el AIC Fuel Flow Master 5004. El GPS utilizado es el incorporado dentro del data logger Freematics ONE+, que almacena los datos de latitud (Lat), longitud (Long), altura (Alt) y velocidad (VGPSo) del vehículo en una tarjeta SD en formato CSV. A más de los datos del GPS el equipo almacena los datos de conducción provenientes de OBD como el régimen de giro del motor (RPM), presión de colector de admisión (MAP), y la velocidad del vehículo (VOBD).

## Rutas de prueba

Las 2 rutas seleccionadas para la recolección de datos se sitúan en la ciudad de Cuenca-Ecuador como se muestra en la figura 1, y se componen de los segmentos urbano, rural y en autopista, realizados en el Centro Histórico, la carretera Panamericana Norte y la Autopista Cuenca Azogues respectivamente, según lo dispuesto en [4] para determinar las emisiones en condiciones reales de conducción (RDE).

Mapa

Descripción generada automáticamente **Figura 1.** Rutas de Prueba. Fuente: elaboración propia.

Los resultados obtenidos en la ejecución de las rutas se muestran en la Tabla 1. Las 2 pruebas se efectuaron bajo las mismas condiciones, sin presencia de lluvias o vientos fuertes, a una temperatura ambiente media de 14°C, el motor en caliente y una masa total de *m* =1719.5 kg que corresponde al vehículo con el tanque lleno de combustible, la instrumentación necesaria y dos pasajeros.

Tabla . Características de las Rutas.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Características | Requisitos RDE |  | Ruta 1 | Ruta 2 | Unidad |
| Distancia Urbana | >16 |  | 21.63 | 17.87 | [km] |
| Distancia Rural | >16 |  | 21.24 | 18.77 | [km] |
| Distancia en Autopista | >16 |  | 19.61 | 18.76 | [km] |
| Proporción Urbana | 29 – 44 |  | 34.61 | 32.25 | [%] |
| Proporción Rural | 23 – 43 |  | 34.01 | 33.87 | [%] |
| Proporción Autopista | 23 – 43 |  | 31.38 | 33.88 | [%] |
| Tiempo muerto | 10 – 30 |  | 11.61 | 14.45 | [%] |
| Diferencia de altitud entre partida y llegada | <100 |  | -4.4 | 54 | [m] |

**2.4 Estimación de contaminantes**

Las emisiones generadas en condiciones reales de conducción en las 2 rutas establecidas son medidas por el analizador de gases en concentración volumétrica siendo necesario convertirlas a concentración másica mediante el procedimiento descrito en [4].

El flujo másico de escape (g/s) se determina mediante (1).

(1)

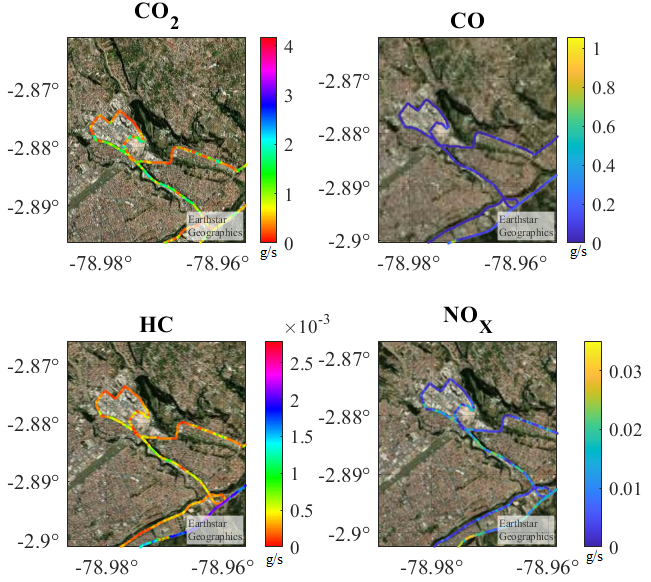
Donde es el flujo másico de aire obtenido desde OBD y el flujo de combustible medido mediante el flujómetro ubicado en la línea de combustible. La corrección de las emisiones de base seca a húmeda se efectúa mediante (2).

(2)

En la que es la concentración seca del contaminante *j* medida por el analizador de gases y y son las concentraciones en base seca de CO2 y CO respectivamente. Las emisiones másicas [g/s] de cada contaminante se obtienen mediante la ecuación (3).

(3)

En la que *i* representa el número de cada medición, *cj* la concentración instantánea del contaminante *j* en [ppm], la relación entre la densidad de cada componente y la densidad global de los gases de escape expresados como: . Los resultados instantáneos obtenidos pueden verse en la figura 2.



**Figura 2**. Emisiones Instantáneas en RDE

La emisión de cada contaminante (g) en el ciclo de conducción es igual a la sumatoria de su emisión instantánea respecto al tiempo (4).

donde es el flujo másico instantáneo del contaminante *j*, *n* es el número de muestras del dataset y es el tiempo de muestreo que es igual a 0.1 s. Los factores de emisión de cada contaminante (g/km) en la sección *k* del RDE se determina por (5).

(5)

En la que es la masa del contaminante *j* y la distancia recorrida en la sección *k* del RDE, *k* asume los valores de *u, r, a* para las secciones urbana, rural y autopista respectivamente. Los resultados obtenidos en la ruta 1 se muestran en la tabla 3.

Tabla 3. Factores de Emisión

|  |  |  |  |
| --- | --- | --- | --- |
| F | Urbano  g/km | Rural  g/km | Autopista  g/km |
| CO2 | 82.574 | 79.410 | 45.598 |
| CO | 2.8442 | 6.6942 | 6.6363 |
| HC | 0.0314 | 0.0427 | 0.0736 |
| NOx | 0.6597 | 0.8088 | 0.6821 |

**2.5 Procesamiento de la señal de Velocidad**

Según lo mostrado en [7] las prestaciones del motor se pueden determinar a partir de datos del recorrido tomados por el GPS, como la pendiente recorrida y la velocidad, por ello es muy importante que la velocidad utilizada en la generación y aplicación del modelo sea lo más estable y con el menor error posible ya que muchas de las variables de análisis se determinan a partir de ella[13] [14]. La velocidad del vehículo VGPSo es mayor a VOBD, a más de que presenta oscilaciones fuertes que generan error en el cálculo de la aceleración del vehículo, por tal motivo se aplica el algoritmo de Savitzky-Golay para disminuir las oscilaciones y suavizar los escalones generados por la adquisición de datos discretos [15]. Para disminuir el error de VGPSo frente VOBD se aplica una regresión lineal mostrada en la ecuación (6). La velocidad procesada obtenida a partir del GPS VGPS se expresa como:

(6)

El error medio disminuye de 1.59 a 0.08 km/h, como puede verse en la figura 3.



**Figura 3.** Procesamiento de señales

El procesamiento generado a partir de los datos de la ruta 1 se aplica para su validación a los datos medidos en la ruta 2 obteniendo los resultados que se muestran en la tabla 2.

**2.5 Estimación de predictores**

Tabla 2. Validación de resultados. Fuente: elaboración propia.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Características Ruta 2** | **VOBD** | **VGPSo** | **VGPS** | **Unidad** | **Error [%] VGPSo-VOBD** | **Error [%] VGPS-VOBD** |
| Velocidad Máxima | 105 | 109,42 | 105,15 | [km/h] | 4,21 | 0,14 |
| Velocidad Media | 39,16 | 40,75 | 39,24 | [km/h] | 4,06 | 0,2 |
| Velocidad Media Urbana | 22,63 | 23,69 | 22,79 | [km/h] | 4,68 | 0,69 |
| Velocidad Media Rural | 50,31 | 52,24 | 50,31 | [km/h] | 3,84 | 0 |
| Velocidad Media Autopista | 84,73 | 87,86 | 84,64 | [km/h] | 3,68 | -0,11 |

Según lo mostrado en [16] las variables más influyentes en la generación de emisiones contaminantes son las que se muestran en la tabla 2.

Tabla 2. Principales predictores

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Emisión | Predictores | | | |
| CO2 | GEAR | RPM | TPS | MAP |
| CO | TPS | MAP | VOBD | RPM |
| HC | TPS | MAP | VOBD | GEAR |
| NOx | MAP | RPM | GEAR | VOBD |

Estos predictores se obtienen de manera directa mediante OBD, por lo que no pueden ser utilizados para entrenar el modelo basado en GPS. La marcha utilizada por el conductor no se puede determinar de manera directa por VGPS ya que como se muestra en la figura 4, las marchas seleccionadas no dependen de la velocidad de circulación. Por tal efecto se deben buscar relaciones entre los datos proporcionados por el GPS y el desempeño del motor.



**Figura 4.** Marchas según la velocidad

La velocidad del vehículo se relaciona con la velocidad del motor mediante las relaciones de transmisión de cada una de las 6 marchas del vehículo. Con el fin de obtener una etiqueta para el entrenamiento de un árbol de clasificación que estime la marcha utilizada por el conductor se utilizan los datos obtenidos en la ruta 1, a los se le aplican el algoritmo de Kmeans, en específico al vector (7):

que genera una etiqueta para cada uno de los 7 grupos *G* obtenidos a partir de sus centroides [17], los grupos generados corresponden a cada una de las 6 marchas del vehículo y a la posición neutral. El valor de *RTG* se obtiene estimando la mediana de *Ri*para cada marcha. Para el vehículo de análisis la 5ta marcha (*G=5*) es directa por lo que la relación de transmisión *R5*=1, a partir de ello se determina la relación de transmisión del grupo corona diferencial RC como se muestra en (8).

En donde *RN = 0.3607 m* que corresponde al valor del radio efectivo del neumático. La relación de transmisión *RG* para cada marcha se calcula mediante (9).

El factor de masas rotativas se determina y la eficiencia de cada marcha se determinan por (10) y (11) respectivamente.

Los valores de las relaciones de transmisión, factor de masas rotativas y eficiencia de cada marcha se muestran en la tabla 3.

Tabla 3. Valores característicos del vehículo

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| G | 1 | 2 | 3 | 4 | 5 | 6 |
|  | 7.3 | 13.1 | 19.9 | 25.7 | 31.4 | 37.3 |
|  | 4.31 | 2.39 | 1.58 | 1.22 | 1 | 0.84 |
|  | 1.086 | 1.054 | 1.046 | 1.043 | 1.042 | 1.041 |
|  | 0.803 | 0.806 | 0.809 | 0.811 | 0.814 | 0.816 |

En la figura 5 se muestran los valores de torque obtenidos en el banco dinamométrico a diferentes aperturas del acelerador con los que mediante la ecuación (12) se determina el valor de aceleración del vehículo, en donde T es el torque entregado por el motor en función de la marcha escogida por el conductor.

Gráfico, Gráfico de líneas

Descripción generada automáticamente

**Figura 5**. Torque del motor

La fuerza tractora del vehículo *FT* está relacionada con la aceleración longitudinal aX, la fuerza de frenado *Fbrk*, la resistencia a la rodadura y aerodinámica *Fres*, y la resistencia a las pendientes *Fslope* , como se muestra en la ecuación 13.

(13)

La aceleración longitudinal se obtiene a partir de la velocidad *VGPS* mediante la ecuación 14.

La resistencia a la rodadura y aerodinámica se obtienen mediante (9), en la que los coeficientes de adherencia estática y dinámica son y respectivamente, la densidad del aire , el coeficiente aerodinámico , y el área frontal del vehículo . La resistencia a la pendiente se expresa en (15).

La resistencia que se presenta por el ascenso en pendiente se muestra en la ecuación (16).

La fuerza tractora del vehículo se expresa en (17), y puede ser determinada con los datos que provienen del GPS. La fuerza de frenado *Fbrk* se aplica cuando *FT* es nula, es decir no se puede accionar el acelerador y el freno al mismo tiempo, y la resistencia aerodinámica y a la rodadura nunca serán negativas pues dependen de la velocidad, por lo que las fuerzas tractoras negativas se deben a frenados y a ascensos en pendientes.

La fuerza tractora depende del desempeño del motor, que involucra la apertura del acelerador, la presión del colector de admisión y la marcha seleccionada por el conductor, con lo que se obtiene un elemento predictor que al igual que la aceleración longitudinal y la etiqueta de marcha seleccionada se utilizan para entrenar un árbol de clasificación que determine la marcha utilizada por el conductor. El vector tomado en la ruta 1 para el entrenamiento es *In*i=[*FTi, axi, VGPSi*], El árbol de clasificación obtenido tiene 7 splits cuya efectividad en porcentaje se muestra en la matriz de confusión de la tabla 4. La mayor taza de aciertos se da en neutro, 1ª , 2ª y 3ª marchas, mientras que en 4ª, 5ª y 6ª la eficiencia del modelo determinado disminuye, debido a que las prestaciones del vehículo en estas condiciones son muy similares.

Tabla 4. Matriz de Confusión

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| MARCHA VERDADERA | **0** | 99.5 | 2.0 |  |  |  |  |  |
| **1** | 0.5 | 89.3 | 6.2 | 0.1 |  |  |  |
| **2** |  | 8.6 | 82.5 | 8.8 | 0.6 |  |  |
| **3** |  | 0.1 | 10.6 | 83.6 | 19.8 | 3.6 | 2.1 |
| **4** |  |  | 0.7 | 6.4 | 61.5 | 17.7 | 6.3 |
| **5** |  |  |  | 0.7 | 14.3 | 45.3 | 19.6 |
| **6** |  |  |  | 0.3 | 3.8 | 33.4 | 71.9 |
|  |  | **0** | **1** | **2** | **3** | **4** | **5** | **6** |
|  |  | MARCHA PREDICHA | | | | | | |

Una vez determinada la marcha seleccionada por el conductor se puede determinar el régimen de giro del motor a partir de *VGPS* mediante la ecuación (7). Los valores obtenidos se pueden visualizar en la figura 6, que al comparase con los datos obtenidos mediante OBD presentan diferencias significativas en los cambios de marcha. Durante el cambio de marcha el conductor no acciona el acelerador por lo que la velocidad del motor desciende al valor de marcha mínima de 663 RPM determinado por la moda del tiempo muerto en la ruta 1, como se aprecia en la figura 6.

Gráfico

Descripción generada automáticamente

**Figura 6.** Escalones de Velocidad

## Con los datos obtenidos de *Ii* y la etiqueta de cada marcha, se entrenan 4 redes neuronales para la estimación de las emisiones contaminantes. Cada red tiene 3 neuronas en la capa de entrada, 10 en la capa oculta y una a la salida. Los índices de determinación R2 obtenidos son de 0.935, 0.923, 0.914 y 0.943 para las emisiones de CO2, CO, HC y NOx respectivamente. El modelo *MGPS* obtenido a partir de la señal del GPS fue alimentado con los datos medidos en la ruta 2 generando resultados muy similares al modelo RDE como se puede observar en la figura 7.



**Figura 7.** Factores de Emisión MGPS-RDE

Se aplican en el modelo generado los datos obtenidos en 1218 km de recorrido aleatorio en la zona urbana, rural y autopista de la ciudad de Cuenca. Las paradas son consideradas como los períodos en los que la velocidad del vehículo es inferior a 1 *km/h* según lo expuesto en [4] estos resultados no consideran las operaciones especiales del motor durante el arranque en frío, mismas que deben ser estudiadas a futuro. En la figura 8 puede apreciarse que la 1°, 2° y 3° marchas se utilizan en bajas velocidades medias, en pequeños recorridos que se dan de manera mayoritaria en zona urbana y muy pocas veces en rural y autopista, y que a mayores velocidades las emisiones de CO2, CO y NOX disminuyen. El error que se obtiene en el cálculo de la velocidad media de circulación en cada marcha se debe al modelo que determina la marcha seleccionada por el conductor.



**Figura 8.** Factores de Emisión

Los factores de emisión medios para cada modelo se muestran en la tabla 5, los valores estimados por IVE son mayores a los otros modelos analizados, la principal diferencia la constituye el factor de emisión de CO2 que es fuertemente influenciado por las bajas velocidades de circulación en zona urbana.

Tabla 5. Factores de Emisión

|  |  |  |  |
| --- | --- | --- | --- |
| F | IVE [g/km] | RDE g/km] | MGPS [g/km] |
| CO2 | 208.97 | 70.23 | 110.93 |
| CO | 11.83 | 5.33 | 6.95 |
| NOX | 0.661 | 0.7199 | 0.69 |
| HC | 0.1477 | 0.0485 | 0.041 |

# Conclusiones

Este artículo propone un método para la estimación de emisiones contaminantes basado en datos desde GPS aplicando Machine Learning a una gran cantidad de datos, para esto inicialmente se obtuvo un clasificador con alta efectividad para la evaluación de la marcha seleccionada por el conductor basado en la obtención de etiquetas mediante Kmeans y el posterior entrenamiento de un árbol clasificador, los errores generados se producen en los pequeños instantes que dura la transición entre marchas. El cálculo de emisiones contaminantes se dio a partir de la determinación de la importancia de los predictores en los datos obtenidos en dos rutas del RDE test, para después entrenar 4 ANNs que obtuvieron altos índices de determinación R2 que dan como válido el método aplicado.

El modelo obtenido es más robusto a diferentes condiciones de tráfico y presenta mejores resultados en circulación a bajas velocidades medias de circulación que los modelos IVE y RDE, por lo que se recomienda ser utilizado para el cálculo de factores de emisión y estimación de inventarios de emisiones vehiculares.

Como desarrollos futuros, el modelo obtenido se puede ajustar a diferentes parámetros como la longevidad del vehículo, y el funcionamiento en frío. Se debe replicar la metodología planteada en los modelos de vehículos con mayor presencia y actividad del parque automotor.

# Referencias

[1] UN, “The World’s Cities in 2016,” New York, 2013.

[2] S. Wang, Z. Li, J. Tan, W. Guo, and L. Li, “A method for estimating carbon dioxide emissions based on low frequency GPS trajectories,” *Proceedings - 2017 Chinese Automation Congress, CAC 2017*, vol. 2017-Janua, pp. 1960–1964, 2017, doi: 10.1109/CAC.2017.8243091.

[3] Z. Kan, L. Tang, M. P. Kwan, and X. Zhang, “Estimating vehicle fuel consumption and emissions using GPS big data,” *International Journal of Environmental Research and Public Health*, vol. 15, no. 4, pp. 1–23, 2018, doi: 10.3390/ijerph15040566.

[4] R. Ue and D. E. L. A. Comisión, “REGLAMENTO (UE) 2016/427. Por el que se modifica el Reglamento (CE) n.o 692/2008 en lo que concierne a las emisiones procedentes de turismos y vehículos comerciales ligeros (Euro 6),” *Diario Oficial de la Unión Europea*, vol. L 82, pp. 1–98, 2016.

[5] K. Kurtyka and J. Pielecha, “The evaluation of exhaust emission in RDE tests including dynamic driving conditions,” *Transportation Research Procedia*, vol. 40, pp. 338–345, 2019, doi: 10.1016/j.trpro.2019.07.050.

[6] G. Fontaras, N. G. Zacharof, and B. Ciuffo, “Fuel consumption and CO2 emissions from passenger cars in Europe – Laboratory versus real-world emissions,” *Progress in Energy and Combustion Science*, vol. 60, pp. 97–131, 2017, doi: 10.1016/j.pecs.2016.12.004.

[7] F. Ortenzi and M. A. Costagliola, “A new method to calculate instantaneous vehicle emissions using OBD data,” *SAE Technical Papers*, no. April 2010, 2010, doi: 10.4271/2010-01-1289.

[8] J. Castermans, A. Brusselmans, and C. Jean-Pandazis, “Cooperative Mobility Systems and Services for Energy Efficiency,” 2010. [Online]. Available: https://cordis.europa.eu/project/id/247908/es

[9] H. Hajmohammadi, G. Marra, and B. Heydecker, “Data-driven models for microscopic vehicle emissions,” *Transportation Research Part D: Transport and Environment*, vol. 76, no. October, pp. 138–154, 2019, doi: 10.1016/j.trd.2019.09.013.

[10] P. G. Boulter, T. J. Barlow, I. S. Mccrae, S. Latham, and C. Parkin, “Emission factors 2009: Report 1 – a review of methods for determining hot exhaust emission factors for road vehicles; PPR353,” p. 116, 2009.

[11] C. Samaras, D. Tsokolis, S. Toffolo, G. Magra, L. Ntziachristos, and Z. Samaras, “Enhancing average speed emission models to account for congestion impacts in traffic network link-based simulations,” *Transportation Research Part D: Transport and Environment*, vol. 75, pp. 197–210, 2019, doi: 10.1016/j.trd.2019.08.029.

[12] Asociación de Empresas Automotrices del Ecuador, “Automotive in figures,” Quito, 2018.

[13] L. Thibault, P. Degeilh, O. Lepreux, L. Voise, G. Alix, and G. Corde, “A new GPS-based method to estimate real driving emissions,” *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pp. 1628–1633, 2016, doi: 10.1109/ITSC.2016.7795776.

[14] T. H. Witte and A. M. Wilson, “Accuracy of non-differential GPS for the determination of speed over ground,” *Journal of Biomechanics*, vol. 37, no. 12, pp. 1891–1898, 2004, doi: 10.1016/j.jbiomech.2004.02.031.

[15] R. W. Schafer, “What Is a Savitzky-Golay Filter? [Lecture Notes],” no. July, pp. 111–117, 2011.

[16] N. D. Rivera-Campoverde, J. L. Muñoz-Sanz, and B. D. V. Arenas-Ramirez, “Estimation of pollutant emissions in real driving conditions based on data from OBD and machine learning,” *Sensors*, vol. 21, no. 19, Oct. 2021, doi: 10.3390/S21196344.

[17] A. K. Jain, “Data clustering: 50 years beyond K-means,” *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010, doi: 10.1016/j.patrec.2009.09.011.