

ANÁLISIS DE PATRONES DE CONDUCCIÓN DE VEHÍCULOS.

MARIO ALBERTO SALAMANCA RÍOS
JUAN FELIPE MORA MARTÍNEZ

PONTIFICIA UNIVERSIDAD JAVERIANA.
FACULTAD DE INGENIERÍA.
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA
BOGOTÁ DC.
2016.

ANÁLISIS DE PATRONES DE CONDUCCIÓN DE VEHÍCULOS.

MARIO ALBERTO SALAMANCA RIOS

JUAN FELIPE MORA MARTINEZ

Trabajo presentado como requisito para
Optar al título de Ingeniero electrónico.

Director: ING. DANIEL JARAMILLO RAMÍREZ, PhD

PONTIFICIA UNIVERSIDAD JAVERIANA.
FACULTAD DE INGENIERÍA.
DEPARTAMENTO DE INGENIERÍA ELECTRÓNICA
BOGOTÁ DC.

2016.

Dedicado a Olga Patricia Ríos, Cristina Martínez y Oscar Mora,
Amigos y demás familiares.

Agradecimientos.

Gracias a Dios, por sus bendiciones, al ingeniero Daniel Jaramillo por su ayuda y orientación profesional para alcanzar este logro. A la Pontificia Universidad Javeriana por ser nuestra alma mater.

Agradecemos a un continuo de personas que han estado con nosotros en todas las situaciones y que nos han ofrecido su apoyo y su cariño incondicional; especialmente agradecer a nuestras familias, que son el motor de nuestras vidas. Además, nuestros amigos y compañeros con los que hemos compartido experiencias de vida, a mis profesores que me brindaron sus conocimientos para nuestra formación como profesionales.

Contenido

1. INTRODUCCIÓN.....	8
2. MARCO TEÓRICO:.....	9
2.1. Dinámica del movimiento.....	9
2.2. Filtro digital.....	10
2.2.1. Filtro pasa bajos:.....	10
2.3. Clasificación no supervisada.....	11
2.4. Algoritmo de agrupamiento (clustering).....	11
2.4.1. K-means:.....	12
2.5. Análisis de Componentes Principales (PCA).....	13
2.6. Acelerómetro.....	13
2.7. iPhone 6.....	13
2.8. Sensor Log:.....	14
3. OBJETIVO DEL PROYECTO.....	15
4. DESARROLLO.....	16
4.1. Fase uno: Estado de Arte.....	16
4.2. Fase dos: Construcción de la base de datos.....	16
4.2.1. Experimento.....	20
4.3. Fase tres: Acondicionamiento y clasificación de los patrones obtenidos.....	20
4.3.1. Acondicionamiento de las muestras.....	20
4.3.2. Clasificación de eventos.....	23
4.3.2.1. Aceleraciones.....	23
4.3.2.2. Frenadas.....	23
4.3.2.3. Cambio de carril a Izquierda y a Derecha.....	24
4.3.2.4. Baches o Huecos.....	24
4.3.3. Clasificación no supervisada.....	25
4.3.3.1. Frecuencia de Picos Que Superan los 0.09G	25
4.3.3.2. Promedio Sobrepasso en algún factor los picos mayores a Referencia mínima	26
4.3.3.3. Porcentaje de tiempo que dura frenando (Frenos mayores a 0.09G).....	27
4.3.3.4. Frecuencia de Picos Que Superan los 0.15G	27
4.3.3.5. Frecuencia de picos que superan los 0.06 G/s.....	28
4.3.3.6. Porcentaje de tiempo que dura frenando (Frenos mayores a 0.15G).....	29
4.3.3.7. Porcentaje de frenadas mayores a 0.09G	30
4.3.3.8. Porcentaje de frenadas mayores a 0.15G	31
4.3.3.9. Promedio Sobre Paso en algún factor los picos mayores a 0.09G	32
4.3.3.10. Promedio Sobrepasso en algún factor los picos mayores al 5% del promedio de la señal	32

4.3.3.11.	Frecuencia de Picos Que Superan los 0.09G.....	33
4.3.3.12.	Promedio Sobre Paso en algún factor los picos mayores a 0.09G	34
4.3.3.13.	Porcentaje de tiempo que dura Acelerando (Aceleraciones mayores a 0.09G).....	34
4.3.3.14.	Frecuencia de Picos Que Superan los 40 Km/H.....	35
4.3.3.15.	Promedio de los picos previos a una detención.....	35
4.3.3.16.	Promedio de W cuando los picos de aceleración supera 0.09 G	36
4.3.3.17.	Frecuencia de W cuando los picos de aceleración supera 0.09 G	37
4.4.	Fase Cuatro: Análisis estadístico y clasificación de patrones.	37
4.4.1.	Reducción de dimensionalidad	37
4.4.2.	Agrupamiento de datos.	38
5.	PROTOCOLO DE PRUEBAS Y ANÁLISIS DE RESULTADOS.	40
6.	CONCLUSIONES Y RECOMENDACIONES:.....	48
7.	Bibliografía	49
8.	ANEXOS	50

Tabla de ilustraciones.

Fig. 1 Variables y dinámicas de desplazamiento de estación a estación para vehículos eléctricos.	10
Fig. 2 Especificación de un filtro pasó bajo. ⁴	11
Fig. 3. Dendrograma resultante de clasificación no supervisada [7].....	12
Fig. 4 Coordenadas acelerómetro iPhone.....	14
Fig. 5. Buses del estudio.	17
Fig. 6 Ruta del estudio.	17
Fig. 7. Anclaje al bus teléfono celular.....	18
Fig. 8. Comparación de acelerómetros incluidos en iPhone 6.	18
Fig. 9. Sistema de coordenadas	19
Fig. 10. Traza de aceleraciones de un bus en los ejes X, Y y Z.....	21
Fig. 11. Periodograma PSD de una traza de prueba.	22
Fig. 12. Traza de aceleración original.	22
Fig. 13 traza de aceleración con filtrado de promedio	23
Fig. 14 Traza de aceleración filtrada.	23
Fig. 15. Acondicionamiento de una muestra por medio del Filtrado	23
Fig. 16 Traza de aceleración (eje Y)- Aceleraciones.	23
Fig. 17. Traza de aceleración (eje Y)-Desaceleraciones.	24
Fig. 18 Aceleración en el eje X- Giros.....	24
Fig. 19 Aceleración en el eje Z	25
Fig. 20: Frecuencia de Picos que Superan los $ 0.09 G $	26
Fig. 21: Porcentaje de tiempo que dura frenando (Frenos mayores a $ 0.09 $)	27
Fig. 22: Frecuencia de Picos que superan los $ 0.15 G $	28
Fig. 23: Frecuencia de Picos que superan los $0.06 G/s$	29
Fig. 24: Porcentaje de tiempo que dura frenando (Frenos mayores a $ 0.15 G $)	30
Fig. 25: Porcentaje de frenadas mayores a $ 0.09 G $	31
Fig. 26: Porcentaje de frenadas mayores a $ 0.15G $	32
Fig. 27: Promedio de sobrepaso en algún factor los picos mayores al $ 5\%$ del promedio de la señal 	33
Fig. 28: Frecuencia de picos que superan los $0.09G$	34
Fig. 29: Frecuencia de picos que superan los $40Kh/h$	35
Fig. 30: Promedio de los picos previos a una detención	36
Fig. 31. Histograma general de las 17 características	40
Fig. 32. Metrica Squeulidean.....	42
Fig. 33, Metrica Cityblock	42
Fig. 34. Metrica Cosine.....	42
Fig. 35. Metrica Correlation.....	42
Fig. 36 Metricas Tecnica de agrupamiento.	42
Fig. 37. Valores de Silueta Características principales.	43
Fig. 38. PCA para las tres componentes principales, con 4 clusters en tres dimensiones.....	45
Fig. 39 PCA para las tres componentes principales, con 4 clusters en dos dimensiones 2PCA vs 1PCA ..	45
Fig. 40. PCA para las tres componentes principales, con 4 clusters en dos dimensiones 3PCA vs 1PCA. .	46
Fig. 41. PCA para las tres componentes principales, con 4 clusters en dos dimensiones 3PCA vs 2PCA. .	46

1. INTRODUCCIÓN.

El transporte público en bus urbano está pensado, de manera general, como una sucesión de paradas en las que los viajeros suben y bajan del vehículo en función de su destino, y donde los usuarios pueden viajar de pie o sentados. Estas situaciones del transporte urbano provocan que, en los momentos de desaceleración o aceleración exagerada, se multipliquen las posibilidades de producirse accidentes. Además, las personas que viajan de pie son susceptibles de sufrir caídas y golpes, en especial las personas mayores. Según un estudio de la Cámara de Comercio de Bogotá, publicado en el periódico EL TIEMPO [1] la mayor cantidad de quejas y reclamos que los ciudadanos hacen respecto al servicio, se debe entre otras a la forma como se conducen los vehículos de transporte público: frenados bruscos, aceleraciones indebidas, exceso de velocidad, cambios abruptos de carril. Este problema no solo va en contra de los estándares mínimos de calidad que el servicio público debe prestar, sino que además, va en contra de la comodidad y seguridad de los usuarios. Según SURA empresa de seguros “...el 85% de todos los accidentes de tránsito que ocurren son responsabilidad directa o indirecta del conductor y de las decisiones que toma y por ello, recae en él la responsabilidad de conducir siempre en forma preventiva” [2].

Con el fin de ayudar a solucionar en parte este inconveniente, se vio necesario desarrollar un estudio que de una manera más objetiva muestre la forma como son conducidos los vehículos de servicio público en la ciudad de Bogotá, ya que no existe información disponible sobre estos estudios en la ciudad. Este proyecto se adelantó con la toma de muestras de aceleraciones, desaceleraciones, velocidad y posición de los vehículos (que llamaremos en adelante dinámica del movimiento), en una selección de buses duales sobre la carrera séptima y un análisis de las cifras obtenidas. Igualmente este estudio servirá de base de recomendaciones en la forma de conducción de los vehículos de transporte público y generará una mejora en la calidad de vida de los usuarios del Sistema Integrado de Transporte Público (Transmilenio dual), resultando como objetivo principal; investigar dinámicas del movimiento en vehículos de transporte público, en el sector de la carrera séptima entre calles cuarenta y cinco a treinta y dos de la localidad de Chapinero de Bogotá, con el uso de acelerómetros y GPS embebidos en teléfonos celulares.

Diferentes estudios de seguridad realizados en ciudades como Valencia y Zaragoza, España [3], La Universidad del Estado de Pennsylvania [4], y en San Diego, California [5], de transporte público y vehículos particulares han analizado las dinámicas de movimiento, encontrado valores límites sugeridos y en algunos casos estandarizándolos como normas locales para la aceleración máxima y velocidad máxima dentro del casco urbano.

El presente informe documenta del trabajo desarrollado durante el presente año, el cual contó con un anteproyecto elaborado y un trabajo desarrollado en las fases de estado del arte, fase construcción de la base de datos, Fase de Procesamiento y clasificación de 17 patrones obtenidos, análisis de resultados y recomendaciones finales.

Las gráficas obtenidas muestran las dinámicas del movimiento en el transporte público de un sector de Chapinero, mostrando un trabajo de modelamiento de datos y una generación de graficas interesantes.

Esperamos que este trabajo sea del agrado de todos.

2. MARCO TEÓRICO:

El propósito del presente capítulo es el de recolectar la información aportada y contenida en la literatura relacionada con dinámica del movimiento, análisis estadístico, clasificación no supervisada, algoritmo de agrupamiento (clustering), Principal Component Analysis (PCA), el filtro digital, acelerómetro, iPhone 6, sensor log, los cuales serán explicados más adelante.

El presente modelo es de tipo microscópico, enfocándose en la descripción del comportamiento del flujo del tráfico vehicular describiendo las entidades discretas individuales y atómicas que interactúan unas con otras (en este caso cada vehículo individual). Es un modelo discreto, denominado como modelo con autómatas celulares (de los cuales los más importantes son los modelos Nagel-Schreckenberg y Fukui-Ishibashi). Kai Nagel, Michael Schreckenberg, M. Bando, P. G. Gipps, M. Fukui y Y. Ishibashi son algunos de los mayores investigadores en esta área¹.

El modelo de Nagel y Schreckenberg (Na-Sch) es un modelo de flujo de tránsito vehicular con un autómata celular (AC) probabilístico. Por ende, es un modelo de espacio y tiempo discretos, donde cada célula del autómata equivale ya sea a un vehículo en movimiento con cierta velocidad v o a un espacio vacío de la avenida donde se encuentran los vehículos.

El modelo Na-Sch original sirve para modelar autopistas de un carril (ya sea abiertas o en circuito) con vehículos homogéneos². Creado en 1992 por los científicos Kai Nagel y Michael Schreckenberg, el modelo Na-Sch se ha convertido en la base de muchos otros modelos discretos de tráfico vehicular, debido a su sencillez que sin embargo es capaz de modelar adecuadamente los fenómenos de congestión en autopistas. Esto sucede ya que las gráficas de densidad de tráfico vs. flujo de tráfico son muy similares a las observadas empíricamente en diversas avenidas reales. Además, una simulación de este modelo permite observar las ondas de tráfico comunes en el flujo vehicular.

2.1. Dinámica del movimiento.

La mayoría de los análisis de las características del vehículo, el tiempo de viaje, la programación, el confort de los pasajeros, se realizaron a través de las variables básicas del movimiento del vehículo. Estas variables son la posición y sus tres derivadas con respecto al tiempo:

$$\text{Posición: } s = f(t); \quad (1)$$

$$\text{Velocidad: } v = \frac{ds}{dt}; \quad (2)$$

$$\text{Aceleración: } a = \frac{d^2s}{dt^2}; \quad (3)$$

$$\text{Jerk: } z = \frac{d^3s}{dt^3}; \quad (4)$$

La posición s [m], es analizada a través de diagramas tiempo-distancia, este diagrama es frecuentemente utilizado en la programación de líneas de tránsito, la determinación de señales de bloque para el tránsito rápido, el análisis del movimiento del vehículo en una calle con intersecciones y el análisis de tiempo de viaje estación-estación.

La velocidad v [m/s], está representada en la distancia sobre el tiempo, también puede ser definida matemáticamente como el área bajo la curva de la coordenada de distancia sobre el tiempo Ecuación (5) esto se puede ver mejor en la Fig. 1.

$$s = \int v dt; \quad (5)$$

¹ Tomado de: https://es.wikipedia.org/wiki/Tr%C3%A1nsito_vehicular, el día 10 de noviembre de 2016

² Tomado de: https://es.wikipedia.org/wiki/Modelo_Nagel-Schreckenberg, el día 10 de noviembre de 2016

La aceleración [m/s²], es la derivada de la velocidad con respecto al tiempo, se representa en el diagrama como una función del tiempo, y es representada por la pendiente de la curva de velocidad. Este parámetro influye en el tiempo del viaje, el consumo de combustible, especialmente en servicios con paradas frecuentes. Los valores máximos suelen estar limitados por el confort y la seguridad de los pasajeros que van de pie al interior del bus.

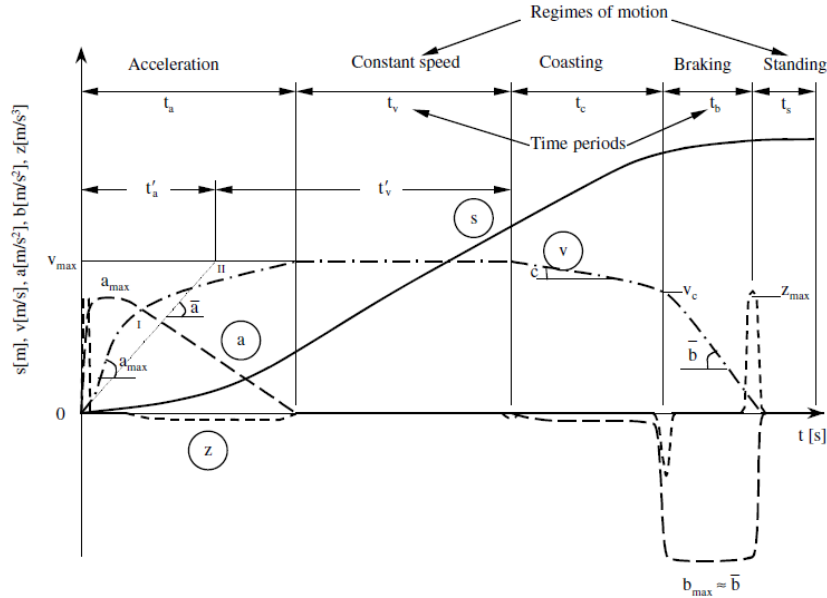


Fig. 1 Variables y dinámicas de desplazamiento de estación a estación para vehículos eléctricos.³

Jerk z [m/s³] también conocido como jalón, es el cambio de aceleración por unidad de tiempo, es en realidad el factor más importante para la comodidad de conducción de pasajeros. Las personas que van al interior de un vehículo a una velocidad constante no sienten fuerzas incómodas, en cambio el jerk es la fuerza que tiende a desviar a los pasajeros de su posición inicial. [6]

2.2. Filtro digital.

Un filtro es un elemento que discrimina una determinada frecuencia o gama de frecuencia de una señal eléctrica que pasa a través de él, pudiendo modificar su amplitud y fase. Un filtro digital es un filtro que opera en señales discretas y cuantizadas, estos implementados en circuitos digitales o en programas informáticos. Generalmente son un procesamiento matemático sobre una señal y funcionan mediante el uso de la Transformada de Fourier, obteniendo a la salida el resultado del procesamiento matemático. En este análisis se utilizó el filtro pasa bajos, explicado a continuación:

2.2.1. Filtro pasa bajos:

Es un tipo de filtro electrónico caracterizado principalmente por el paso de frecuencias bajas situadas por debajo de una frecuencia de corte, atenuando las frecuencias altas encontradas por encima de la frecuencia de corte en fase y en magnitud (Fig. 2). La ecuación de un filtro pasa bajos digital de primer orden se encuentra en la ecuación (6).

³Vukan R. Vuchic, URBAN TRANSIT SYSTEMS AND TECHNOLOGY , 1 ED Pag 122.

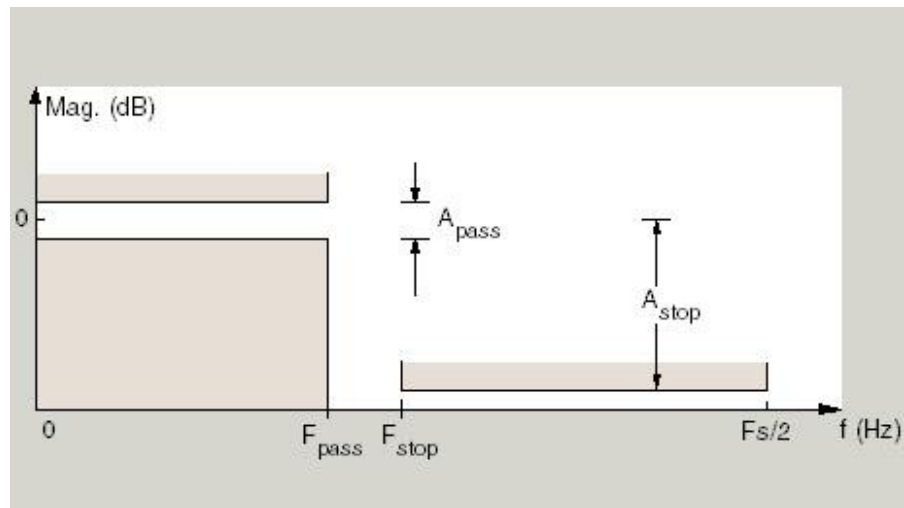


Fig. 2 Especificación de un filtro pasó bajo.⁴

$$y[n] = \frac{x[n]+x[n-1]}{A} \quad (6)$$

Donde A ha de ser mayor que uno. Los filtros pasó bajo, en especial los filtros de media móvil o los filtros de Savitzky-Golay, se utilizan a menudo para limpiar las señales, eliminar el ruido, calcular los promedios de las señales, diseñar diezmadores e interpoladores y descubrir patrones importantes.

Otros métodos habituales de diseño de filtros paso bajo basados en FIR son la ventana de Kaiser, los mínimos cuadrados y el de igual rizado. Entre los métodos de diseño de filtros basados en IIR se cuentan el de Butterworth, el de Chebyshev (tipo I y tipo II) y el elíptico.⁴ Los filtros mencionados anteriormente se encuentran en la herramienta MATLAB®.

2.3. Clasificación no supervisada.

Las clasificaciones no supervisadas son aquellas en las que el algoritmo clasificador no necesita de más información que la escena a clasificar y algunos parámetros que limiten el número de clases. Estos mecanismos de clasificación basan su efecto en la búsqueda de clases con suficiente separabilidad espectral como para conseguir diferenciar unos elementos de otros.

2.4. Algoritmo de agrupamiento (clustering).

Un algoritmo de agrupamiento (de su nombre en inglés, Clustering) es un método de agrupación de una serie de vectores de acuerdo a un criterio. Esos criterios por lo general son distancia o similitud. La cercanía se define en términos de una determinada función de distancia, como la euclidiana, aunque existen otras más robustas.

El resultado final de un proceso de clustering suele ser un dendrograma (Fig. 3) en el que puede verse como los diversos individuos se aglutinan en clases, primero los que están a una menor distancia (los más parecidos), y como posteriormente las clases se unen entre sí. A partir de un dendrograma podemos elegir el número de clases que queremos mantener en función de diferentes criterios [7]

⁴ Tomado de: <https://es.mathworks.com/discovery/filtro-paso-bajo.html>, el día 15 de noviembre de 2016

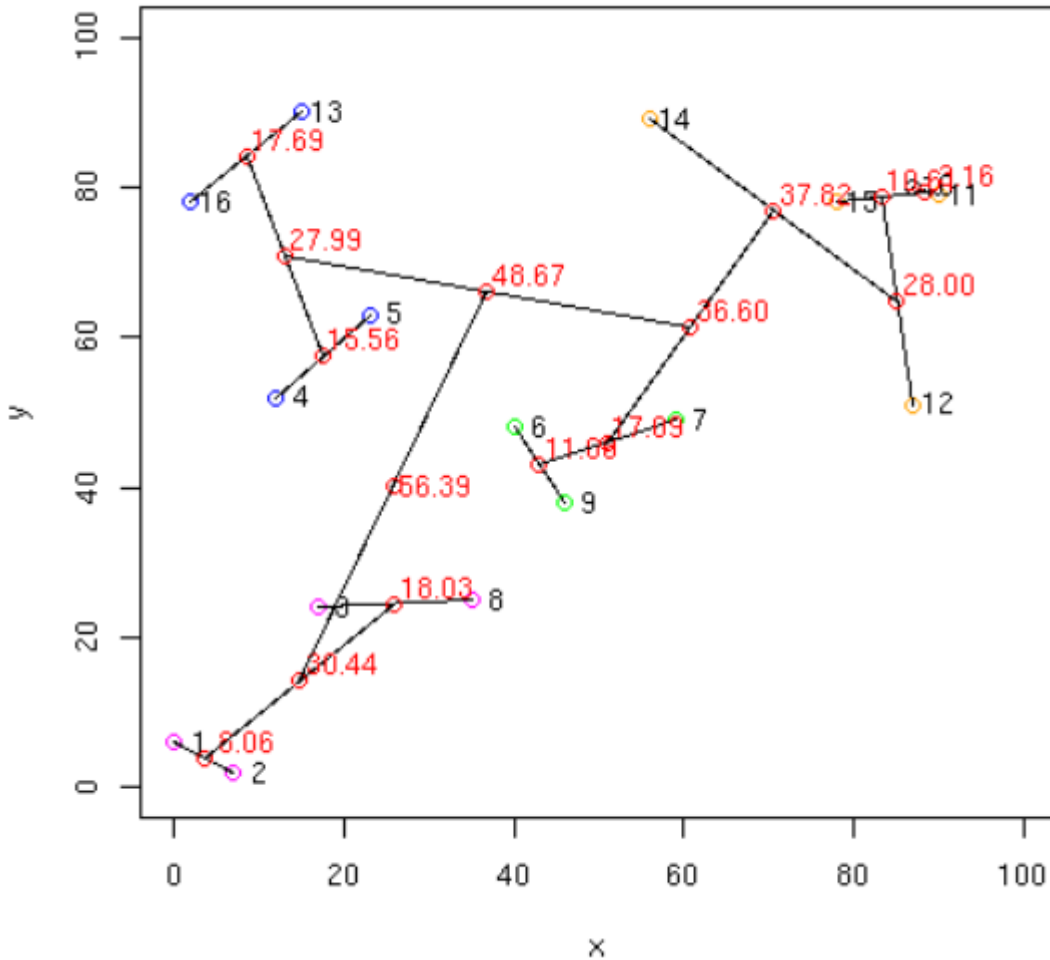


Fig. 3. Dendrograma resultante de clasificación no supervisada [7]

2.4.1.K-means:

Es un método de agrupamiento que tiene como objetivo la partición de un conjunto de n características en k grupos, en donde cada característica pertenece al grupo cuyo valor medio es más cercano. Es comúnmente utilizada en minería de datos.

El algoritmo escoge los centroides de la siguiente manera, suponiendo que el número de grupos es k .

1. Selecciona una observación uniforme al azar del conjunto de datos, X . La observación elegida es el primer centroeide y se denota con c_1 .
2. Calcula las distancias de cada observación c_1 . Denota la distancia entre c_j y la observación m como $d(x_m, c_j)$
3. Selecciona el siguiente centroeide, c_2 al azar de X con probabilidad

$$\frac{d^2(x_m, c_1)}{\sum_{j=1}^n d^2(x_j, c_1)} \quad (7)$$

4. Para elegir el centro j :
 - a. Calcula las distancias de observación a cada centroeide, y asigna a cada observación de su centroeide más cercano.
 - b. Para $m=1, \dots, n$ y $p=1, \dots, j-1$, selecciona el centroeide j aleatoria, con una probabilidad de X .

$$\frac{d^2(x_m, c_p)}{\sum_{\{h_i | x_{h_i} \in C_p\}} d^2(x_{h_i}, c_p)} \quad (8)$$

Donde C_p es el conjunto de todas las observaciones más próximas al centroide c_p y x_m que pertenece a C_p .

Es decir, selecciona cada centro posterior con una probabilidad proporcional a la distancia del mismo al centro más cercano que ya eligió.

5. Repita el paso 4 hasta encontrar k centroides

2.5. Análisis de Componentes Principales (PCA).

El análisis de componentes principales es una técnica utilizada para reducir la dimensionalidad de un conjunto de datos, esta técnica busca la proyección según la cual los datos queden mejor representados en términos de mínimos cuadrados. Esta convierte un conjunto de observaciones de variables posiblemente correlacionadas en un conjunto de valores de variables sin correlación lineal llamadas componentes principales.

El PCA se emplea principalmente en análisis exploratorio de datos o para construir modelos predictivos. Este comporta el cálculo de la descomposición en auto valores de la matriz de covarianza, normalmente tras centrar los datos en la media de cada atributo.

2.6. Acelerómetro

Es un dispositivo que mide la aceleración y la fuerza inducida por la gravedad. Nos permite detectar el movimiento y el giro. Se construye uniendo una masa a un dinamómetro cuyo eje está en la misma dirección que la aceleración que se desea medir. De acuerdo con la Ley Fundamental de la Dinámica o Segunda Ley de Newton.

$$F = m \cdot a \quad (9)$$

Son comúnmente usados en teléfonos de telefonía celular, cámaras digitales, consolas de video juegos entre otros.

2.7. iPhone 6

Es un teléfono inteligente de gama alta desarrollado por Apple. Es un dispositivo que combina una pantalla de 4,7 pulgadas, un teléfono móvil y diversas prestaciones. Su sistema operativo es 10, para obtener los datos del acelerómetro hace uso de la clase UIAccelerometer de la API. Esta clase es la responsable de registrar los movimientos en los ejes x, y, x que se producen en nuestro dispositivo iOS.

El sensor giroscopio (Fig. 4) fue introducido a partir del iPhone 4 e iPad 2. Mediante este se mejora considerablemente los resultados que nos ofrece el acelerómetro y, por tanto, incrementa el nivel de precisión en el posicionamiento. También el giroscopio añade un nivel más incorporando la detección de nuestros movimientos (los que hacemos con el dispositivo en la mano).

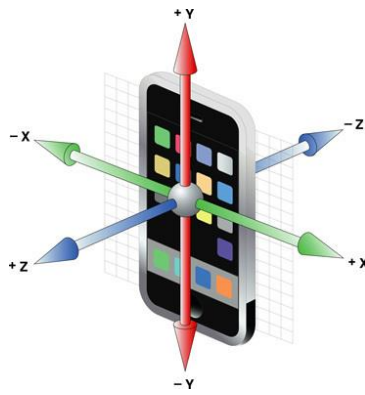


Fig. 4 Coordenadas acelerómetro iPhone⁵

2.8. Sensor Log:

El software lee datos de los sensores de un dispositivo iOS y enviarlos por correo electrónico como archivo csv. Los archivos CSV (del inglés comma-separated values) son un tipo de documento en formato abierto sencillo para representar datos en forma de tabla, en las que las columnas se separan por punto y coma y las filas por saltos de línea.

⁵ <http://www.genbetadev.com/desarrollo-aplicaciones-moviles/programando-en-ios-aceleracion-y-localizacion-nos-ponemos-en-movimiento>

3. OBJETIVO DEL PROYECTO

El objetivo planteado al inicio de este proyecto era:

Investigar dinámicas del movimiento en vehículos de transporte público, en un sector de la localidad de Chapinero de Bogotá, con el uso de acelerómetros y GPS embebidos en teléfonos celulares.

Para cumplir con ese objetivo era preciso cumplir con una serie de objetivos específicos relevantes.

1. Conocer las dinámicas del movimiento en una zona de la localidad de Chapinero.
2. Encontrar características en las señales recogidas, que permitan determinar la calidad de la conducción del vehículo.
3. Analizar y clasificar patrones de dinámicas del movimiento, mediante comparaciones y clasificación no supervisada de situaciones anómalas o repetitivas que vayan en contra de la seguridad de los usuarios dentro del vehículo.

Al final de este proyecto se logró conocer, detectar e identificar dinámicas de movimiento de las muestras realizadas en el trayecto definido en la localidad de Chapinero. Se crearon y se programaron 17 características específicas, las cuales permitieron diferenciar rasgos entre los conductores y se encontraron patrones de conducción existentes, por medio de la reducción de dimensionalidad (PCA) y el agrupamiento (clustering) con el método K-means.

Especificaciones finales del análisis de patrones:

- Construcción de la base de datos por medio del acelerómetro embebido en un teléfono celular.
- Clasificación no supervisada con la definición de diecisiete características específicas.
- Reducción de dimensionalidad por medio de la técnica PCA y posterior a esto algoritmo de agrupamiento K-means.

4. DESARROLLO

El Análisis de patrones de conducción de vehículos se desarrolló en cuatro fases explicadas a continuación.

4.1. Fase uno: Estado de Arte

Se requirió realizar una investigación de estándares, estudios y libros en donde se hablara de dinámicas de movimiento del tránsito urbano, seguridad al interior de medios de transporte y utilización de celulares en aplicaciones parecidas a la nuestra.

Varios estudios de seguridad internacionales al interior de medios de transporte público han analizado las dinámicas de movimiento, encontrando valores límites sugeridos y en algunos casos estandarizándolos como normas locales para la aceleración máxima. Tres ejemplos son:

1. En las ciudades de Valencia y Zaragoza, España se realizaron estudios sobre la aceleración y desaceleración de los buses para determinar la relación que tiene el número de accidentes registrados en esas ciudades con la forma en que conducen los buses. Establecieron que la aceleración máxima que resiste un usuario antes de ser susceptible a caerse es de 0.15G [3].
2. La Universidad del Estado de Pennsylvania realizó un estudio de las aceleraciones y los jalones que perciben los usuarios del transporte público. Establecieron que la aceleración máxima que debe sufrir un usuario que está de pie en el vehículo es de 0.09G y valor máximo de jalón de 0.06 G/s[4].
3. En el libro URBAN TRANSIT SYSTEMS AND TECHNOLOGY del autor Vukan R. Vuchic, enuncian como parámetros máximos de aceleración un rango de 1.0 a 1.8 m/s² en aceleración positiva y en momentos de frenados de emergencia un rango de 2.0 a 3.0 m/s² [6].

Estos resultados sirvieron como base en nuestro análisis definiendo los límites máximos de aceleración y desaceleración.

Existen estudios que avalan el uso de los acelerómetros y receptor GPS incluidos en celulares para la medición de la dinámica del movimiento de vehículos. Para mejorar la precisión de las mediciones, en varios estudios se hace énfasis en el ruido característico de cada equipo celular [8] [9] [10] [5].

4.2. Fase dos: Construcción de la base de datos.

La población de buses que fue escogida para realizar las muestras para el análisis de patrones de conducción de vehículos, son los buses de Transmilenio híbridos (Fig. 5) que transitan sobre la carrera séptima en la localidad de chapinero. La ruta que se analizó fue desde la estación *Calle 45* ubicada en la calle 45 con carrera 7, hasta la estación *Museo Nacional* ubicada en la calle 32 con carrera 7, en ambas direcciones, con una distancia total de 1,6 Km por trayecto (Fig. 6).



Fig. 5. Buses del estudio.⁶

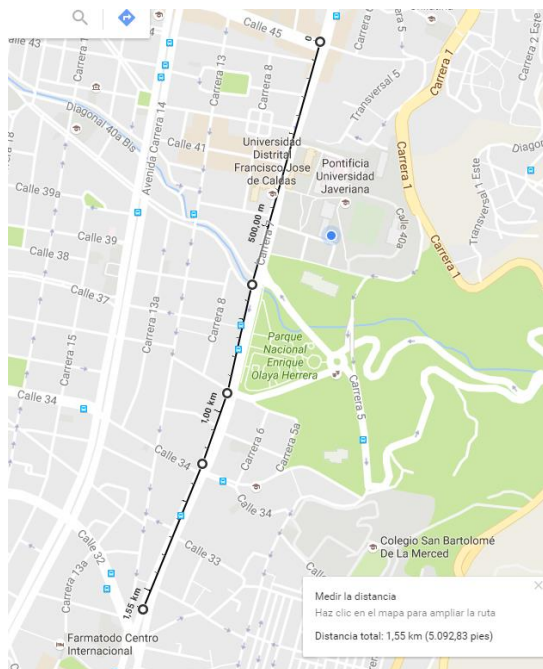


Fig. 6 Ruta del estudio.⁷

Fue necesario diseñar y construir un objeto que garantizara la estabilidad del celular dentro del vehículo. Este anclaje se realizó mediante la caja donde venía el teléfono celular y se le añadió cinta para que el dispositivo no tuviera ningún tipo de movimiento extra debido a la inercia sentida dentro de un vehículo en movimiento. Esta caja se colocó en la parte trasera de cada uno de los buses de la siguiente manera.

⁶ Tomado de: <http://www.portafolio.co/economia/finanzas/superpolo-alista-primeros-buses-hibridos-85998>, el día 23 de noviembre de 2016

⁷ Tomado de: <https://www.google.com.co/maps/@4.6247022,-74.0618294,16.08z?hl=es>, el día 23 de noviembre de 2016

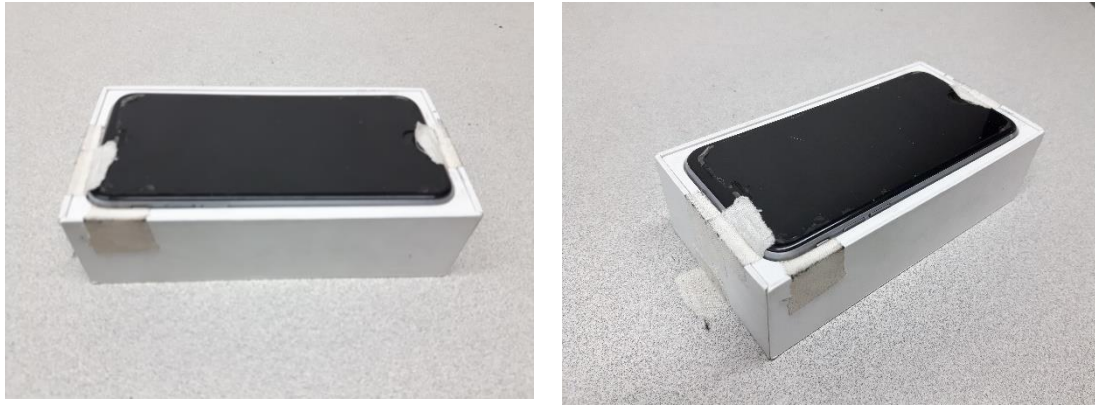


Fig. 7. Anclaje al bus teléfono celular.

El celular utilizado para el análisis fue un iPhone 6, el cual cuenta con dos acelerómetros independientes, el primero Bosch BMA280 de tres ejes y el segundo un MPU-6700 de seis ejes de InvenSense, las características principales de los dos sensores se muestran en la Fig. 8.

Comparison of Bosch BMA280 and InvenSense MPU-6500 Specifications			
Parameter	Bosch BMA280	InvenSense MPU-6500	Units
ADC range	14	16	bit
Acceleration range	$\pm 2, \pm 4, \pm 8, \pm 16$	$\pm 2, \pm 4, \pm 8, \pm 16$	g
Maximum sensitivity	4096	16684	LSB/g
Temperature sensitivity	0.015	0.026	%/K
Cross axis sensitivity	1	2	%
Nonlinearity	0.5	0.5	%
Output data rate	2000	4000	Hz
Cold start up time	3	30	ms
Sleep mode start up time	1	20	ms
Accelerometer start up time from sleep mode	1.3	20	ms
Accelerometer supply current in normal mode	130	450	μA
Accelerometer supply current in low power mode 1	6.5	7.27	μA
Accelerometer supply current in low power mode 2	66	18.65	μA
Sleep mode current	2.1	6	μA

Fig. 8. Comparación de acelerómetros incluidos en iPhone 6.⁸

Para la adquisición de la información necesaria para este proyecto se utilizó la aplicación SensorLog® [11], disponible en el App Store de Apple. Esta aplicación recopiló la medición de las siguientes variables.

- Aceleración en el eje X: Se interpretará como posibles cambios de carril o cambios de dirección de izquierda a derecha o viceversa.
- Aceleración en el eje Y: Se interpretará como las aceleraciones o desaceleraciones que se realicen en línea recta.
- Aceleración en el eje Z: Se interpretará como posibles baches o saltos que encuentre el vehículo en el trayecto.
- Velocidad.

⁸ <http://www.macrumors.com/2014/09/26/iphone-6-6-plus-two-accelerometers/>

Para ilustrar lo anterior, la siguiente imagen muestra el sistema de coordenadas referenciadas (Fig. 9).

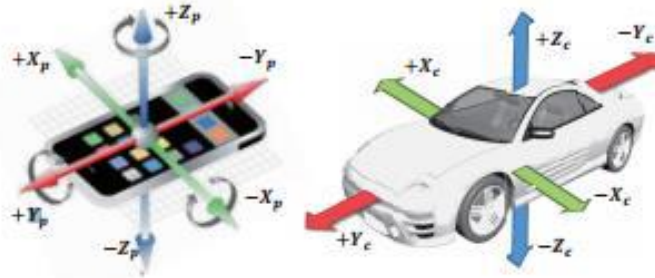


Fig. 9. Sistema de coordenadas

Esta app permite configurar el muestreo según cada aplicación que el usuario necesite. Al inicio del análisis se tomaron muestras con una tasa de 30 Hz, obteniendo archivos extensos. Se decidió que para las siguientes muestras configurar la app con una tasa de muestreo de 15 Hz.

Cada una de las muestras obtenidas, aparte de la información correspondiente a cada variable, contiene la siguiente información adicional.

1. Fecha del día que se tomaron los datos.
2. Hora de iniciación y hora de finalización del trayecto.
3. Número de la ruta seleccionada.
4. Sentido de la ruta. “Javeriana-Museo Nacional” si fue sentido norte-sur o “Museo Nacional” si fue sentido sur-norte.

Cada muestra corresponde a un trayecto y está compuesta por dos archivos digitales, el primero en formato de hoja de cálculo que contiene todas las mediciones de las variables previamente descritas y la información adicional y el segundo corresponde a una grabación de voz en formato mp3. En este archivo de voz se nombró cada uno de los eventos que sucedieron durante el trayecto, con la siguiente nomenclatura:

- Aceleración: Cada vez que se sentía o se escuchaba el motor del bus acelerar.
- Giro Izquierda: Cada vez que el bus cambiaba de carril hacia la izquierda.
- Giro Derecha: Cada vez que el bus cambiaba de carril hacia la derecha.
- Freno o Freno Normal: Cada vez que el bus frenaba de manera normal.
- Freno Medio: Cada vez que el bus frenaba de manera más brusca.
- Freno Duro: Cada vez que el bus frenaba de manera extremadamente brusca.
- Hueco o Bache: Cada vez que el bus pasaba por un hueco o bache muy evidente y notorio.
- Detención: cada vez que el bus se detenía por completo.
- Cambio: Cada vez que se escuchaba o se sentía un aumento de cambio.

Esta nomenclatura nos ayudó a comprender y a verificar las dinámicas de movimiento que tienen los buses las cuales serán explicadas más adelante.

Los archivos correspondientes a cada trayecto se nombraron de la misma manera para ambos casos (hoja de cálculo y formato mp3) de la siguiente manera. Un prefijo “g” seguido de una numeración consecutiva para cada muestra, es decir *g001*, *g002*, *g003* y así sucesivamente.

Se realizaron 72 muestras, las cuales 17 sirvieron para el entrenamiento del software y 50 fueron las utilizadas para el análisis de resultados y 5 fueron descartadas. En el Anexo 1 se puede observar un ejemplo del formato establecido para un archivo en hoja de cálculo.

4.2.1. Experimento

La toma de muestras se realizó bajo las siguientes condiciones.

- Los horarios en los que se tomaron las muestras correspondieron a los horarios en los que no se tenía clase en la Universidad, las horas fueron entre 9:00 am - 1:00 pm y 2:00 pm – 4:00 pm.
- Las muestras se obtuvieron de lunes a viernes.
- El anclaje del celular al bus se realizó en la parte trasera del bus, esta ubicación dispone de una superficie plana encima del motor del vehículo. Aunque existe mayor vibración que en otras partes es el lugar más estable para ubicar el celular.
- El inicio de la grabación en la aplicación Sensor Log y el inicio de la grabación de voz se hizo al mismo tiempo. Esto con el fin de tener una sincronía entre los eventos descritos por la voz y los eventos visualizados por el algoritmo. Se realizaron algunas grabaciones sin sincronismo inicial pero
- No se tuvo en cuenta variables como temperatura, flujo de vehículos y cantidad de personas en el bus.
- Las condiciones iniciales y finales del vehículo (detenido o en movimiento) no determinaban el inicio o final de la grabación.

Para la obtención de las muestras dentro de cada uno de los buses no se pidió ningún permiso, por lo que las grabaciones tenían que ser realizadas con discreción para no llegar a tener problemas con las personas encargadas del vehículo.

4.3. Fase tres: Acondicionamiento y clasificación de los patrones obtenidos.

En esta fase se enunciarán los procesos previos al reconocimiento de patrones de conducción, que son el acondicionamiento de las muestras, clasificación de eventos y clasificación no supervisada.

4.3.1. Acondicionamiento de las muestras.

Antes de realizar cualquier tipo de clasificación o análisis de las muestras estudiadas fue necesario diseñar un software en MATLAB, que facilitara la visualización y el reconocimiento de las dinámicas de movimiento presentes en cada trayecto de un bus.

El primer paso fue definir el sistema de coordenadas que se utilizó en todo el análisis y verificar que cada una de las muestras respetara esta referencia estipulada. El sistema de coordenadas utilizado se describirá a continuación:

- Todas aquellas aceleraciones en el eje Y positivas corresponden al aumento de velocidad respecto al tiempo, se graficaron en el cuadrante positivo.
- Todas aquellas aceleraciones en el eje Y negativas corresponden a la disminución de la velocidad respecto al tiempo, se graficaron en el cuadrante negativo.
- Todas aquellas aceleraciones en el eje X positivas corresponden a los giros a la derecha, se graficaron en el cuadrante positivo.
- Todas aquellas aceleraciones en el eje X negativas corresponden a los giros a la izquierda, se graficaron en el cuadrante negativo.
- La aceleración en el eje Z se graficó en el eje Z. corresponde a los baches encontrados durante el trayecto.

Para ilustrar lo anterior, se muestra un ejemplo de una de las muestras obtenidas en la Fig. 10.

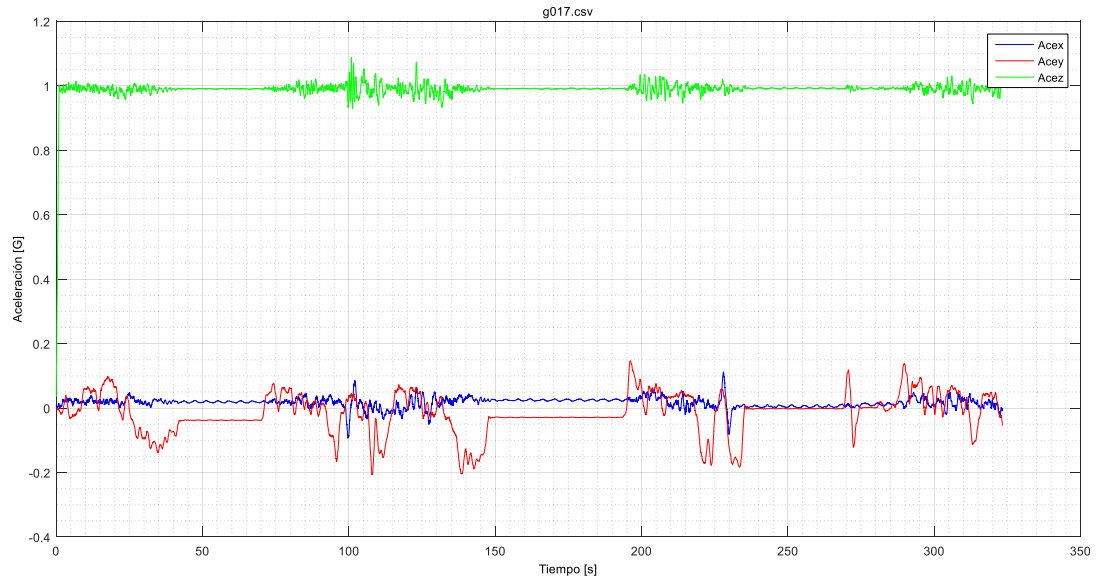


Fig. 10. Traza de aceleraciones de un bus en los ejes X, Y y Z

Todas las pendientes positivas en la traza roja (eje Y) corresponden a la aceleración positiva que tuvo el bus, las pendientes negativas significan frenadas. En la traza azul (eje x) las pendientes positivas corresponden a cambios de carril a la derecha y las pendientes negativas cambios de carril a la izquierda. La traza verde es la aceleración en el eje Z y su referencia es 1 G.

Como se puede ver en la Fig. 10 las unidades que se van a trabajar durante todo el proyecto son Gravedades [G] para las aceleraciones y Kilómetros por hora [Km/h] para la velocidad.

El siguiente paso que se realizó fue filtrar todos los datos con el fin de eliminar el ruido existente. En este tipo de mediciones se encontraron dos tipos de ruido, el primero es el ruido del sensor y el segundo es el ruido debido a las vibraciones del vehículo.

Los datos recolectados se pasaron por dos filtros. El primer filtro fue un Moving_Average Filter, este se utilizó con el fin de suavizar la traza y poder ver mejor los eventos existentes. La ecuación que define este filtro es la siguiente.

$$y(n) = \frac{1}{\text{TamanoVentana}}(x(n) + x(n - 1) + \dots + x(n - (\text{TamanoVentana} - 1))). \quad (10)$$

El tamaño de la ventana que se utilizó fue el mismo al de las muestras por segundo según la grabación, con el fin de promediar las mismas muestras/segundo que el sensor medía.

El segundo filtro se diseñó de tal manera que eliminara el ruido presente por la vibración del vehículo y el ruido característico del sensor. Para calcular la frecuencia de corte de este filtro se estimó la densidad de potencia espectral (PSD), con el fin de ver a qué frecuencia estaban las componentes con mayor PSD y filtrar aquellas que no eran parte de este análisis.

A continuación se muestra el periodograma usando la transformada de Fourier de una de las muestras de prueba existente.

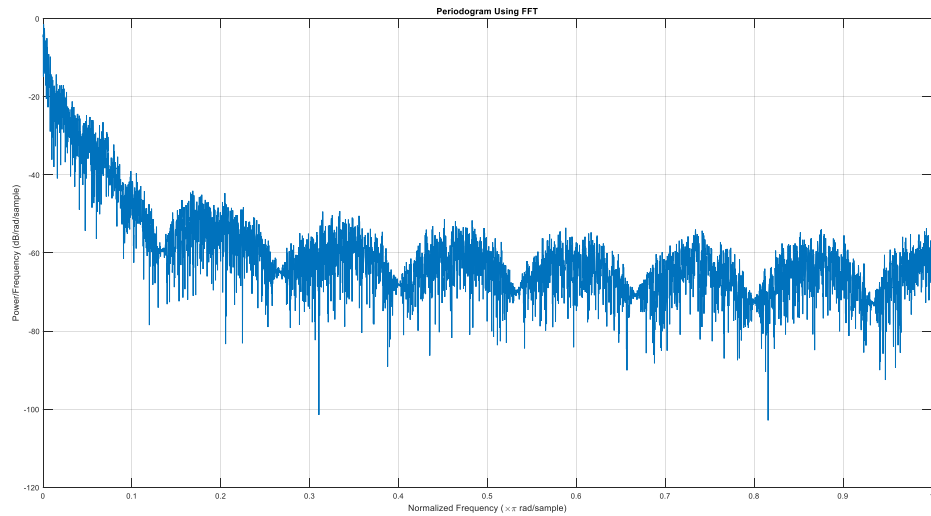


Fig. 11. Periodograma PSD de una traza de prueba.

Se determinó que la frecuencia de corte es de:

$$\omega_n = 0.03 \text{ rad/sample} \quad (11)$$

Este valor se escogió para filtrar las componentes de potencia espectral que superaron los -40 dB y -50 dB según la muestra.

El orden del filtro, como su definición lo explica, será la pendiente de caída que tiene el diagrama de bode en magnitud. Para esta aplicación se escogió un orden de 50.

En la figura 15 se muestran el proceso de filtrado que tuvo cada muestra de la base de datos.

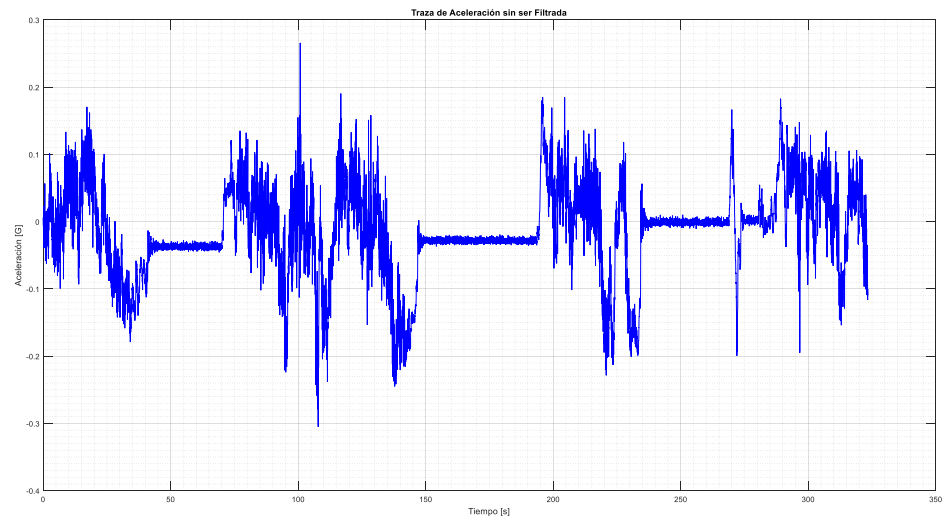


Fig. 12. Traza de aceleración original.

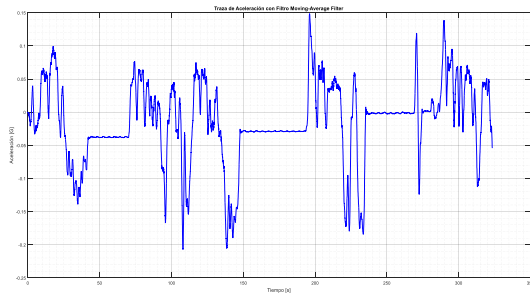


Fig. 13 traza de aceleración con filtrado de promedio

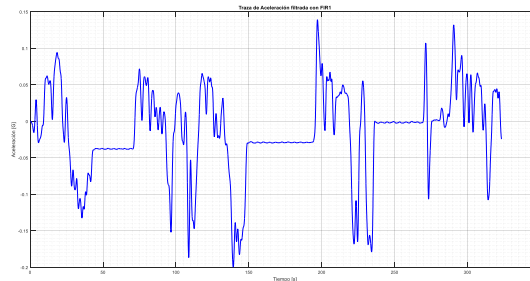


Fig. 14 Traza de aceleración filtrada.

Fig. 15. Acondicionamiento de una muestra por medio del Filtrado

En el Anexo 2 se puede observar el código utilizado para el procesamiento realizado.

Posterior al filtrado de cada señal se continuó con el análisis de las dinámicas de movimiento, verificando las trazas registradas con las grabaciones de voz correspondientes a cada una, esto con el fin de conocer los eventos que aparecían en las gráficas, para definir las características de la clasificación no supervisada.

4.3.2. Clasificación de eventos.

Los eventos que se utilizaron en el análisis de patrones de conducción son los siguientes:

4.3.2.1. Aceleraciones

En la Fig. 16 se muestran encerrados en los círculos azules, los eventos de aceleración que se detectaron en el trayecto, en este caso se seleccionaron las aceleraciones mayores a 0.09G, los puntos azules señalados son cambios de velocidad positiva respecto al tiempo menores a 0.09G.

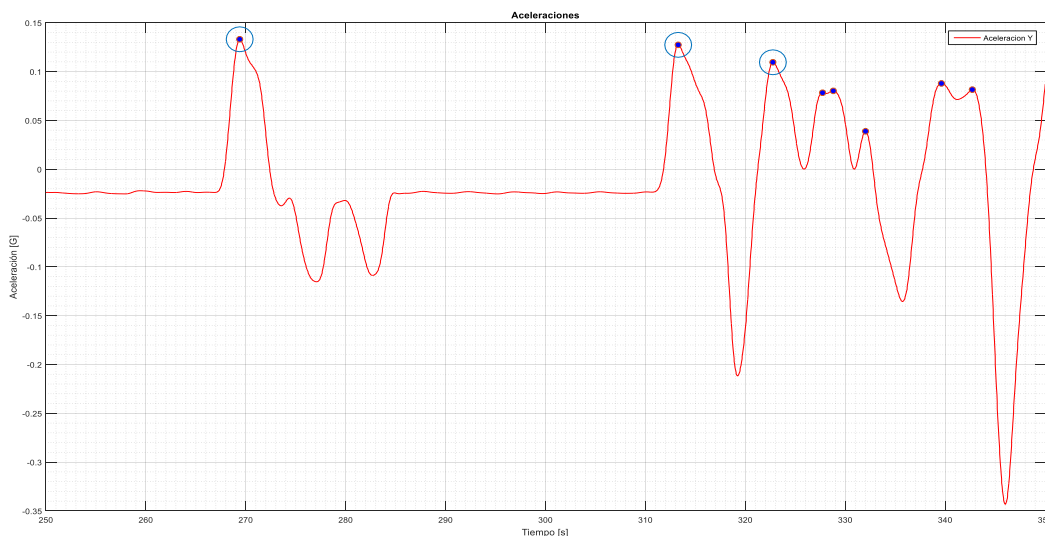


Fig. 16 Traza de aceleración (eje Y)- Aceleraciones.

4.3.2.2. Frenadas

En la Fig. 17 se pueden observar encerrados en los círculos azules, eventos de frenado que sucedieron en el trayecto seleccionado. En este tipo de eventos se encontraron mayores diferencias en la magnitud de los picos, como se puede evidenciar en la Fig 17., que corresponde a las diferentes maneras en que un conductor frena según cada situación durante el trayecto.

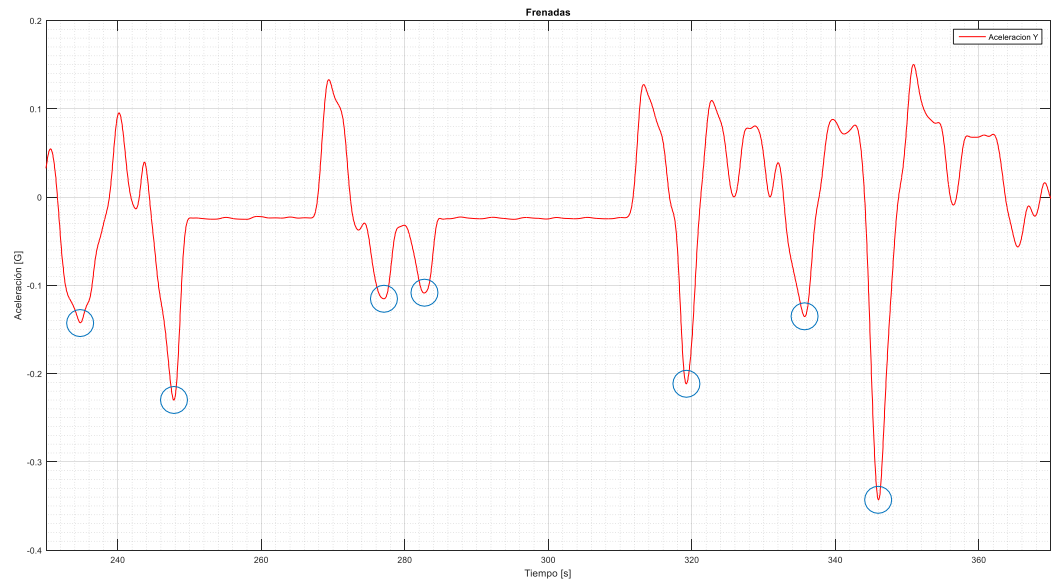


Fig. 17. Trazo de aceleración (eje Y)-Desaceleraciones.

4.3.2.3. Cambio de carril a Izquierda y a Derecha

Como los trayectos estudiados no tienen giros de 90 grados, solo se analizó los cambios de carril que tuvieron los buses en cada muestra. En la Fig. 18 se muestra un cambio de carril a izquierda y otra a la derecha. Estos eventos se ven reflejados en la traza de la Aceleración en el eje X como se explicó en la definición de variables.

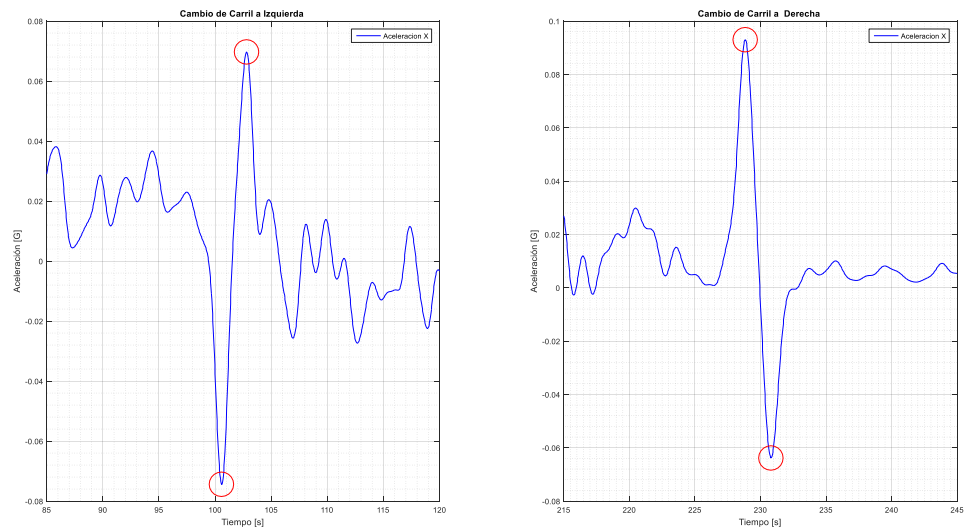


Fig. 18 Aceleración en el eje X- Giros.

4.3.2.4. Baches o Huecos

En la Fig. 19 se observa como es una traza de la aceleración en el eje Z. Esta es la variable que se ve más afectada por las vibraciones dentro del bus y sobre todo por las irregularidades que posee la carrera séptima. Entre mayor sea el pico más grande fue la irregularidad por la que se pasó.

Existen otros dos tipos de gráficas que se analizaron las cuales serán explicadas más adelante en la definición de las características para la clasificación no supervisada, las cuales son la velocidad y el jerk o jalón.

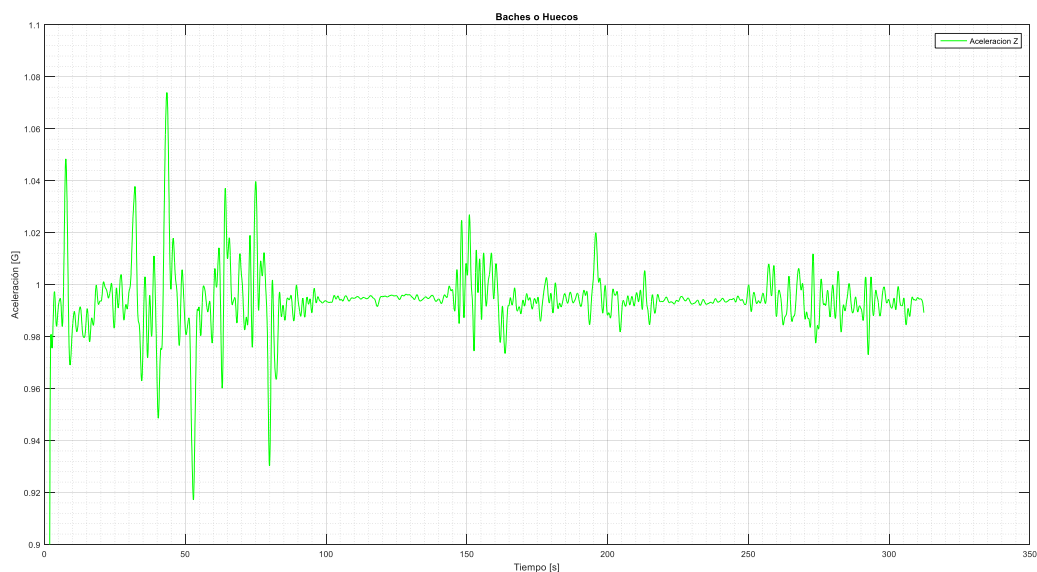


Fig. 19 Aceleración en el eje Z

4.3.3. Clasificación no supervisada

Luego de comprender y entender los eventos que son importantes para el análisis, se definieron 17 características que nos permitieron comenzar a clasificar las 50 muestras analizadas. Estas características tuvieron como fundamento los eventos y las variables que se han venido definiendo a lo largo del texto y tres rasgos importantes: el tiempo de duración del evento, la frecuencia de repetición del evento y el factor de sobrepaso de un límite establecido en cualquier tipo de variable.

Los límites que se definieron se estipularon según el riesgo que pueden llegar a generar en un pasajero y referenciándonos en los artículos estudiados. Estos fueron los límites con los que se realizó el análisis.

- 0.09 [G] Para la Aceleración en Y y en X. Riesgo medio-alto [4].
- 0.15 [G] Para la Aceleración en Y y en X. Riesgo muy alto [3].
- 40 [Km/h] Para la velocidad. Riesgo medio [12].
- 0.06 [G/s] Para el jalón. Riesgo Medio-alto [4].

Cada una de las muestras se clasificó entre tres rangos; “Bueno”, “Regular” y “Malo”, en cada una de las características definidas, esto facilitó el análisis y se comenzó a evidenciar los posibles grupos de conductores que tenían formas de conducción similares.

Las características definidas fueron las siguientes:

4.3.3.1. Frecuencia de Picos Que Superan los |0.09G|

Eje: Aceleración Y - Frenadas

Unidades: [# / min]

Descripción: Esta característica define en número de veces que en la muestra el conductor superó los 0.09 G a la hora de frenar sobre el tiempo de duración del trayecto.

Valor Mínimo: 0.

Valor Máximo: Infinito.

Ecuación:

$$C1 = \frac{\text{número de picos}}{\text{Duración de la Grabación}} \quad (12)$$

Rangos de clasificación:

- “Bueno”: [0 - 2]. De 0 a 2 frenadas mayores a 0.09G por min.
- “Regular”: [2 - 4]. De 2 a 4 frenadas mayores a 0.09G por min.
- “Malo”: [4 - Inf]. Más de 4 frenadas mayores a 0.09G por min.

Con el fin de poder ver cómo fue la programación de algunas variables se mostrarán los resultados obtenidos de las variables que fueron posibles graficarlas. En el Anexo 3 se observa el código programado para desarrollar cada una de las características (C1 – C17)

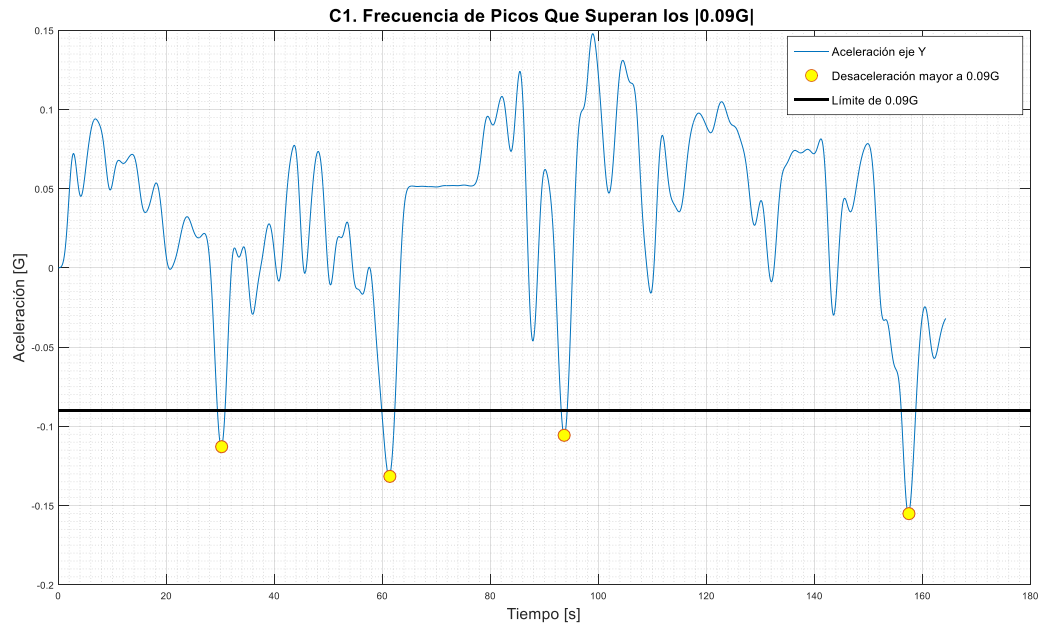


Fig. 20: Frecuencia de Picos que Superan los |0.09G|

4.3.3.2. Promedio Sobrepasso en algún factor los picos mayores a |Referencia mínima⁹

Eje: Aceleración Y - Frenadas

Unidades: [Adimensional]

Descripción: Esta característica define el promedio de los factores de sobrepasso de los picos respecto a 0.09G. Solo se tuvieron en cuenta los picos mayores a la referencia mínima.

Valor Mínimo: Referencia Mínima/0.09 G.

Valor Máximo: Infinito.

Ecuación:

$$C2 = \text{Promedio} \left(\frac{\text{Valor del Pico}}{0.09G} \right) \quad (13)$$

Rangos de clasificación:

- “Bueno”: [0 - 1]. Promedio de frenado menor a 0.09 G.

⁹ Referencia Mínima: valor mínimo que tiene la traza cuando el bus está detenido, si este valor es mayor a -0.05G la referencia mínima es igual a -0.05G.

- “Regular”: [1 – 1.66]. Promedio de frenado entre 0.09 G y 0.15 G.
- “Malo”: [1.66 - Inf]. Promedio de frenado mayor a 0.15 G.

4.3.3.3. Porcentaje de tiempo que dura frenando (Frenos mayores a |0.09G|)

Eje: Aceleración Y - Frenadas

Unidades: [Adimensional]

Descripción: Esta característica define el porcentaje del tiempo total que el conductor realizo frenadas mayores a 0.09 G.

Valor Mínimo: 0 %

Valor Máximo: 100%.

Ecuación:

$$C3 = \frac{\text{tiempo de frenado Final} - \text{tiempo de frenado inicial}}{\text{Duración Total del Trayecto}} * 100 \quad (14)$$

Rangos de clasificación:

- “Bueno”: [0 - 10]. Del 0% al 10% del tiempo frenando.
- “Regular”: [20 – 50]. Del 20% al 50% del tiempo frenando.
- “Malo”: [50 - 100]. Más del 50% del tiempo frenando.

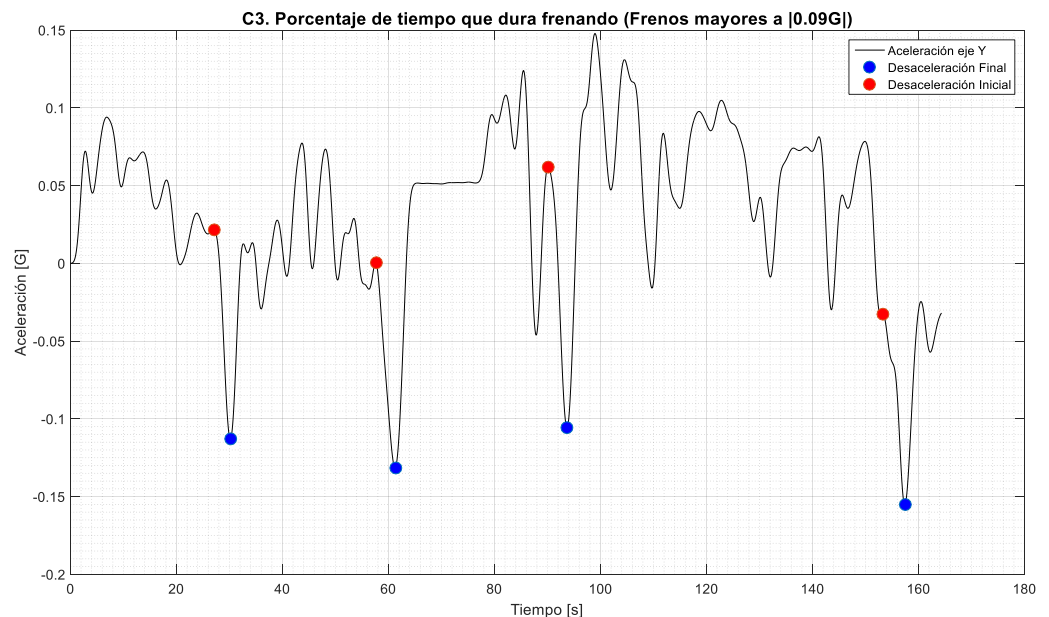


Fig. 21: Porcentaje de tiempo que dura frenando (Frenos mayores a |0.09|)

4.3.3.4. Frecuencia de Picos Que Superan los |0.15G|

Eje: Aceleración Y - Frenadas

Unidades: [# / min]

Descripción: Esta característica define en número de veces que en la muestra el conductor superó los 0.15 G a la hora de frenar sobre el tiempo de duración del trayecto.

Valor Mínimo: 0.

Valor Máximo: Infinito.

Ecuación:

$$C4 = \frac{\text{número de picos}}{\text{Duración de la Grabación}} \quad (15)$$

Rangos de clasificación:

- “Bueno”: [0 – 0.3]. El tiempo entre frenadas es igual o mayor a 3 minutos.
- “Regular”: [0.3 - 1]. El tiempo entre frenadas es menor a 3 y mayor a 1 minuto.
- “Malo”: [1 - Inf]. El tiempo entre frenadas es igual o menor a 1 minuto.

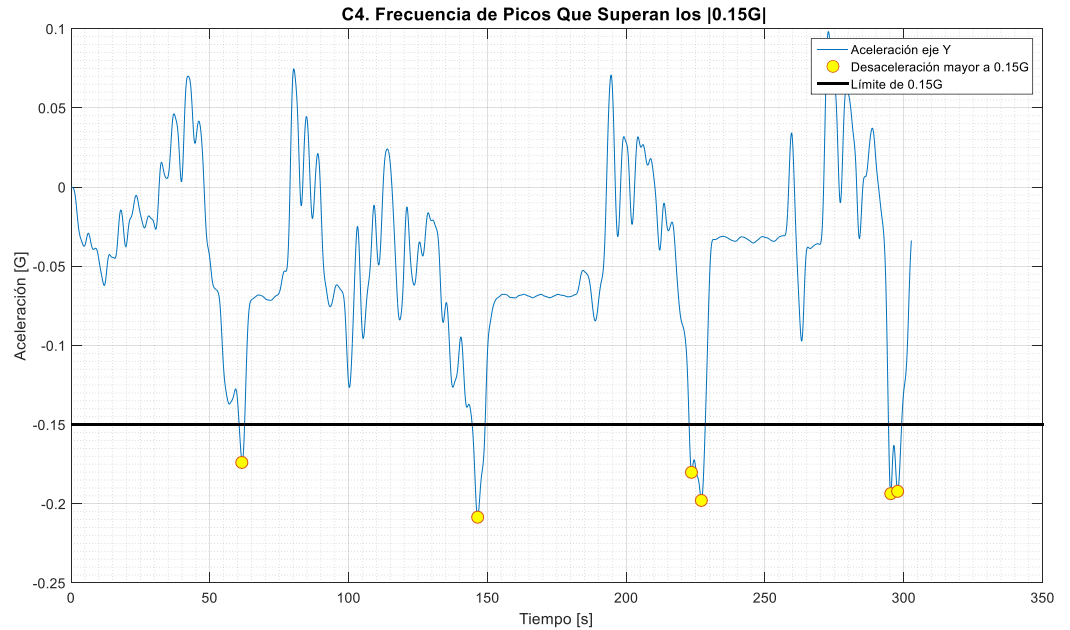


Fig. 22: Frecuencia de Picos que superan los |0.15 G|

4.3.3.5. Frecuencia de picos que superan los 0.06 G/s

Eje: Jalón

Unidades: [Adimensional]

Descripción: Esta característica define el número de veces en que durante el trayecto se supera 0.06G/s en la gráfica del jalón

Valor Mínimo: 0 %

Valor Máximo: 100%.

Ecuación:

$$C5 = \frac{\text{número de picos}}{\text{Duración de la Grabación}} \quad (16)$$

Rangos de clasificación:

- “Bueno”: [0 - 5]. De 0 a 5 picos por minutos.
- “Regular”: [5 – 10]. De 5 a 10 picos por minutos.
- “Malo”: [10 - 100]. Más de 10 picos por minutos.

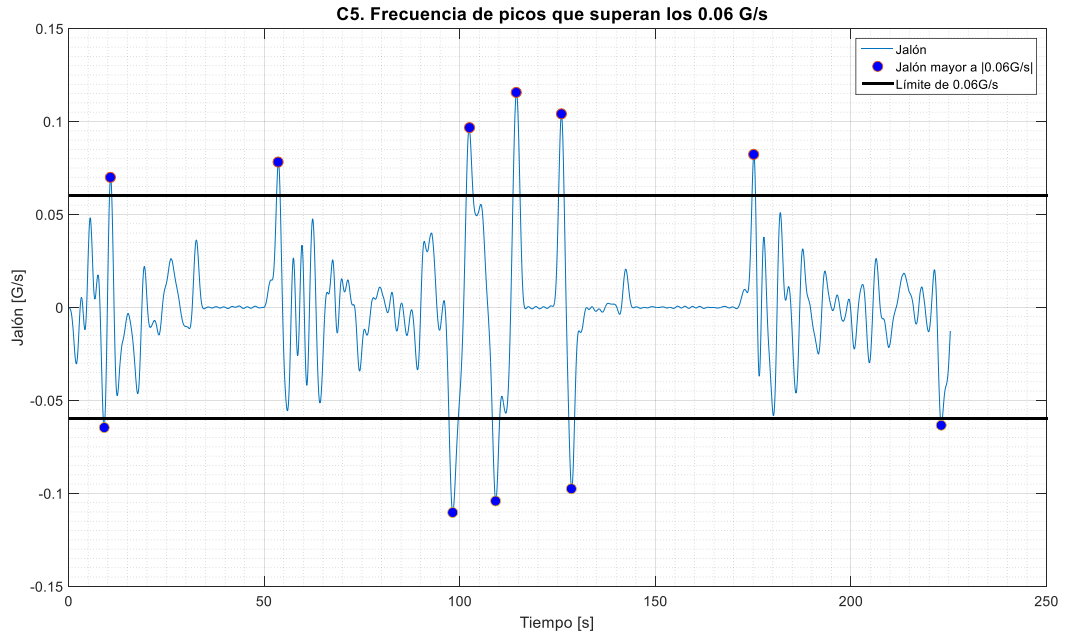


Fig. 23: Frecuencia de Picos que superan los 0.06 G/s

4.3.3.6. Porcentaje de tiempo que dura frenando (Frenos mayores a |0.15G|)

Eje: Aceleración Y - Frenadas

Unidades: [Adimensional]

Descripción: Esta característica define el porcentaje del tiempo total que el conductor realizo frenadas mayores a 0.15 G.

Valor Mínimo: 0 %

Valor Máximo: 100%.

Ecuación:

$$C6 = \frac{\text{Duración Total de Frenado}}{\text{Duración Total del Trayecto}} * 100 \quad (17)$$

Rangos de clasificación:

- “Bueno”: [0 - 5]. Del 0% al 5% del tiempo frenando.
- “Regular”: [5 – 10]. Del 5% al 10% del tiempo frenando.
- “Malo”: [10 - 100]. Más del 10% del tiempo frenando.

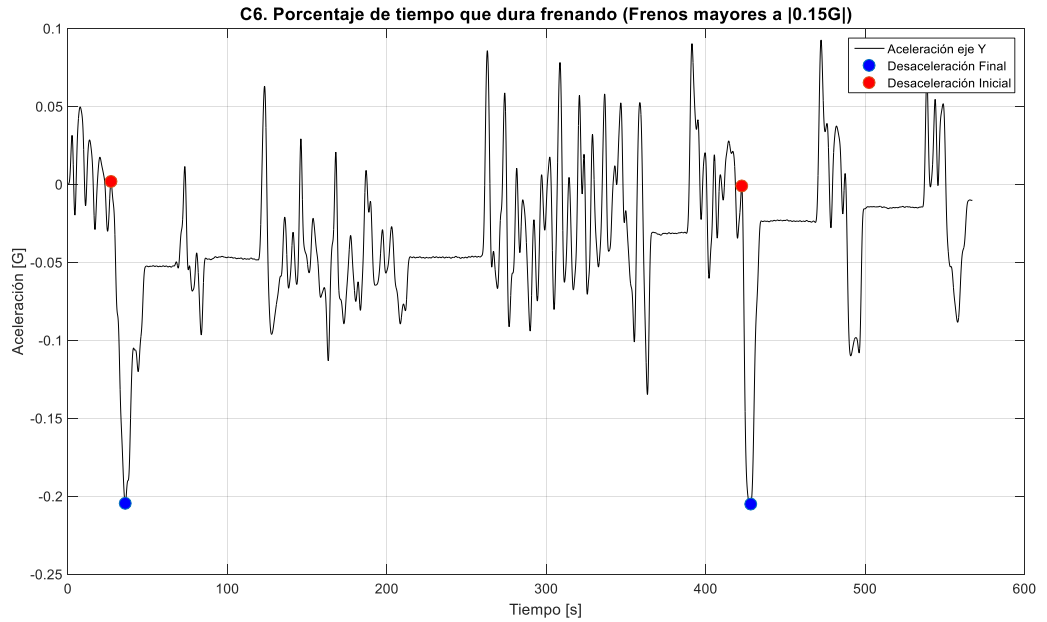


Fig. 24: Porcentaje de tiempo que dura frenando (Frenos mayores a |0.15 G|)

4.3.3.7. Porcentaje de frenadas mayores a |0.09G|

Eje: Aceleración Y - Frenadas

Unidades: [Adimensional]

Descripción: Esta característica define el porcentaje de picos mayores a 0.09 G respecto al total de picos mayores a la referencia mínima.

Valor Mínimo: 0 %

Valor Máximo: 100%.

Ecuación:

$$C7 = \frac{\text{picos mayores a } 0.09G}{\text{Picos mayores a la referencia mínima}} * 100 \quad (18)$$

Rangos de clasificación:

- “Bueno”: [0 - 50]. Menos de la mitad de las frenadas son mayores a 0.09 G.
- “Regular”: [50 – 75]. Entre en 50% y el 75% de las frenadas son mayores a 0.09 G.
- “Malo”: [75 - 100]. Más del 75% de las frenadas son mayores a 0.09G.

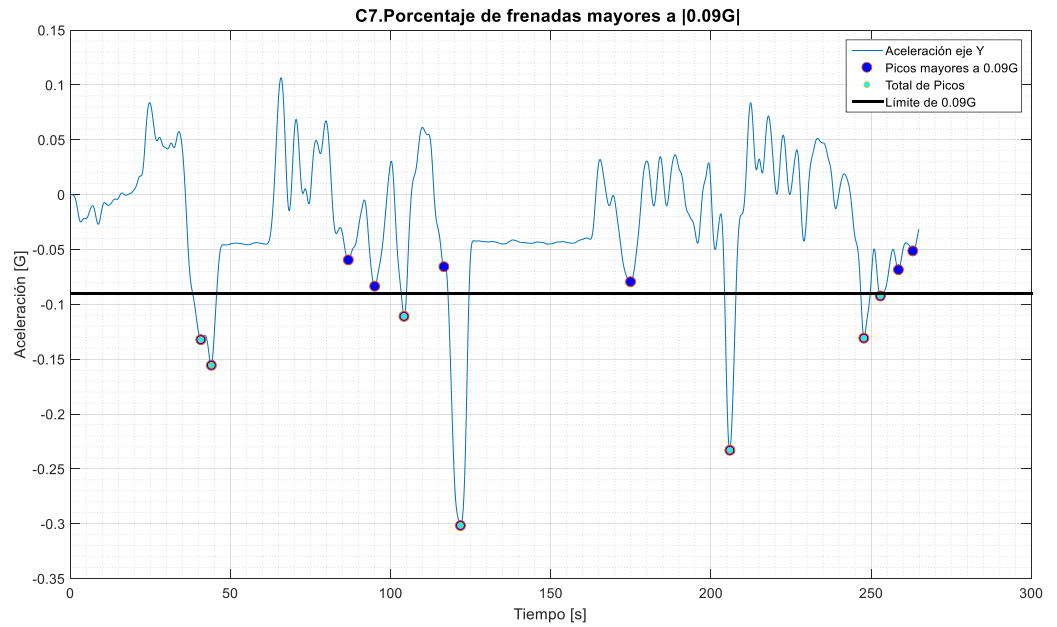


Fig. 25: Porcentaje de frenadas mayores a |0.09 G/

4.3.3.8. Porcentaje de frenadas mayores a |0.15G|

Eje: Aceleración Y - Frenadas

Unidades: [Adimensional]

Descripción: Esta característica define el porcentaje de picos mayores a 0.15 G respecto al total de picos mayores a la referencia mínima.

Valor Mínimo: 0 %

Valor Máximo: 100%.

Ecuación:

$$C8 = \frac{\text{picos mayores a } 0.15G}{\text{Picos mayores a la referencia mínima}} * 100 \quad (19)$$

Rangos de clasificación:

- “Bueno”: [0 - 15]. Menos de la sexta parte de las frenadas son mayores a 0.15 G.
- “Regular”: [15 – 33]. Entre en 15% y el 33% de las frenadas son mayores a 0.15 G.
- “Malo”: [33 - 100]. Más del 33% de las frenadas son mayores a 0.15G.

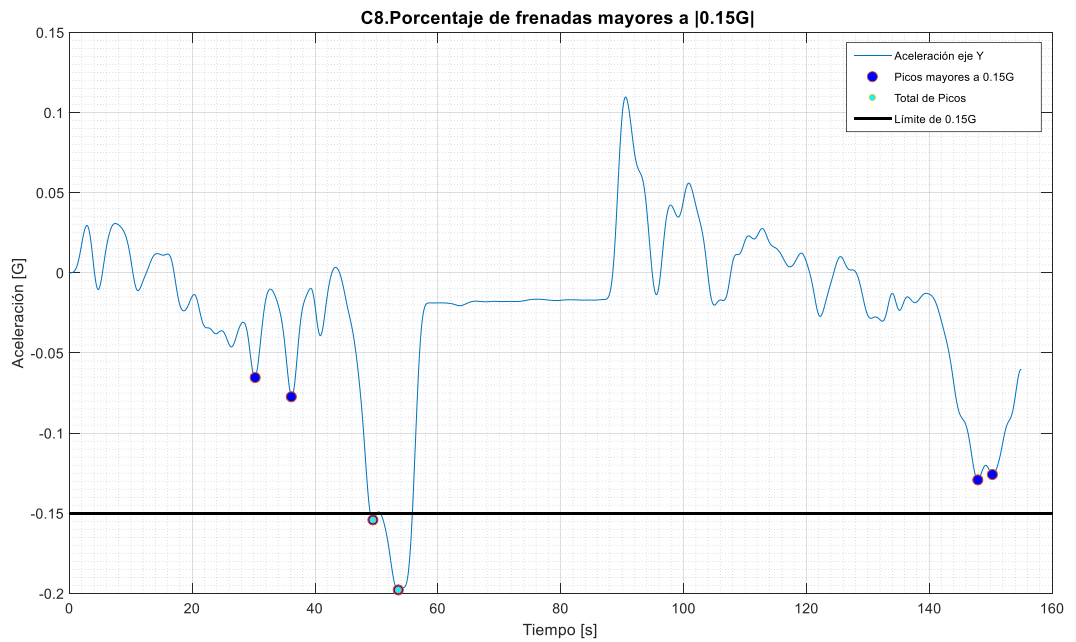


Fig. 26: Porcentaje de frenadas mayores a |0.15G|

4.3.3.9. Promedio Sobre Paso en algún factor los picos mayores a |0.09G|

Eje: Aceleración X - Giros

Unidades: [Adimensional]

Descripción: Esta característica define el promedio de los factores de sobre paso de los picos respecto a 0.09G. Solo se tuvieron en cuenta los picos mayores a 0.09G.

Valor Mínimo: 1.

Valor Máximo: Infinito.

Ecuación:

$$C9 = \text{Promedio} \left(\frac{\text{Valor del Pico}}{0.09G} \right) \quad (20)$$

Rangos de clasificación:

- “Bueno”: [1 – 1.33]. Promedio de aceleración en X entre 0.09 G y 0.12 G.
- “Regular”: [1.33 – 1.66]. Promedio de aceleración en X entre 0.12 G y 0.15 G.
- “Malo”: [1.66 - Inf]. Promedio de aceleración en X mayor a 0.15 G.

4.3.3.10. Promedio Sobrepasso en algún factor los picos mayores al |5% del promedio de la señal|

Eje: Aceleración Z - Baches

Unidades: [Adimensional]

Descripción: Esta característica define el promedio de los factores de sobrepasso de los picos respecto al 5% del promedio de la señal en la aceleración en Z. Solo se tuvieron en cuenta los picos mayores al 5% del promedio de la señal.

Valor Mínimo: 1.

Valor Máximo: Infinito.

Ecuación:

$$C10 = \text{Promedio} \left(\frac{\text{Valor del Pico}}{5\% \text{ del promedio de la señal}} \right) \quad (21)$$

Rangos de clasificación:

- “Bueno”: [1 – 1.2]. Promedio de aceleración en Z hasta un 20% mayor a la aceleración ideal.
- “Regular”: [1.2 – 1.4]. Promedio de aceleración en Z entre el 20% y 40% mayor a la aceleración ideal.
- “Malo”: [1.4 - Inf]. Promedio de aceleración en Z arriba del 40% mayor a la aceleración ideal.

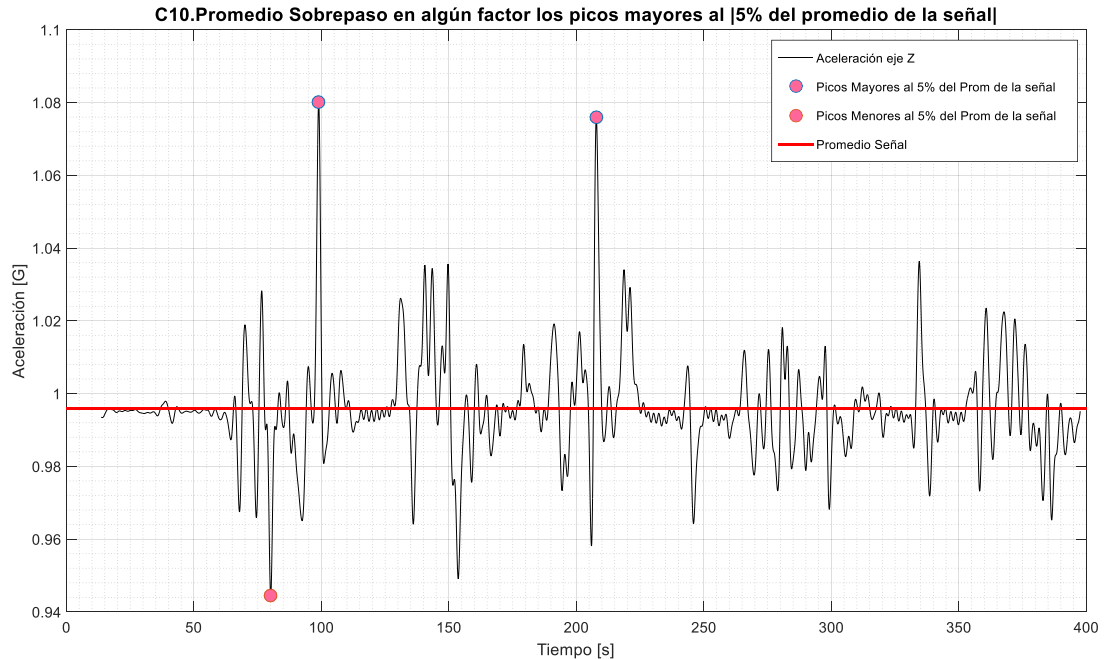


Fig. 27: Promedio de sobrepasso en algún factor los picos mayores al |5% del promedio de la señal|

4.3.3.11. Frecuencia de Picos Que Superan los 0.09G

Eje: Aceleración Y - Aceleraciones

Unidades: [# / min]

Descripción: Esta característica define en número de veces que en la muestra el conductor superó los 0.09 G a la hora de acelerar sobre el tiempo de duración del trayecto.

Valor Mínimo: 0.

Valor Máximo: Infinito.

Ecuación:

$$C11 = \frac{\text{número de picos}}{\text{Duración de la Grabación}} \quad (22)$$

Rangos de clasificación:

- “Bueno”: [0 - 1]. De 0 a 1 acelerada mayor a 0.09G por min.
- “Regular”: [1 - 3]. De 1 a 3 aceleradas mayores a 0.09G por min.
- “Malo”: [3 - Inf]. Más de 3 aceleradas mayores a 0.09G por min.

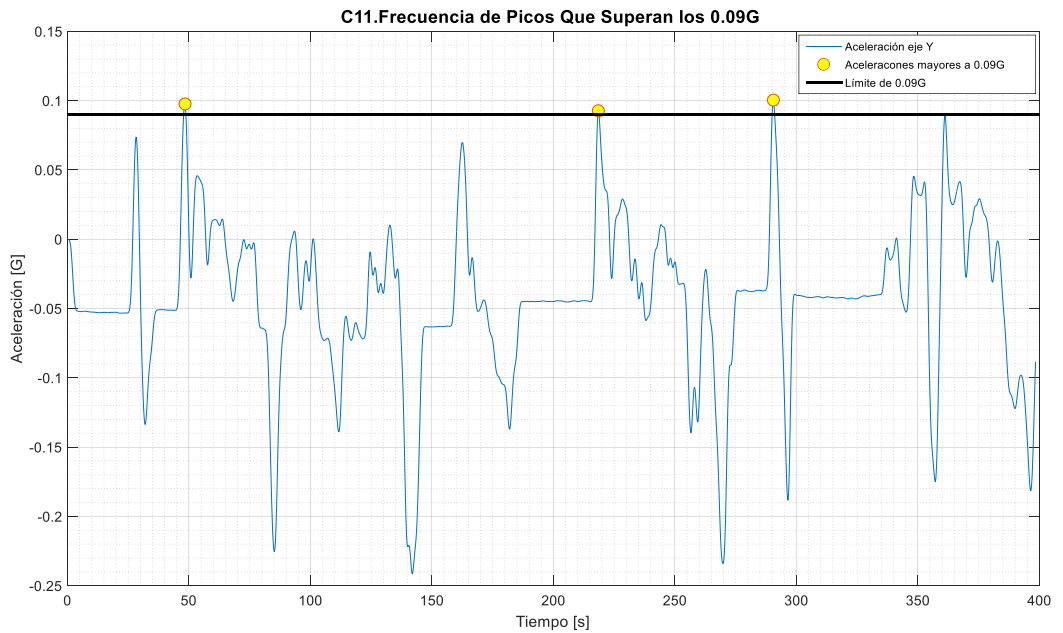


Fig. 28: Frecuencia de picos que superan los 0.09G

4.3.3.12. Promedio Sobre Paso en algún factor los picos mayores a 0.09G

Eje: Aceleración Y - Aceleraciones

Unidades: [Adimensional]

Descripción: Esta característica define el promedio de los factores de sobre paso de los picos respecto a 0.09G. Solo se tuvieron en cuenta las aceleraciones mayores a 0.09 G.

Valor Mínimo: 1.

Valor Máximo: Infinito.

Ecuación:

$$C12 = \text{Promedio} \left(\frac{\text{Valor del Pico}}{0.09G} \right) \quad (23)$$

Rangos de clasificación:

- “Bueno”: [1 – 1.33]. Promedio de frenado entre 0.09 G a 0.1 G.
- “Regular”: [1.33 – 1.67]. Promedio de frenado entre 0.1 G a 0.15 G.
- “Malo”: [1.67 - Inf]. Promedio de frenado mayor a 0.15 G.

4.3.3.13. Porcentaje de tiempo que dura Acelerando (Aceleraciones mayores a 0.09G)

Eje: Aceleración Y - Aceleraciones

Unidades: [Adimensional]

Descripción: Esta característica define el porcentaje del tiempo total que el conductor realizo aceleraciones mayores a 0.09 G.

Valor Mínimo: 0 %

Valor Máximo: 100%.

Ecuación:

$$C13 = \frac{\text{Duración Total de Frenado}}{\text{Duración Total del Trayecto}} * 100 \quad (24)$$

Rangos de clasificación:

- “Bueno”: [0 - 5]. Del 0% al 5% del tiempo frenando.
- “Regular”: [5 - 10]. Del 5% al 10% del tiempo frenando.
- “Malo”: [10 - 100]. Más del 10% del tiempo frenando.

4.3.3.14. Frecuencia de Picos Que Superan los 40 Km/H

Eje: Velocidad

Unidades: [# / min]

Descripción: Esta característica define en número de veces que en la muestra el conductor superó los 40 Km/h sobre el tiempo de duración del trayecto.

Valor Mínimo: 0.

Valor Máximo: Infinito.

Ecuación:

$$C14 = \frac{\text{número de veces}}{\text{Duración de la Grabación}} \quad (25)$$

Rangos de clasificación:

- “Bueno”: [0 - 1]. El tiempo entre frenadas es igual o mayor a 1 minuto.
- “Regular”: [1 - 2]. El tiempo entre frenadas es menor a 1 y mayor a medio minuto.
- “Malo”: [2 - Inf]. El tiempo entre frenadas es igual o menor a medio minuto.

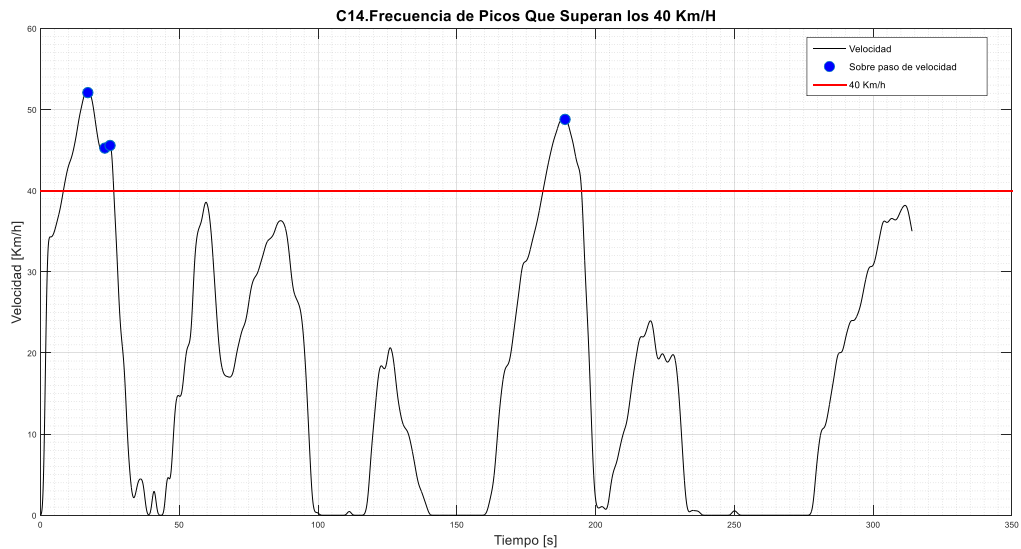


Fig. 29: Frecuencia de picos que superan los 40Km/h

4.3.3.15. Promedio de los picos previos a una detención

Eje: Aceleración Y - Frenadas

Unidades: [G]

Descripción: Esta característica define el promedio de las frenadas que se realizaron previas a cualquier detención durante el trayecto.

Valor Mínimo: 0

Valor Máximo: Infinito.

Ecuación:

$$C15 = \text{Promedio}(\text{Picos previos a una detención}) \quad (26)$$

Rangos de clasificación:

- “Bueno”: [0 – 0.09]. Promedio de frenado previo a una detención menor a 0.09 G.
- “Regular”: [0.09 – 0.15]. Promedio de frenado previo a una detención entre 0.09 G y 0.15G.
- “Malo”: [0.15 - Inf]. Promedio de frenado previo a una detención mayor a 0.15 G

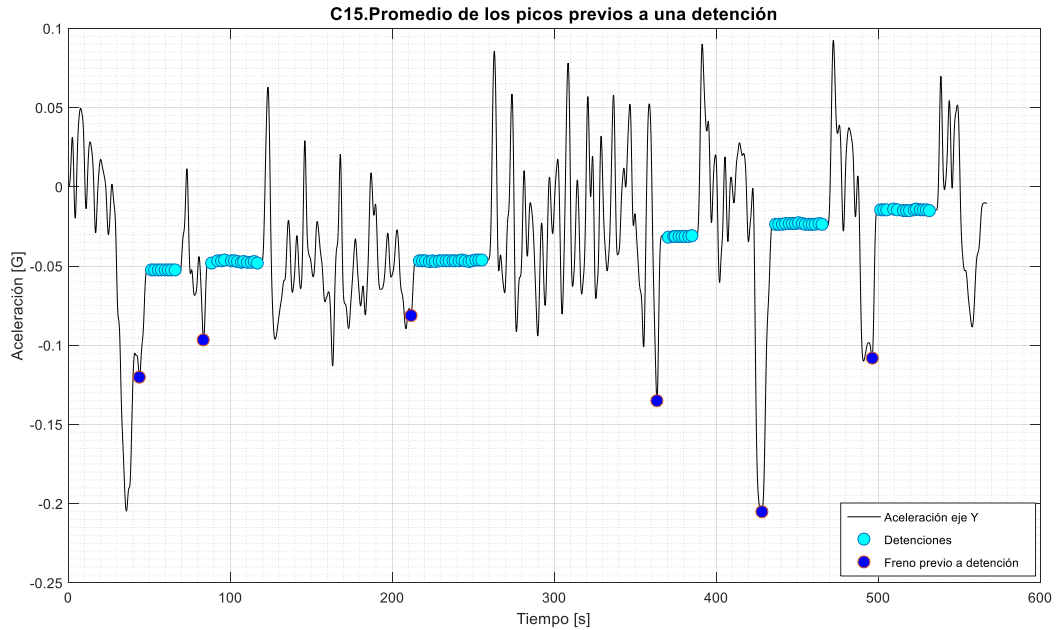


Fig. 30: Promedio de los picos previos a una detención

Para la definición de las siguientes 2 características se calculó una variable nueva con la siguiente fórmula:

$$W = \text{Velocidad}(\text{Antes de la frenada}) * \text{Aceleración}(\text{Frenada}) \quad (27)$$

Esto con el fin de poder analizar otro fenómeno que ocurre durante el trayecto y es la combinación entre la velocidad que tiene el bus y la frenada que este realiza.

4.3.3.16. Promedio de W cuando los picos de aceleración supera |0.09 G|.

Eje: Aceleración Y – Frenadas y Velocidad

Unidades: [(m²/s³)]

Descripción: Esta característica define el promedio de la variable W medida cada vez que se realizó una frenada mayor a 0.09 G.

Valor Mínimo: 0

Valor Máximo: Inf

Ecuación:

$$C16 = \frac{\text{Promedio}(\text{Aceleración} * \text{Velocidad})}{\text{Duración Total del Trayecto}} \quad (28)$$

Rangos de clasificación:

- “Bueno”: [0 – 3.6]. Promedio de W cuando es menor a 3.6.
- “Regular”: [3.6 – 4.5]. Promedio de W cuando está entre 3.6 y 4.5
- “Malo”: [4.5 - 100]. Promedio de W cuando es mayor a 4.5.

4.3.3.17. Frecuencia de W cuando los picos de aceleración supera |0.09 G|.

Eje: Aceleración Y – Frenadas y Velocidad

Unidades: [(m/s²) ^ 2]

Descripción: Esta característica define el número de veces que, durante el trayecto, se supera el valor de la variable $W = 3,6 (m/s^2)^2$

Valor Mínimo: 0

Valor Máximo: Inf

Ecuación:

$$C17 = \text{Numero de veces} (\text{Velocidad} * \text{Aceleración}) \quad (29)$$

Rangos de clasificación:

- “Bueno”: [0 – 3]. Supera los 3.6 menos de 3 veces por minuto
- “Regular”: [3 – 6]. Supera los 3.6 entre 3 a 6 veces por minuto
- “Malo”: [6 - 100]. Supera los 3.6 más de 6 veces por minuto.

4.4. Fase Cuatro: Análisis estadístico y clasificación de patrones.

En esta fase se enunciará la reducción de dimensionalidad por medio de PCA y la técnica de agrupamiento (clustering).

4.4.1.Reducción de dimensionalidad

Debido a la cantidad de características que se definieron en la fase 3 y las 50 muestras recolectadas, la cantidad de datos obtenidos es un espacio dimensional grande, por lo que se hizo complicado realizar un análisis a priori de estos puntos. La dimensionalidad del problema hace que la información relevante se oculte entre el nube de datos. Otra razón que refuerza por que fue necesario la reducción de dimensionalidad fue porque las características definidas están relacionadas entre sí, es decir que la correlación entre ellas es diferente de 0. Con ayuda de MATLAB se analizaron los coeficientes de correlación entre las características obteniendo el siguiente resultado.

Tabla 1. Tabla de correlaciones de las 17 características.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
C1	1,0000	0,2304	0,7743	0,5633	0,4510	0,3316	0,5049	0,1270	-0,0353	-0,1383	-0,0873	-0,0964	-0,1128	0,0867	-0,1400	-0,0123	-0,0374
C2	0,2304	1,0000	0,3732	0,6075	0,0587	0,5716	0,7266	0,8527	0,1990	0,2252	-0,1462	-0,2128	-0,0749	0,2234	0,5147	0,4959	0,2556
C3	0,7743	0,3732	1,0000	0,6389	0,0948	0,6611	0,4986	0,3592	-0,0875	0,1367	-0,1705	-0,2436	-0,1643	0,1392	-0,0486	0,0683	0,1234
C4	0,5633	0,6075	0,6389	1,0000	0,2783	0,8453	0,3505	0,7990	0,0554	0,2142	-0,2836	-0,3169	-0,2424	-0,0119	0,1657	0,2688	-0,0040
C5	0,4510	0,0587	0,0948	0,2783	1,0000	0,0278	0,1759	0,0722	0,0907	-0,3244	0,2163	-0,0076	0,2059	-0,1995	-0,0126	-0,2360	-0,4943
C6	0,3316	0,5716	0,6611	0,8453	0,0278	1,0000	0,2479	0,7846	0,0132	0,3480	-0,2401	-0,2743	-0,1859	0,0723	0,1269	0,2747	0,1357
C7	0,5049	0,7266	0,4986	0,3505	0,1759	0,2479	1,0000	0,4795	0,0831	0,0646	0,1716	-0,0278	0,2214	0,3019	0,1504	0,2348	0,2753
C8	0,1270	0,8527	0,3592	0,7990	0,0722	0,7846	0,4795	1,0000	0,1026	0,3389	-0,1562	-0,2525	-0,0785	0,1346	0,3856	0,4049	0,1702
C9	-0,0353	0,1990	-0,0875	0,0554	0,0907	0,0132	0,0831	0,1026	1,0000	-0,1601	-0,0018	0,0138	-0,0597	0,2025	0,2814	0,2230	-0,0257
C10	-0,1383	0,2252	0,1367	0,2142	-0,3244	0,3480	0,0646	0,3389	-0,1601	1,0000	0,0209	-0,2349	0,0723	0,1926	0,0416	0,2746	0,3143
C11	-0,0873	-0,1462	-0,1705	-0,2836	0,2163	-0,2401	0,1716	-0,1562	-0,0018	0,0209	1,0000	0,4013	0,9693	0,0645	-0,2661	-0,1884	0,0556
C12	-0,0964	-0,2128	-0,2436	-0,3169	-0,0076	-0,2743	-0,0278	-0,2525	0,0138	-0,2349	0,4013	1,0000	0,4419	0,2153	-0,3419	-0,0270	0,2489
C13	-0,1128	-0,0749	-0,1643	-0,2424	0,2059	-0,1859	0,2214	-0,0785	-0,0597	0,0723	0,9693	0,4419	1,0000	0,0946	-0,2234	-0,1204	0,0978
C14	0,0867	0,2234	0,1392	-0,0119	-0,1995	0,0723	0,3019	0,1346	0,2025	0,1926	0,0645	0,2153	0,0946	1,0000	0,1141	0,4271	0,7064
C15	-0,1400	0,5147	-0,0486	0,1657	-0,0126	0,1269	0,1504	0,3856	0,2814	0,0416	-0,2661	-0,3419	-0,2234	0,1141	1,0000	0,3234	0,0483
C16	-0,0123	0,4959	0,0683	0,2688	-0,2360	0,2747	0,2348	0,4049	0,2230	0,2746	-0,1884	-0,0270	-0,1204	0,4271	0,3234	1,0000	0,6473
C17	-0,0374	0,2556	0,1234	-0,0040	-0,4943	0,1357	0,2753	0,1702	-0,0257	0,3143	0,0556	0,2489	0,0978	0,7064	0,0483	0,6473	1,0000

Como se puede observar existen características que tienen un alto coeficiente de correlación entre ellas, lo cual significa que existe información redundante.

La técnica que se utilizó para la reducción de la dimensionalidad de los datos fue análisis de componentes principales (PCA) explicado anteriormente. MATLAB cuenta con la función *PCA*, la cual calcula las componentes principales de los datos que sean ingresados, pero antes de ingresar la matriz fue necesario normalizar los datos, esto a causa de que las escalas que representan cada una de las características son diferentes, por lo que se ajustaron todos los datos a una escala común.

La normalización que se utilizó para la reducción de normalidad, fue la siguiente:

$$\frac{X_{i,j} - \mu_i}{\sigma_i} \quad (30)$$

j: Representa cualquiera de las muestras.

i: Representa cualquiera de las características.

X: Es el dato que se está analizando perteneciente a una muestra j y a una característica i.

μ : Es la media correspondiente a la característica i.

σ : Es la desviación estándar correspondiente a la característica i.

Luego de utilizar PCA en el análisis de patrones fue necesario determinar la cantidad de componentes principales con los que se va a trabajar. Se utilizó el criterio de la proporción de varianza explicada de su nombre en inglés The Proportion of Variance Explained Criterion) [13], con el fin de saber cuánta fue la información que quedó después de hacer PCA.

El criterio de proporción de varianza explicada que se utilizó fue mayor al 80% de la información.

4.4.2. Agrupamiento de datos.

Posterior a la reducción de dimensionalidad, se realizó el agrupamiento de datos por medio de la técnica k-means, este método de agrupamiento requiere que se le provea un criterio de medida de similitud entre los datos y los centroides estipulados.

Para este análisis y de manera general, la técnica de k-means utiliza la métrica euclidiana para la medición de las distancias y aunque la distorsión en los resultados es menor utilizando otras métricas, se obtiene mejor resultado con la distancia euclidiana [14].

Las 4 opciones de métricas que MATLAB tiene para este tipo de análisis son las siguientes:

- Euclidiana

$$d(x, c) = (x - c)(x - c)' \quad (31)$$

- Cityblock

$$d(x, c) = \sum_{j=1}^p |x_j - c_j| \quad (32)$$

- Cosine

$$d(x, c) = 1 - \frac{xc'}{\sqrt{(xx')(cc')}} \quad (33)$$

- Correlation

$$d(x, c) = 1 - \frac{(x - \bar{x})(c - \bar{c})}{\sqrt{(x - \bar{x})(x - \bar{x})'} \sqrt{(c - \bar{c})(c - \bar{c})'}} \quad (34)$$

$c = \text{centroide}; x = \text{muestra}$

Otro parámetro necesario para el funcionamiento de esta técnica, es determinar el número de clusters, centroides o k-means. Para determinar el número correcto de clusters que se debe utilizar según la muestra que se tiene, se realizó el siguiente procedimiento.

1. Se simuló el agrupamiento de las muestras analizadas con 2, 3, 4 y 5 clusters.
2. Se verificó cuál de las 4 opciones es la correcta con ayuda de la función *Silhouette*.

5. PROTOCOLO DE PRUEBAS YANÁLISIS DE RESULTADOS.

En este capítulo se analizarán y se interpretarán los resultados que se obtuvieron durante el análisis, respecto a la información obtenida por la clasificación no supervisada y los datos obtenidos después de la realización de PCA junto con los clusters establecidos. Estos resultados se obtuvieron luego de ingresar una base de datos de 50 muestras al sistema. Es importante aclarar que en el momento que se agregue o se elimine una nueva muestra los resultados van a cambiar.

El primer análisis que se realizó fue ver el comportamiento y la información estadística de cada una de las características definidas. A continuación se muestran los resultados obtenidos.

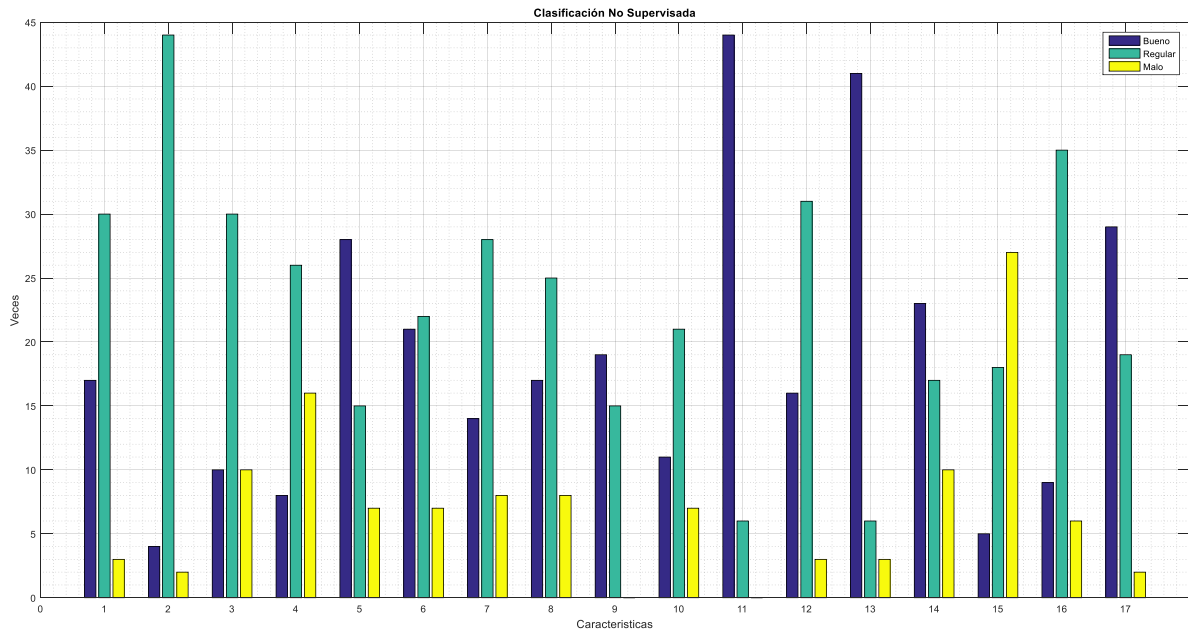


Fig. 31. Histograma general de las 17 características

La Fig. 31 ilustra cómo está distribuida cada una de las características, respecto a los rangos definidos en la sección 4.3.3 y por la cantidad de muestras tomadas se analizó la probabilidad de que un conductor pueda o no pertenecer a un rango específico. Observamos que la probabilidad de que un conductor pertenezca al rango “Regular”, si es analizado con la característica 2 es alta, del mismo modo es un poco más probable que si se analiza al mismo conductor con la característica 15 puede llegar a pertenecer al rango “Malo”. Otro análisis que se puede observar es que es poco probable que en la característica 11 se clasifique a un conductor en el rango “Malo”. Se puede inferir que esto es debido a la transmisión automática del vehículo que hace que este no pueda exceder los límites de aceleración explicados en la Fase uno.

Para el análisis de componentes principales se debe escoger la cantidad de componentes con los que se va a trabajar, a continuación se muestran los valores propios y la porción de varianza de cada uno de los componentes.

Tabla 2 Valores propios y porción de varianza de cada uno de los componentes.

Valor Propio	% de Varianza	Acumulativa %
4,9649	29,2056	29,2056
2,7430	16,1352	45,3408
2,4696	14,5270	59,8678
1,6431	9,6651	69,5329
1,4355	8,4441	77,9770
0,8669	5,0992	83,0762
0,7320	4,3061	87,3823
0,5759	3,3878	90,7701
0,4710	2,7706	93,5407
0,4051	2,3830	95,9237
0,3193	1,8780	97,8017
0,1428	0,8398	98,6415
0,1127	0,6629	99,3044
0,0554	0,3258	99,6302
0,0393	0,2314	99,8616
0,0171	0,1003	99,9619
0,0065	0,0382	100,0001

Con estos resultados obtenidos se determinó trabajar con los 6 primeros componentes principales, trabajando con el 83,0762% de la información de las muestras. También se va a realizar un análisis con los 3 primeros componentes principales con el fin de poder visualizar el resultado. Con este segundo análisis se trabajó con el 60% de la información total.

El siguiente paso que se realizó fue interpretar los resultados obtenidos del agrupamiento de datos, con el fin de verificar que la métrica Euclidiana es la mejor para este trabajo, se experimentó cambiando la métrica de medición de distancia.

Como se explicó en el inciso 4.4.2, se simularon 2, 3, 4 y 5 clusters para cada una de las métricas y por medio de la función *Silhouette* se compararon los resultados. En la Fig. 36 se muestran las siluetas obtenidas.

La interpretación de este tipo de gráficas es la siguiente; cada una de las figuras muestra la cantidad de clusters que se simuló respectivamente en el eje Y. El eje X representa los valores de silueta. Idealmente se quiere que el promedio de estos valores de silueta sea muy parecido al valor de cada cluster, como se evidencia en la Fig. 32 de métrica Euclidiana y 2 clusters. Es decisión del investigador decidir si quiere o no que las siluetas sea parecidas. Como la finalidad de este análisis es investigar si existe alguna forma de clasificar los patrones de conducción y sobretodo de poder ver y clasificar casos extremos, no deseamos que los valores de silueta sean parecidos.

Las barras que se encuentran en el eje X negativo indican que los clusters no están separados de la mejor manera y que posiblemente estas muestras pertenezcan a otro cluster. En las gráficas de métricas Cosine, Correlation y Cityblock se ve que utilizar este tipo de medición no es adecuado para este trabajo, ya que la posición de los centros no va a ser la correcta distorsionando el resultado. La métrica Euclidiana tiene unas muestras fuera de sus clusters pero no están tan alejadas de su verdadero grupo, por lo que no generan mucha distorsión en los resultados.

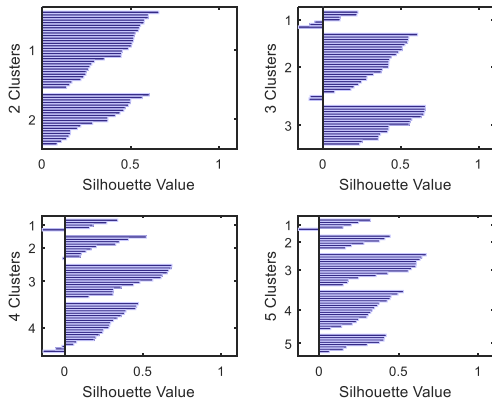


Fig. 32. *Metrica Sqeuclidean*

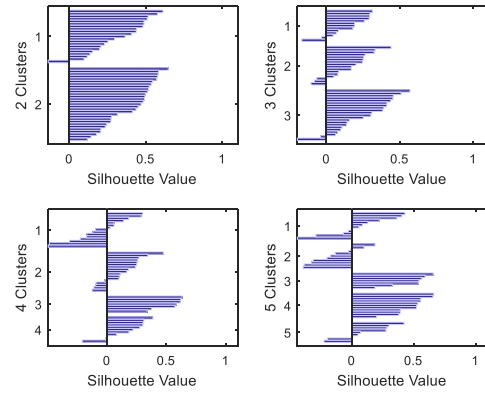


Fig. 33. *Metrica Cityblock*

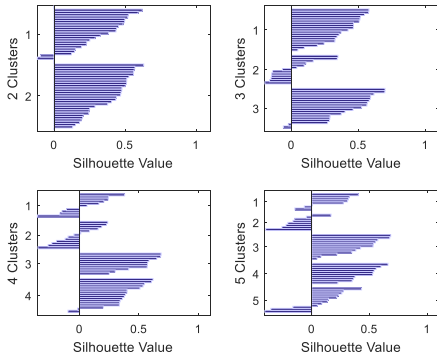


Fig. 34. *Metrica Cosine*

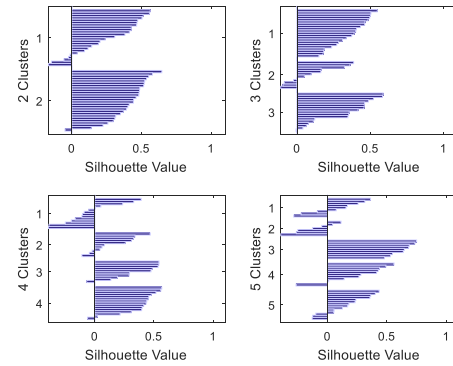


Fig. 35. *Metrica Correlation.*

Fig. 36 *Metricas Tecnica de agrupamiento.*

Como se explicó en la Tabla 2, el análisis de patrones de conducción requirió utilizar seis componentes principales agrupándolos en cinco clusters, como este análisis no es posible visualizarlo de forma gráfica, también se va a realizar un análisis reduciendo los componentes principales a tres y agrupándolos en cuatro clusters.

Para verificar la importancia de utilizar PCA se utilizó clustering con las características enunciadas en la sección 4.3.3., luego de realizar la clasificación no supervisada y encontrar los valores de las muestras normalizadas respecto a las características definidas se agruparon los datos en los 5 clusters establecidos. Los resultados obtenidos fueron los siguientes:

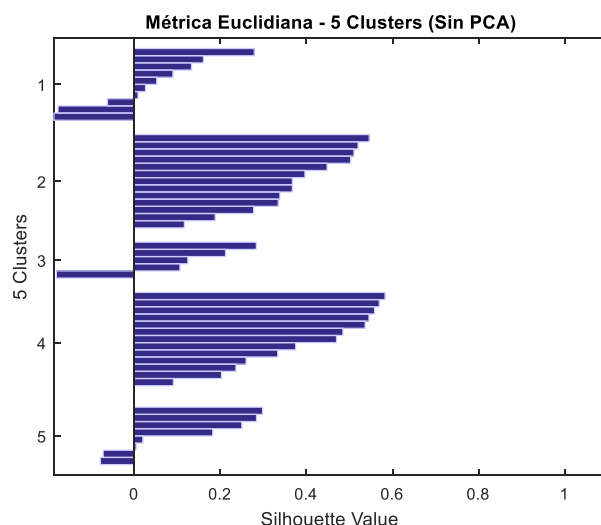


Fig. 37. Valores de Silueta Características principales.

Tabla 3. Ubicación de los centroides

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
Centroide 1	0.5271	0.3085	0.6786	-0.0399	-0.5374	0.1600	0.7253	0.0368	-0.4046	0.2254	-0.2008	0.6638	-0.1255	1.2259	-0.2748	0.4592	1.2458
Centroide 2	-0.3278	0.6066	0.0812	0.3487	-0.4476	0.4895	-0.1737	0.5501	0.1857	0.3614	-0.4275	-0.5404	-0.3984	-0.1696	0.8823	0.7249	0.1133
Centroide 3	-0.2723	-0.6497	-0.4965	-0.8912	0.1977	-0.7550	0.2962	-0.6153	-0.1640	-0.0134	2.6215	0.8045	2.4546	-0.1835	-0.7594	-0.8561	0.0504
Centroide 4	-0.7582	-0.9400	-0.8810	-0.8023	-0.1615	-0.7621	-0.8870	-0.8062	0.0717	-0.3621	-0.2684	0.0735	-0.2980	-0.4887	-0.4376	-0.5483	-0.6064
Centroide 5	1.3708	0.6797	0.8719	1.4443	1.5580	0.8100	0.7429	0.8556	0.1811	-0.2270	-0.2232	-0.5830	-0.2083	-0.2868	0.1501	-0.2574	-0.7117
Desv Estan	0.8449	0.7468	0.7478	0.9548	0.8526	0.7196	0.6862	0.7184	0.2548	0.3017	1.3006	0.65	1.2172	0.6864	0.6337	0.669	0.781

Con el cálculo de la desviación estándar se puede evidenciar que los puntos de cada uno de los clusters no están separados como se espera, esto demuestra que hay información redundante y que los agrupamientos realizados no tienen diferencias entre unos y otros. Los valores de la Fig. 37 tienen 6 muestras que no se encuentran en su cluster respectivo.

La suma, dentro del mismo cluster, de los puntos al centro se encuentra en la Tabla 4.

Tabla 4. Within-cluster sums of point-to-centroid distances

	Suma
Centroide 1	109,1031
Centroide 2	91,7558
Centroide 3	67,1896
Centroide 4	106,6945
Centroide 5	93,0283

Se puede observar que los centroides están muy separados de las muestras que pertenecen a cada uno de los grupos, lo que significa que el centro y las muestras del mismo cluster no necesariamente tengan un rasgo en común que los identifique. Además el hecho de tener un agrupamiento en 17 dimensiones, y sabiendo que estas variables dependen unas de otras, resulta más difícil concluir al respecto.

Como contraparte al análisis de no utilizar PCA en el Proyecto, se realizó una prueba con las especificaciones que se plantearon anteriormente, que son el uso de los 6 componentes principales y 5 clusters, es decir el 83% de la información estudiada. A continuación se muestran los resultados obtenidos.

Tabla 5. Centroides PCA.

	1° CP	2° CP	3° CP	4° CP	5° CP	6° CP
Centroide 1	-2,4097	1,4041	2,7659	0,2510	1,1833	0,2769
Centroide 2	0,2059	0,1157	0,0480	-2,2290	-0,9747	0,4967
Centroide 3	-2,2695	-0,5483	-0,8471	0,0773	-0,0247	-0,2217
Centroide 4	1,4673	0,8906	-0,7526	0,4260	0,0954	-0,0232
Centroide 5	2,2054	-1,8973	1,3171	0,4744	-0,1679	-0,1082
Desv Estan	2,1147	1,2921	1,53213	1,14497	0,77228	0,29556

Tabla 6. Within-cluster sums of point-to-centroid distances en PCA

	Suma
Centroide 1	53,769674
Centroide 2	33,902921
Centroide 3	79,642735
Centroide 4	105,92454
Centroide 5	62,788789

Evidentemente al realizar análisis de componentes principales los cinco centroides tienen más distancia entre ellos que el resultado observado en la Tabla 3 y la información dada por cada cluster es diferente, además se observa que los centroides están más cerca a las muestras pertenecientes, totalmente contrario a lo sucedido en la Tabla 4. A continuación se interpretará cada uno de los componentes principales y posteriormente el significado de cada centroide.

Para la interpretación de los componentes principales es necesario remitirse al ANEXO MATRIZ DE COEFF. Esta matriz contiene los coeficientes de las transformaciones lineales de variables originales que dan un lugar a la componente principal ¹⁰

- El primer componente da mayor peso a las características C12, C14 y C17 y por otra parte contrapone las características C5. Este componente tiene en cuenta positivamente las variables que tiene que ver con velocidad.
- El segundo componente da mayor peso a las características C1, C14 y C16 y por otra parte contrapone las características C15. Este componente tiene cuenta positivamente las variables de frenado y velocidad, pero negativamente los frenos previos a detenciones.
- El tercer componente contrapone las características C4, C10, C12 y C15. Este componente tiene en cuenta negativamente eventos variados.
- El cuarto componente da mayor peso a las características C1 y C12 y por otra parte contrapone las características C6 y C17.
- El quinto componente da mayor peso a las características C3, C4, C8 y C13. Este componente tiene en cuenta positivamente los frenos bruscos y el tiempo que dura acelerando y frenando.

¹⁰ Tomado de: <http://simfit.usal.es/casos/curso2009/caso17.pdf> (24 de octubre del 2016)

- El sexto componente da mayor peso a las características C1 y C15 y por otra parte contrapone las características C6 y C12. ¹¹

Al reducir a 3 los componentes principales con el fin de visualizar los agrupamientos realizados se pudo interpretar mejor el significado de cada cluster.

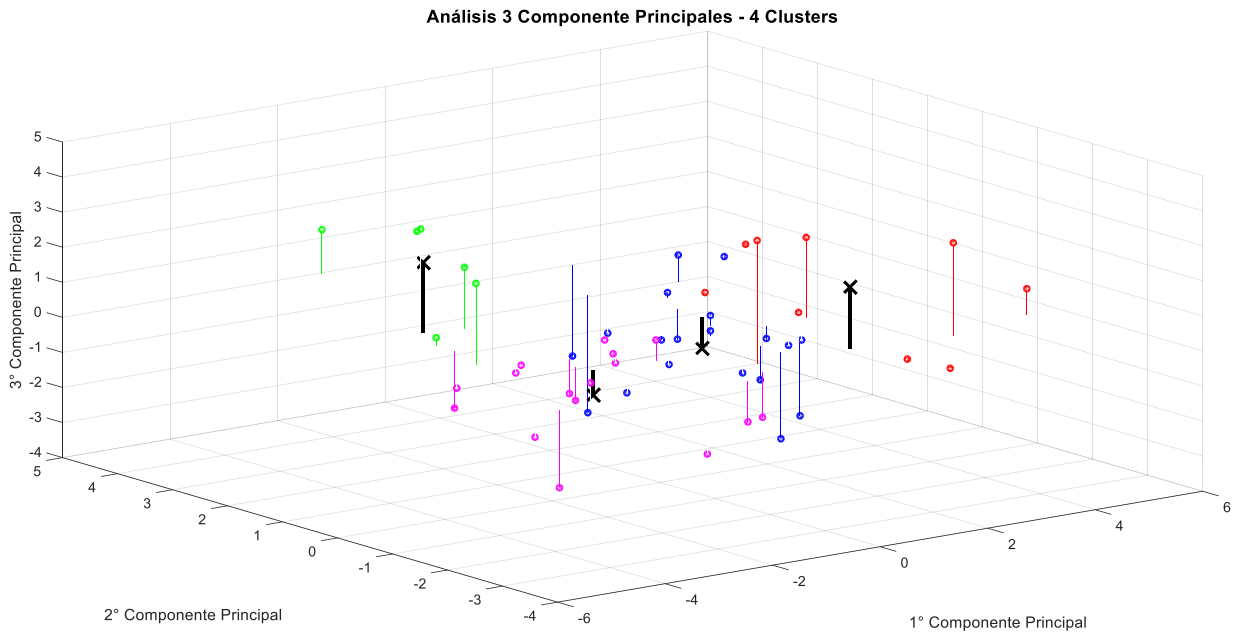


Fig. 38. PCA para las tres componentes principales, con 4 clusters en tres dimensiones.

En la Fig. 38 se observa el agrupamiento realizado, ahora se va a mostrar el resultado en 2 dimensiones

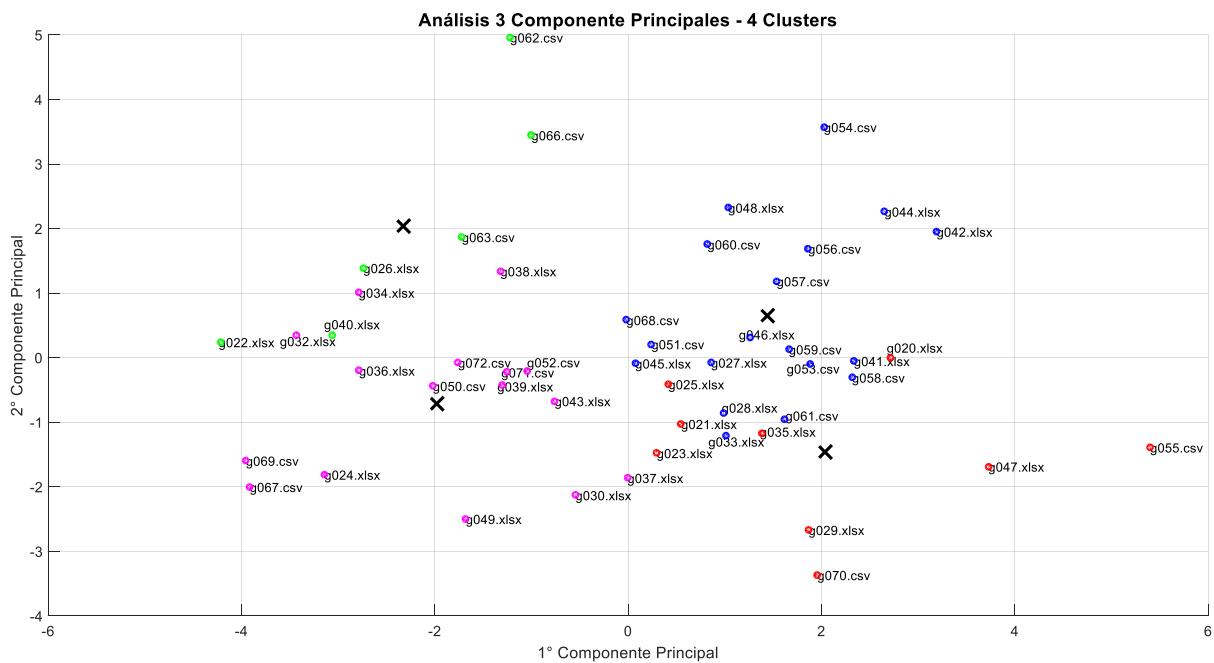


Fig. 39 PCA para las tres componentes principales, con 4 clusters en dos dimensiones 2PCA vs 1PCA

¹¹ Interpretación tomada de: https://www.uam.es/personal_pdi/ciencias/dfaraco/docencia/AD/ComponentesPrincipales.pdf (20 de noviembre de 2016)

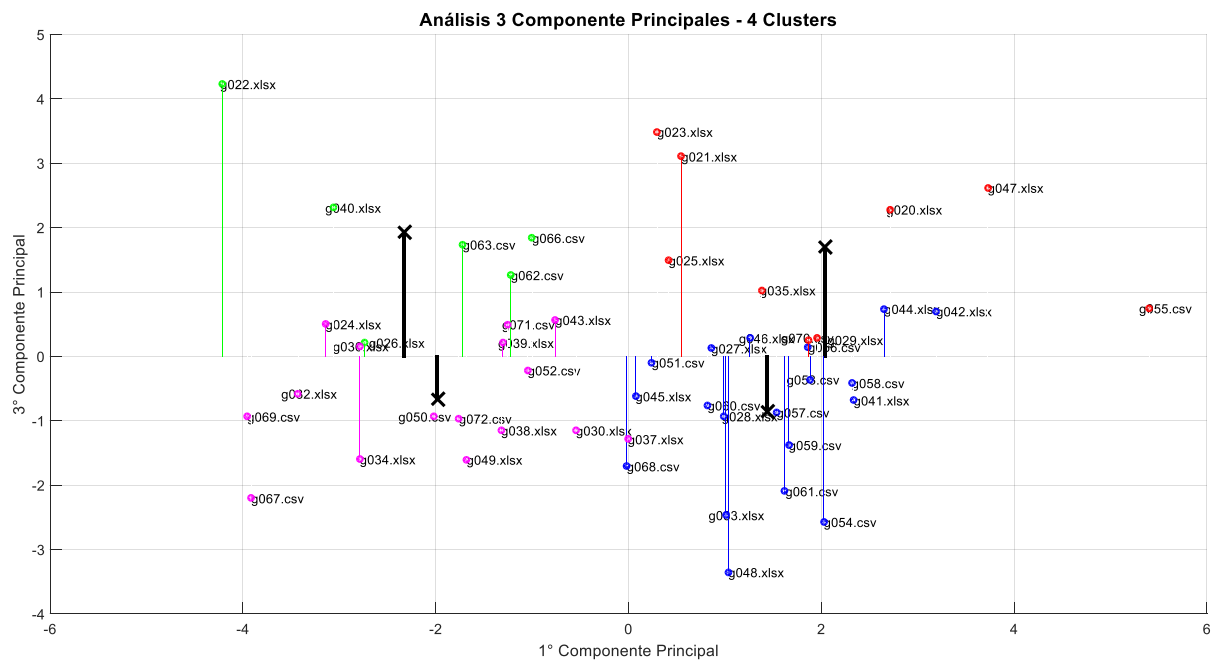


Fig. 40. PCA para las tres componentes principales, con 4 clusters en dos dimensiones 3PCA vs 1PCA.

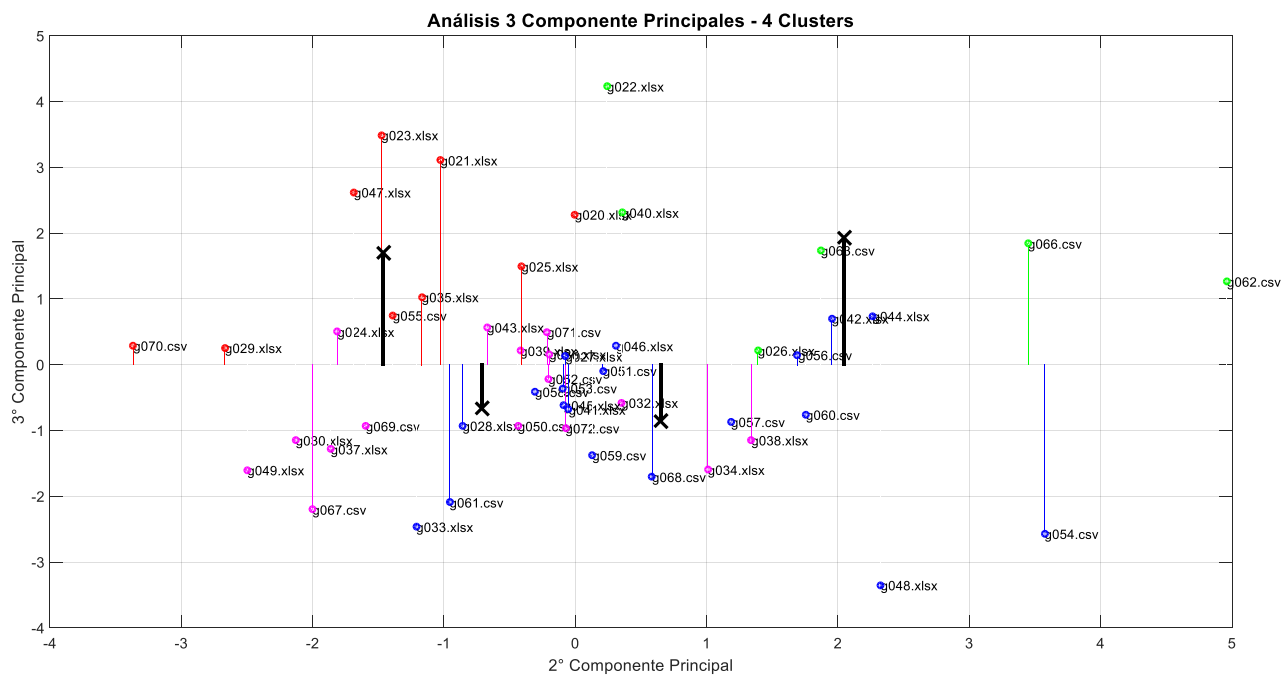


Fig. 41. PCA para las tres componentes principales, con 4 clusters en dos dimensiones 3PCA vs 2PCA.

Ahora se prosigue con interpretar el significado de cada cluster.

Tabla 7. Ubicación de centroides.

	1° CP	2° CP	3° CP
Centroide 1	1,7172	2,3909	-0,1586
Centroide 2	1,7042	-0,904	0,0791
Centroide 3	-2,41	1,4041	2,7659
Centroide 4	-1,892	-0,315	-0,7946

Teniendo en cuenta los datos en la Tabla 7 podemos concluir que cada agrupación tiene el siguiente significado.

- El primer cluster (color rojo Fig. 39) tiene mayor relación con la segunda componente principal, esto significa, que este centroide tiene en cuenta positivamente las variables de frenado y velocidad, pero negativamente los frenos previos a detenciones.

Este grupo de vehículos, según las características descritas, corresponde a conductores que frenan con una aceleración mayor a los estándares del numeral 4.3.3., sin llegar a ser peligrosa y manejan con velocidades mayores a 40 Km/h. Estos conductores según los resultados de la Fig. 16 tienen una mayor probabilidad de que manejen en el rango “Regular”.

Analizando un poco más profundo las muestras de este cluster, la muestra g054 tiene mucha dependencia de la segunda componente (Fig. 39) por lo que se puede inferir que su estilo de manejo sea brusco a comparación del resto del cluster.

La muestra g048 tiene mucho más dependencia negativa en la 3 componente lo cual significa que mientras las otras muestras contraponen eventos variados, este conductor puede ser considerado más brusco por las posibles maniobras que hace con respecto al resto del grupo.
- El segundo cluster (color azul) tiene mayor relación con el primer componente principal, esto significa que este centroide tiene en cuenta las variables que dependen de la velocidad. Estos conductores tienen una probabilidad mayor de manejar con velocidades superiores a las enunciadas en la sección 4.3.3., pero no poseen maniobras bruscas comparándolo con el cluster rojo.

La muestra representada con el nombre de g055 tiene más acentuado la característica de manejar con velocidades altas.
- El tercer cluster (color verde) es una agrupación que depende mucho de la tercera componente que a su vez se interpreta como conductores que dependen negativamente de variables como frenos peligrosos y frenos bruscos antes de detenciones, es decir que a medida que es mayor la dependencia sobre la tercera componente, crece con signo contrario la dependencia a estas características.

Este cluster tiene una particularidad, es que las muestras se distribuyen a lo largo de la segunda componente, esto se interpreta como un grupo de conductores que no tienen rasgos de manejar agresivamente pero que pueden tener maniobras bruscas de vez en cuando, sobre todo cuando manejan a velocidades mayores a 30 Km/h.
- El cuarto cluster (color magenta) es el grupo que más cerca tiene las muestras respecto al centroide y tiene una característica importante y es que depende negativamente de la primera componente y esta a su vez depende negativamente de la característica que contiene Jerk, por lo que estas muestras tienen un rasgo muy definido y son aquellas muestras que reportaron gran cantidad de eventos superiores al límite establecido en la gráfica de jerk o Jalón.

6. CONCLUSIONES Y RECOMENDACIONES:

Para poder tener una mayor cantidad de muestras y que el estudio tenga más validez estadística, se recomienda que el dispositivo con el que se toman los datos este embebido en el vehículo, obteniendo una base de datos más grande, con mediciones constantes y a diferentes horas del día. Esta recomendación hace que la aleatoriedad de la muestra sea mayor por ende un mejor análisis.

A priori se conocía que la manera en que los conductores manejan en la ciudad de Bogotá es muy variable. Al realizar la reducción de dimensionalidad la primera componente encontrada solo resumía el 30 % de la variabilidad total de las muestras (Tabla 2), esto confirma que se tiene que analizar diferentes eventos de las dinámicas de movimiento para dar juico respecto a cómo conducen los vehículos ya que en una sola componente solo se tiene la tercera parte de la variabilidad del modo de conducción.

Respecto al análisis realizado con las técnicas trabajadas es necesario entender, que se están analizando ejes coordenados compuestos por una combinación lineal de varias características conocidas. Es complicado hacer juzgamientos numéricos sobre este tipo de trabajos ya que se están manejando demasiadas variables, las cuales tienen correlación entre ellas. Como proceso de análisis se planteó una metodología la cual se fundamenta en la búsqueda de aquellos factores que más dependencia generan, ya sea positiva o negativa y de ahí comenzar a realizar el análisis cualitativo de las muestras.

Existen más eventos que se pueden analizarse para conocer las dinámicas de movimiento y no solo los expuestos en la sección 4.3.2, como por ejemplo dinámicas combinadas entre los diferentes ejes de la aceleración. El análisis que se realizó estuvo limitado por la trayectoria que se escogió, un trayecto más complejo incrementaría las dinámicas de movimiento por estudiar.

7. Bibliografía

- [1] Bogota, «EL tiempo,» 25 Mayo 2014. [En línea]. Available: <http://www.eltiempo.com/bogota/las-quejas-contra-el-sistema-integrado-de-transporte-publico-/14035757>. [Último acceso: 15 09 2015].
- [2] A. Loaiza, «www.sura.com,» 01 Enero 2015. [En línea]. Available: <http://www.sura.com/blogs/autos/accidentes-transito-colombia.aspx>. [Último acceso: 05 Septiembre 2015].
- [3] C. C. S.A., «La Seguridad Viaja en Autobus,» Instituto Biomedica de Valencia, Valencia y Zaragoza.
- [4] «AN INVESTIGATION OF ACCELERATION AND JERK PROFILES OF PUBLIC TRNSPORTATION VEHICLES,» Pensylvania State University-Berks, Pensylvania, 2008.
- [5] D. A. J. a. M. M. Trivedi, «Driving Style Recognition Using a Smartphone as a Sensor Platform,» University of California, San Diego, California, 2011.
- [6] V. R. Vuchic, URBAN TRANSIT SYSTEMS AND TECHNOLOGY, New Jersey : John Wiley & Sons, Inc., 2007.
- [7] Anonimo, «Universidad de Murcia,» .. [En línea]. Available: <http://www.um.es/geograf/sigmur/teledet/tema09.pdf>. [Último acceso: 3 10 2016].
- [8] J. Y. H. I. Y. C. M. G. a. R. M. Y. Wang, «Sensing Vehicle Dynamics for Determining Driver Phone use.».
- [9] T. C. a. P. B. A. Chowdhury, «, A Novel Approach To Improve Vehicle Speed Estimation Using Smartphoneâ€™s INS/GPS Sensors,» in 8th International Conference on Sensing Technology, Liverpool, UK, 2014..
- [10] «VEHICLE ACCELERATION ESTIMATION USING SMARTPHONE-BASED SENSORS.,» Department of E&E Engineering, Stellenbosch University.
- [11] Thomas, «SensorLog on the App Store,» App Store, 2015. [En línea]. Available: <https://itunes.apple.com/us/app/sensorlog/id388014573?mt=8>. [Último acceso: 09 Nov 2015].
- [12] A. M. d. B. D.C., «Alcaldía Mayor de Bogotá D.C.,» 02 05 2006. [En línea]. Available: <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=20190>. [Último acceso: 2016 05 10].
- [13] C. D. L. Daniel T. Larose, Data Mining and Predictive Analytics, Hoboken, New Jersey : John Wiley & Sons Inc., 2015.
- [14] A. Y. Archana Singh, «K-means with Three different Distance Metrics, Ajay Rana.,» Noida India, 2013.
- [15] J. B. MacQueen, «"Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability",» California, 1967.

8. ANEXOS

ANEXO 1

Tiempo	Acele X	Acele Y	Acele Y	Velocidad		
0,0666667	0,0699615	0,0142212	-1,0012970	5,9760000		
0,1333333	0,0636444	0,0246735	-0,9875946	5,9760000	Fecha	18/08/2016
0,2000000	0,0554199	0,0344086	-0,9776154	5,9760000	Hora	4:49 - 4:57
0,2666667	0,0450592	0,0237732	-0,9795990	5,9760000	Ruta	L80
0,3333333	0,0230865	0,0248871	-0,9864960	1,7640000	Trayecto	Jave - MN
0,4000000	0,0403595	0,0191803	-1,0120850	1,7640000		
0,4666667	0,0510254	0,0141296	-1,0002899	1,7640000		
0,5333333	0,0529785	0,0166931	-0,9971771	1,7640000		
0,6000000	0,0639191	0,0162506	-1,0025635	1,7640000		
0,6666667	0,0625153	0,0273132	-0,9981384	1,7640000		
0,7333333	0,0720215	0,0257874	-0,9928894	1,7640000		
0,8000000	0,0747528	0,0231018	-0,9839630	1,7640000		
0,8666667	0,0417633	0,0310516	-0,9690704	1,7640000		
0,9333333	0,0370941	0,0251617	-0,9880829	1,7640000		
1,0000000	0,0360565	0,0228882	-0,9926910	1,7640000		
1,0666667	0,0370941	0,0230713	-1,0006866	1,7640000		
1,1333333	0,0579681	0,0142212	-0,9993896	1,7640000		
1,2000000	0,0574799	0,0191650	-0,9937134	1,7640000		
1,2666667	0,0705719	0,0190887	-1,0088043	2,3760000		
1,3333333	0,0802155	0,0186615	-1,0005035	2,3760000		
1,4000000	0,0580139	0,0286865	-0,9872131	2,3760000		
1,4666667	0,0544586	0,0266724	-0,9749908	2,3760000		
1,5333333	0,0402069	0,0218658	-0,9796600	2,3760000		
1,6000000	0,0276489	0,0240021	-0,9994507	2,3760000		
1,6666667	0,0468140	0,0171814	-1,0066681	2,3760000		
1,7333333	0,0407410	0,0196991	-1,0036163	2,3760000		
1,8000000	0,0633850	0,0158997	-1,0137329	2,3760000		
1,8666667	0,0700378	0,0145111	-0,9937744	2,3760000		
1,9333333	0,0579987	0,0216217	-0,9936523	2,3760000		
2,0000000	0,0694885	0,0230103	-0,9924927	2,3760000		
2,0666667	0,0538788	0,0305634	-0,9718018	2,3760000		
2,1333333	0,0411682	0,0300293	-0,9799500	2,3760000		
2,2000000	0,0442047	0,0220642	-0,9881744	2,3760000		
2,2666667	0,0274048	0,0254974	-0,9982758	0,0000000		
2,3333333	0,0535278	0,0199432	-1,0097351	0,0000000		
2,4000000	0,0578766	0,0135040	-0,9986267	0,0000000		
2,4666667	0,0506744	0,0200043	-0,9932098	0,0000000		
2,5333333	0,0739288	0,0172119	-0,9954529	0,0000000		
2,6000000	0,0671997	0,0243378	-0,9924316	0,0000000		
2,6666667	0,0679321	0,0321045	-0,9890442	0,0000000		
2,7333333	0,0541077	0,0271454	-0,9760742	0,0000000		
2,8000000	0,0289917	0,0284882	-0,9777985	0,0000000		
2,8666667	0,0419006	0,0218353	-1,0084991	0,0000000		
2,9333333	0,0426331	0,0129700	-0,9972687	0,0000000		
3,0000000	0,0433960	0,0179749	-0,9951019	0,0000000		
3,0666667	0,0664520	0,0171967	-0,9948425	0,0000000		
3,1333333	0,0584564	0,0260162	-0,9947052	0,0000000		
3,2000000	0,0720215	0,0194244	-1,0010986	0,0000000		
3,2666667	0,0749207	0,0204163	-0,9900360	0,0000000		
3,3333333	0,0482025	0,0321350	-0,9772797	0,0000000		

ANEXO 2

```
%%--*****--
%%--          PONTIFICIA UNIVERSIDAD JAVERIANA          --
%%--          Trabajo de Grado #1554                    --
%%-- Análisis de Patrones de Conducción de Vehículos   --
%%--          Juan Felipe Mora Martínez                --
%%--          y                                         --
%%--          Mario Alberto Salamanca Ríos              --
%%--          INGENIERÍA ELECTRÓNICA                    --
%%--          2016                                       --
%%--*****--

clear all;
% Iniciación del software

% Se ingresa al software todas las muestras que se obtuvieron a lo largo
% del trabajo.

%Es necesario eliminar de la variable list todas aquellas funciones
%externas, archivos y muestras no deseadas para continuar el análisis

list = dir
list([1:22 34 67 68 76],:)=[];

% Procesamiento de señales

% En esta primera parte se aplican los filtros Moving-Average Filter y FIR1
% para eliminar en gran parte el ruido de las muestras.

%Se halla la densidad de potencia espectral (PSD) para la señal antes de
%filtrar y despues de hacerlo.

%DATOS = La estructura que contiene toda la información de cada muestra.
%     Nombre de la muestra
%     Aceleración X
%     Aceleración Y
%     Aceleración Z
%     Velocidad
%     Tiempo de la grabación
%     Rotulación

for i=1:length(list)
name = list(i).name
[N,Tx] = xlsread(name);
T = N(:,1);
L = length(T);
samples = 1/(T(L,1)/L);

Acex = N(:,2);
Acey = N(:,3);
Acez = N(:,4);
```

```

velo = N(:,8);

% Moving-Average Filter
samples = floor(samples);
Acex=filter(ones(1,samples)/samples,1,Acex);
Acey=filter(ones(1,samples)/samples,1,Acey);
Acez=filter(ones(1,samples)/samples,1,Acez);

Long = length(Acey);
fft_Acey = fft(Acey);
fft_Acey = fft_Acey(1:Long/2+1);
PSD_Acey = (1/(2*pi*Long)) * abs(fft_Acey).^2;
PSD_Acey(2:end-1) = 2*PSD_Acey(2:end-1);
freq = 0:(2*pi)/Long:pi;

% FIR1
n = 50;
wn = 0.03;
Filtro = fir1(n,wn);

newAcex = filter(Filtro,1,Acex);
newAcey = filter(Filtro,1,Acey);
newAcez = filter(Filtro,1,Acez);
newVelo = filter(Filtro,1,velo);

Long_filt = length(newAcey);
fft_newAcey = fft(newAcey);
fft_newAcey = fft_newAcey(1:Long_filt/2+1);
PSD_newAcey = (1/(2*pi*Long_filt)) * abs(fft_newAcey).^2;
PSD_newAcey(2:end-1) = 2*PSD_newAcey(2:end-1);
freq_filt = 0:(2*pi)/Long_filt:pi;

DATOS(i).Name = name;
DATOS(i).AceX = newAcex;
DATOS(i).AceY = newAcey;
DATOS(i).AceZ = newAcez;
DATOS(i).Velo = newVelo
DATOS(i).Tiempo = T;
DATOS(i).Info = Tx;

end

```

Published with MATLAB® R2015a

ANEXO 3

```
%Clasificación de los patrones obtenidos.
%Definición de las características

%Cada caracterisitica se simula por separado para cada una de las muestras
%que se tienen en la base de datos

%Primero es necesario compilar la función de Detención, algunas de las
%características necesitan de esta información

for index=1:length(list)
pk_all=0;
lc_all=0;
pk_min_all=0;
lc_min_all=0;

[pk_all,lc_all] = findpeaks(1000*DATOS(index).AceY);
[pk_min_all,lc_min_all] = findpeaks(-1000*DATOS(index).AceY);

%función Detencion: Identifica cuando un bus se detuvo durante el trayecto
%Parámetro de entrada:
%pk_all = Todos los picos que tiene la Aceleración en el eje Y.
%lc_all = La ubicación de todos los picos que tiene la Aceleración en el
%eje Y.
%pk_all_min = Todos los picos que tiene la (-Aceleración en el eje Y).
%index = El valor del índice del ciclo.

%Parámetro de Salida:
%minimo = Referencia minima que tiene la traza.
%pk_det = Los picos pertenecientes a una detención.
%lc_det = La ubiación de los picos pertenecientes a una detención.

[minimo,pk_det,lc_det] = Detencion(pk_all,lc_all,pk_min_all,index);
DATOS(index).Min = minimo;
end

%La ubiación de los picos se halla con el fin de poder gráficar estos
%puntos de ser necesario.

% 1. Frecuencia de Picos Que Superan los |0.09G|

for index=1:length(list)
[pk_C1,lc_C1] = findpeaks(-DATOS(index).AceY, 'MinPeakHeight',0.09);
Car_1 = length (pk_C1);
Car_1 = Car_1/DATOS(index).Tiempo(length(DATOS(index).Tiempo),1);
C1(index)=Car_1;
end
C1 = 60*C1;% Equivalencia en [#/minutos]

% 2.Promedio Sobre Paso en algún factor los picos mayores a |Referencia mínima|

for index=1:length(list)
[pk_C2] = findpeaks(-DATOS(index).AceY, 'MinPeakHeight',-DATOS(index).Min);
pk_C2 = pk_C2/0.09;
```

```

Car_2 = mean(pk_C2);
C2(index)=Car_2;
end

% 3.Porcentaje de tiempo que dura frenando (Frenos mayores a |0.09G|) (EJE Y-Frenados)

for index=1:length(list)
pk_C3_min = 0;
lc_C3_min = 0;
pk_C3_1 = 0;
lc_C3_1 = 0;
lc_C3 = 0;
pc_C3 = 0;

[pk_C3_min,lc_C3_min] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',0.09);
[pk_C3_1,lc_C3_1] = findpeaks(DATOS(index).AceY);

%función pkAnterior: Identifica el pico inmediatamente anterior respecto a
%los picos ingresados como parámetros de entrada
%Parámetro de entrada:
%lc_C3_min = la posición de los picos menos a 0.09G.
%pk_C3_1 = Todos los picos positivos
%lc_C3_1 = La posición de todos los picos positivos .
%index = El valor del índice del ciclo.

%Parámetro de Salida:
%minimo = Referencia minima que tiene la traza.
%pk_C3_max = vector de los picos inediatamente anteriores a lc_C3_min.
%lc_C3_max = vector de la posición de los picos inediatamente anteriores a lc_C3_min..
[pk_C3_max,lc_C3_max] = pkAnterior (lc_C3_min,pk_C3_1,lc_C3_1);

lc_C3_min = lc_C3_min';
pc_C3 = pk_C3_min;
pc_C3 = pc_C3';

%Creamos una nueva variable en DATOS ya que será utilizada en otras
%caracteristicas
DATOS(index).pk_16 = -pc_C3;
DATOS(index).lc_16 = lc_C3_max;

lc_C3 = lc_C3_min - lc_C3_max;
Car3 = sum(lc_C3);
Car3 = (Car3/length(DATOS(index).Tiempo))*100;
C3(index) = Car3;
end

% 4.Frecuencia de Picos Que Superan los |0.15G|

for index=1:length(list)
[pk_C4,lc_C4] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',0.15);
Car_4 = length (pk_C4);
Car_4 = (Car_4/(DATOS(index).Tiempo(length(DATOS(index).Tiempo),1)));
C4(index) = Car_4;
end

```

```

% 5. Frecuencia de picos que superan los 0.06 G/s (Jerk)
for index=1:length(list)

dt = diff(DATOS(index).Tiempo);
derivative = dI./dt;
DATOS(index).Jerk = derivative;
[pk_C5_max,lc_C5_max] = findpeaks(derivative,'MinPeakHeight',0.06);
[pk_C5_min,lc_C5_min] = findpeaks(-derivative,'MinPeakHeight',0.06);
%código tomado de http://stackoverflow.com/questions/27948140/plotting-derivative-of-data-in-matlab

Cat_5 = (length(pk_C5_max)+length(pk_C5_min))/DATOS(i).Tiempo(length(DATOS(i).Tiempo),1);
C5(k) = Cat_5;
end
C5 = 60*C5;

% 6.Porcentaje de tiempo que dura frenando (Frenos mayores a |0.15G|)

for index=1:length(list)
pk_C6_min = 0;
lc_C6_min = 0;
pk_C6_1 = 0;
lc_C6_1 = 0;
lc_C6 = 0;
pc_C6 = 0;
    if (isempty(pk_C4)==0)
[pk_C6_min,lc_C6_min] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',0.15);
[pk_C6_1,lc_C6_1] = findpeaks(DATOS(index).AceY);

[pk_C6_max,lc_C6_max]=pkAnterior(lc_C6_min,pk_C6_1,lc_C6_1);

lc_C6_min = lc_C6_min';
lc_C6 = lc_C6_min - lc_C6_max;
pc_C6 = pk_C6_min;
pc_C6 = pc_C6';
Car_6 = sum(lc_C6);
Car_6 = (Car_6/length(DATOS(index).Tiempo))*100;
    else
Car_6 = 0;
    end
C6(index)=Car_6;
end

% 7. Porcentaje de frenadas mayores a |0.09G|

for index=1:length(list)
pk_C1 = 0;
pk_C7 = 0;
[pk_C1,lc_C1] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',0.09);
[pk_C7,lc_C7] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',-DATOS(index).Min);
Car_7 = (length(pk_C1)/length(pk_C7))*100;
C7(index) = Car_7;
end

```

```
% 8. Porcentaje de frenadas mayores a |0.15G|
```

```
for index=1:length(list)
pk_C4 = 0;
pk_C8 = 0;
[pk_C4,lc_C4] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',0.15);
[pk_C8,lc_C8] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',-DATOS(index).Min);
Car_8 = (length(pk_C4)/length(pk_C8))*100;
C8(index) = Car_8;
```

```
end
```

```
% 9.Promedio Sobre Paso en algún factor los picos mayores a |0.09G|
```

```
for index=1:length(list)
[pk_C9_0] = findpeaks(-DATOS(index).AceX,'MinPeakHeight',0.09);
pk_C9_0 = pk_C9_0/0.09;
C9_0 = mean(pk_C9_0,'omitnan');
[pk_C9_1] = findpeaks(DATOS(index).AceX,'MinPeakHeight',0.09);
pk_C9_1 = pk_C9_1/0.09;
C9_1 = mean(pk_C9_1,'omitnan');
Car_9 = [C9_0 C9_1];
Car_9 = mean(Car_9,'omitnan');
Car_9 = abs(Car_9);
C9(index) = Car_9;
```

```
end
```

```
% 10.Promedio Sobre Paso en algún factor los picos mayores al |5% del promedio de la señal|
```

```
for index=1:length(list)
[pk_C10_0,lc_C10_0] = findpeaks(DATOS(index).AceZ);
DATOS(index).AceZ(1:lc_C10_0(1)) = NaN
New_Prom_Z = mean(DATOS(index).AceZ,'omitnan');

[pk_AceZ,lc_AceZ] = findpeaks(DATOS(index).AceZ-New_Prom_Z,'MinPeakHeight',(New_Prom_Z*0.05));
[pk_AceZ_m,lc_AceZ_m] = findpeaks(-
DATOS(index).AceZ+New_Prom_Z,'MinPeakHeight',(New_Prom_Z*0.05));
```

```
pk_AceZ = pk_AceZ/(New_Prom_Z*0.05);
C10_0 = mean(pk_AceZ,'omitnan');
pk_AceZ_m = pk_AceZ_m/(New_Prom_Z*0.05);
C10_1 = mean(pk_AceZ_m,'omitnan');
Car_10 = [C10_0 C10_1];
Car_10 = mean(Car_10,'omitnan');
Car_10 = abs(Car_10);
```

```
C10(index)=Car_10;
```

```
end
```

```
% 11 Frecuencia de Picos Que Superan los 0.09G
```

```
for index=1:length(list)
```



```

[pk_C11,lc_C11] = findpeaks(DATOS(index).AceY,'MinPeakHeight',0.09);
Car_11 = length(pk_C11);
Car_11 = (Car_11/DATOS(index).Tiempo(length(DATOS(index).Tiempo),1));
C11(index)=Car_11;

end
C11 = 60*C11;

% 12.Promedio Sobre Paso en algún factor los picos mayores a 0.09G

for i=1:length(list)
[pk_C12] = findpeaks(DATOS(i).AceY,'MinPeakHeight',0.09);
pk_C12 = pk_C12/0.09;
Car_12 = mean(pk_C12);
C12(i)=Car_12;
end

% 13. Porcentaje de tiempo que dura Acelerando (Aceleraciones mayores a 0.09G) (EJE Y -
Aceleraciones)

for index=1:length(list)
pk_C13_min = 0;
lc_C13_min = 0;
pk_C13_1 = 0;
lc_C13_1 = 0;
lc_C13 = 0;
pc_C13 = 0;
if (isempty(pk_C12)==0)
[pk_C13_max,lc_C13_max] = findpeaks(DATOS(index).AceY,'MinPeakHeight',0.09);
[pk_C13_1,lc_C13_1] = findpeaks(-DATOS(index).AceY);

[pk_C13_min,lc_C13_min]=pkAnterior(lc_C13_max,pk_C13_1,lc_C13_1);

lc_C13_max = lc_C13_max';

lc_C13 = lc_C13_max - lc_C13_min;
pc_C13 = pk_C13_min;
pc_C13 = pc_C13';
Car_13 = sum(lc_C13);
Car_13 = (Car_13/length(DATOS(index).Tiempo))*100;
else
Car_13 = 0;
end

C13(1)=Car_13;
end

% 15. Promedio de los picos previos a una detención (EJE Y -Frenadas)
for index=1:length(list)
pk_C15 = 0;
lc_C15 = 0;
lc_C15_new=0;
pk_C15_new=0;

```

```

[pk_C15,lc_C15] = findpeaks(-DATOS(index).AceY,'MinPeakHeight',-DATOS(index).Min*1.2);

Num_15=0;
j=1;
for i=1:length(lc_det)
    k = find(lc_C15<lc_det(i));
    if (isempty(k) == 0)
        N1 = k(length(k));
        if(N1 ~= Num_15)
            Num_15 = N1;
            lc_C15_new(j) = lc_C15(Num_15);
            pk_C15_new(j) = pk_C15(Num_15);
            j=j+1;
        end
    end
end

Car_15=mean(pk_C15_new);
C15(index)=Car_15;
end

% 14 Frecuencia de Picos Que Superan los 40 Km/H (EJE Velocidad)

for i=1:length(list)
    if(mean(DATOS(i).Velo ~= 0))

        [pk_C14,lc_C14] = findpeaks(DATOS(i).velo,'MinPeakHeight',40);
        Car_14 = length(pk_C14)/(DATOS(i).Tiempo(length(DATOS(i).Tiempo),1));
        C14(i) = Car_14;

%16. Promedio de combinacion Velocidad-Aceleracion, Picos Que Superan los |0.15G| (EJE Y-Frenados
y Velocidad)

        for j=1:length(DATOS(i).pk_16)
            if(DATOS(i).lc_16(j)~= 0)
                Prom_Ace_Vel(j) = DATOS(i).pk_16(j)*DATOS(i).Velo(DATOS(i).lc_16(j));
            end
        end

        Car_16=mean(-Prom_Ace_Vel);
        C16(i)=Car_16;

%17. Frecuencia de combinacion Velocidad-Aceleracion (40 km/h y 0.09 G), Picos Que Superan los
|0.15G| (EJE Y-Frenados y Velocidad)

        Car_17=length(find(-Prom_Ace_Vel>3.6));
        Car_17 = Car_17/DATOS(i).Tiempo(length(DATOS(i).Tiempo),1);
        C17(i)=Car_17;

        end
    end
    C14 = 60*C14;

%No se colocó código que tenga que ver con gráficas, se deja a disposición
%del lector realizar las gráficas como mejor le parezca

```

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
CP1	0.2201	-0.1653	0.3574 -	0.1747 -	0.4027	-0.0249 -	0.0311	0.1845	0.1684	0.1009	0.1760	0.3699 -	0.1531	0.3905	0.1647 -	0.1586	0.3656
CP2	0.3883	0.0929	0.0042	0.2389	0.1163	-0.2890	-0.0210	-0.1614	0.0139	0.1017	-0.1672	-0.1088	-0.0084	0.6206	-0.3595	0.2848	-0.1186
CP3	0.3093	-0.1365	0.2059	-0.3163	-0.2074	-0.0054	-0.1837	0.0190	-0.2708	0.0762	0.3646	-0.3782	0.0182	-0.2993	-0.4538	0.1097	0.0369
CP4	0.3838	-0.1957	0.1194	-0.0255	0.1212	0.2479	0.1279	-0.0177	0.0752	-0.1097	0.0073	0.5270	-0.1721	-0.2419	-0.0380	0.2687	-0.4993
CP5	0.0207	-0.1461	0.4590	0.2897	-0.0258	0.1344	0.2718	0.4157	0.3115	-0.1308	-0.2142	-0.2741	0.4084	-0.0790	-0.0922	0.0377	-0.0085
CP6	0.3703	-0.1207	0.0206	-0.1451	0.2244	0.3239	0.0335	-0.1386	-0.2451	-0.2117	0.1065	-0.3395	0.1769	0.2829	0.5286	-0.1442	-0.0900
CP7	0.2829	0.1839	0.2577	0.0514	-0.1335	-0.4955	-0.2310	-0.1615	0.0851	0.2463	-0.2332	-0.0662	0.0793	-0.3409	0.4100	-0.1255	-0.2032
CP8	0.3888	0.0114	-0.0094	0.1341	0.3340	0.0204	0.0855	-0.1660	-0.0691	-0.1660	-0.2696	0.1419	-0.0780	-0.2729	-0.2022	-0.3109	0.5835
CP9	0.0618	0.0806	-0.0424	0.4994	-0.2108	0.5437	-0.5415	-0.1219	-0.0110	0.2385	0.0144	0.0683	0.1404	-0.0297	-0.0013	0.0480	0.0748
CP10	0.1518	0.1512	-0.1775	-0.3179	0.4250	0.1119	-0.1761	0.5602	0.0862	0.5063	-0.1072	0.0940	0.0782	0.0227	0.0035	-0.0139	0.0153
CP11	-0.1353	0.3627	0.3751	-0.0183	0.2575	0.0490	-0.2291	-0.0359	0.1038	-0.1921	0.2332	0.0772	-0.0387	0.1356	-0.2627	-0.5431	-0.3080
CP12	-0.0363	0.3651	0.2389	0.1231	-0.1053	0.1736	0.5509	-0.0331	-0.5012	0.4285	-0.0015	0.0663	-0.0506	0.0044	0.0009	-0.0644	-0.0363
CP13	-0.1062	0.3926	0.3655	-0.0200	0.2979	0.0090	-0.1099	-0.0287	0.0572	-0.1630	0.2034	-0.0192	-0.1461	-0.0622	0.2181	0.5963	0.3145
CP14	0.1425	0.3810	-0.1073	-0.0847	-0.3635	0.1240	-0.1226	0.3733	-0.1775	-0.4180	-0.4079	-0.1209	-0.3449	0.0207	0.0004	0.0130	-0.0851
CP15	0.1670	0.0164	-0.2137	0.4928	0.0624	-0.3274	0.0045	0.4502	-0.2519	-0.1609	0.4952	0.1109	0.0324	-0.0649	0.1183	-0.0569	-0.0143
CP16	0.2432	0.2591	-0.2607	0.0829	-0.1153	0.1484	0.3071	-0.0951	0.5904	0.1203	0.3079	-0.3020	-0.3057	-0.0947	0.0293	-0.0890	-0.0042
CP17	0.1547	0.4200	-0.2194	-0.2507	-0.2136	-0.0133	0.1059	-0.1031	0.0831	-0.1899	0.1000	0.2846	0.6866	-0.0061	-0.0777	0.0793	0.0578

Anexos Digitales URL: <https://www.dropbox.com/home/ANEXOS%20TG1554>