

CALIDAD DEL DIESEL DE COSTA RICA ENTRE LOS AÑOS 2006-2010

Víctor Bazán Salazar ^{1*}

¹ Universidad de Costa Rica, Centro de Electroquímica y Energía Química (CELEQ)

Recibido agosto 2015; aceptado diciembre 2015

Abstract

This study explores the behavior and evolution on the quality of automobile diesel distributed in Costa Rica during the 2006-2010 period. Traditionally, fuel behavior studies are done on univariate and bivariate statistical methods but fail to provide sufficient information to describe the relationship between the attributes and multidimensionality of the individuals' behavior information. To overcome this limitation, a principal component analysis (PCA) is performed using the data gathered during the five years of the study. Additionally, a PCA was made for each individual year of the study to determine the evolution of diesel quality. Based on the analysis, it is found that sulfur concentration is the variable that most influences the behavior of individuals and the other variables, and because its progressive elimination in the fuel, its influence is canceled off during the last two years of the study. At the same time there is a decreasing in the fuel distillation temperatures, showing a positive evolution of the diesel quality distributed nationwide.

Resumen

Este estudio explora el comportamiento y evolución de la calidad del diésel automotriz que se distribuyó en Costa Rica durante el periodo 2006-2010. Tradicionalmente el estudio del comportamiento de los combustibles se basa en métodos estadísticos univariantes y bivariantes que no logran proporcionar información suficiente para describir las relaciones entre los atributos y el comportamiento multidimensional de los individuos. Para solventar esta limitante, se realizó un análisis en componentes principales (A.C.P.) en la tabla de datos que reúnen los cinco años del estudio. Adicionalmente, se realizó un A.C.P. en cada uno de los años para determinar la evolución del diésel. Con base en el análisis se encuentra que la concentración de azufre es la variable que más influye en el comportamiento de las otras variables, pero debido a su progresiva eliminación en el combustible, su influencia se anula en los dos últimos años del estudio. Al mismo tiempo, las temperaturas de destilación del combustible disminuyen, evidenciando una evolución positiva en la calidad del diésel distribuido nacionalmente.

Key words: Diesel, principal components, PCA, CELEQ, sulfur, distillation temperatures, cetane, flash point, viscosity, density

Palabras clave: Diésel, componentes principales, ACP, CELEQ, azufre, temperaturas de destilación, cetano, temperatura de inflamación, viscosidad, densidad

* INVID3.CELEQ@ucr.ac.cr

I. INTRODUCCIÓN

Desde el 2001, el Centro de investigación en Electroquímica y Energía Química (CELEQ) es el encargado de la evaluación de la calidad de los combustibles y la verificación de la calibración de los surtidores utilizados en las estaciones de servicio del país. Pero es hasta el año 2006 cuando el “Programa de Verificación de la Calidad de los Combustibles Suministrados por las Estaciones de Servicio Del País” empieza formalmente a registrar los resultados de cada uno de los análisis fisicoquímicos realizados a los combustibles.

Aunque el CELEQ mantiene la base de datos actualizada y realiza ciertos análisis de datos sobre los resultados de las variables fisicoquímicas, éste se circunscribe a un estudio unidimensional sobre cada variable. Esto limita la identificación de interacciones entre las distintas variables y sus posibles relaciones con otras variables suplementarias. Por ejemplo, dentro del estudio unidimensional no se consideran los individuos (estaciones), ya que se toman promedios anuales entre todas las entradas, de manera que la información recolectada se diluye.

El presente estudio es un complemento al estudio estadístico realizado para los primeros cinco años de registro denominado “Estudio Estadístico 2006-Noviembre 2010” [1]. El cual busca extender el análisis de los resultados describiendo tanto las relaciones entre las variables, como el comportamiento multidimensional de los individuos respecto a todas las variables en su conjunto, para ello se utiliza el Análisis en Componentes Principales (A.C.P.).

Análisis en componentes principales

El análisis en componentes principales (A.C.P.) es una técnica multivariante que analiza una tabla de datos en la que se describen las observaciones por varias variables dependientes cuantitativas inter-correlacionadas [2].

Su principal objetivo es el de encontrar, a partir de una tabla de datos con variables cuantitativas, un conjunto de nuevas variables sintéticas ortogonales denominadas componentes principales, cuya información sea lo más parecida a la de las variables originales. De esta forma que se logre visualizar el patrón de similitud de las observaciones y de las variables como los puntos en los mapas [3].

En general, cuando se utilizan tablas de datos de muchas variables e individuos, estas se pueden representar como una nube de puntos en espacios vectoriales de muchas dimensiones (espacios de dimensión de acuerdo al número de variables e individuos que se utilicen, digamos p y n respectivamente); sin embargo si la dimensión del espacio vectorial es superior a 3 ya no es posible visualizar las relaciones en la nube de puntos con el ojo humano. Por lo tanto, la idea central del A.C.P. es el de reducir la dimensionalidad del conjunto de datos, al tiempo que conserva tanto como sea posible de la variación presente en el conjunto [4].

Esto significa que hay una pérdida mínima de información al proyectar los individuos sobre un espacio de dimensión menor, de manera que la forma de la nube se asemeje lo mejor posible a su forma original [3].

II. MATERIALES Y MÉTODOS

Descripción de la base de datos

El CELEQ posee un repositorio de todos los registros del análisis de distintas variables fisicoquímicas que determinan la calidad de los combustibles (diésel, gasolina regular y gasolina superior), para cada una de las estaciones de servicio del país y con una frecuencia de tres veces al año.

Para el siguiente trabajo se utiliza la base de datos que contiene todas las estaciones visitadas en el periodo del 2006 al 2010 y que registra diez de las variables que determinan la calidad del diésel, las cuales son: Temperatura (Punto) Inicial de destilación, Temperatura de destilación al 10 %, Temperatura de destilación al 50 %, Temperatura de destilación al 90 %, Temperatura de destilación Final, Densidad o gravedad API, Temperatura de inflamación, Color, Índice cetano , Contenido de azufre, residuo carbón Conradson y viscosidad.

Pre-tratamiento de los datos

Por simplificación del estudio y debido a la complejidad de trabajar simultáneamente variables tanto cuantitativas continuas como cuantitativas discretas y cualitativas, se decide eliminar las segundas que involucren letras y las últimas en su totalidad, de la base de datos.

Por simplificación del estudio y debido a la complejidad de trabajar simultáneamente variables cuantitativas continuas, variables cuantitativas discretas y variables cualitativas, se decide eliminar las variables cuantitativas discretas que involucren letras y variables cualitativas de la base de datos.

De esta manera, las variables resultantes para análisis son las siguientes: Punto Inicial de destilación, destilación al 10 %, destilación al 50 %, destilación al 90 %, destilación Final, Densidad o gravedad API, Temperatura de inflamación, Contenido de azufre, viscosidad e índice de Cetano. En ocasiones, de la muestra de diesel analizada puede entregar resultados con valores menores a los cuantificables por el método aplicado, por lo tanto el laboratorio indica en el reporte y registra en la base de datos que el valor es inferior al límite cuantificable por el método. En estos casos, ya que no se tiene un valor numérico sino una inecuación, se decide eliminar la inecuación dejando un espacio vacío y sustituyéndolo por el valor del límite inferior cuantificable del método.

Pre-tratamiento de datos para el Análisis en Componentes Principales (A.C.P.)

Para este análisis es imprescindible tomar en cuenta a todos los individuos que cuenten con información sobre los parámetros de calidad de los combustibles, sin importar si éste no repite en todos los años del estudio o si solamente aparece en una ocasión. Por lo tanto, se utilizan todos los individuos que posean al menos un 70 % de información en sus atributos. Lo anterior significa que los individuos en la base de datos de diesel que son candidatos para el A.C.P., son todos aquellos que tengan al menos siete datos registrados de sus nueve variables de estudio. Los restantes individuos son descartados, ya que no aportan información suficiente para el estudio.

La selección de individuos anterior genera que los datos faltantes en las variables se reduzcan considerablemente. En el caso de la base de datos se tiene que las variables tienen valores superiores al 98 % de los datos, por lo tanto no es necesario eliminar ningún atributo por ausencia de información.

Los datos faltantes fueron completados con el promedio de los valores para los cinco años de estudio, es decir se calculó el promedio por atributo de los cinco. Esto se mantuvo así, incluso cuando se realizan los A.C.P. en cada año individual.

Los datos se centran y se estandarizan con el fin que todos utilicen una métrica M igual a la identidad de tamaño $p \times p$, donde p es el número de variables. Además, debido a que todos los individuos poseen el mismo peso, o sea todos los individuos tienen la misma ponderación, la métrica de pesos D es igual a la matriz diagonal de la inversa del número de individuos de cada matriz.

Para el Análisis en componentes Principales se utiliza el programa informático R versión 3.2.1, junto con los paquetes FactoMiner versión 1.31.3, rattle versión 3.4.1 y stats versión 3.2.1.

III. RESULTADOS Y DISCUSIÓN

Análisis en Componentes Principales para el Diésel entre los años 2006-2010:

Al realizar el A.C.P. en toda la matriz de datos, uniendo los cinco años en una sola matriz de tamaño 1660×9 , se debe determinar primero el número de componentes principales necesarias para proyectar la nube de datos en un espacio de menor dimensión y realizar el análisis sobre este subespacio dual creado. Para esto, se grafica en forma decreciente la inercia de cada componente en función del número de la componente y se escoge cuando este gráfico presente un cambio abrupto en la pendiente (el codo de la gráfica). De la Figura 1, se aprecia que a partir del segundo o tercer valor propio la curva presenta el codo, siendo este un indicativo de escogencia del número de componentes principales a usar. Además se obtiene que, el 76.83 % de la varianza total se encuentra en las primeras dos componentes. Asegúranos que, con solo las dos primeras componentes principales, existe una buena representación de los datos y que el restante 23.17 % de la información se reparte en las otras componentes encontradas.

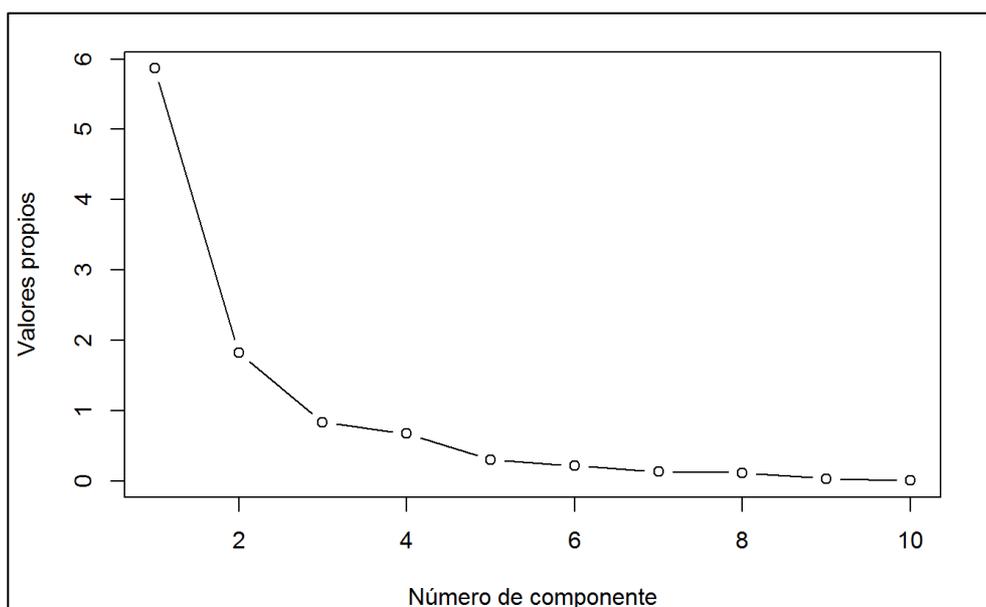


FIGURA 1. Gráfico de sedimentación para la tabla de datos que contiene los 5 años del estudio.

En la Figura 2 se aprecia el plano principal donde se proyecta la nube de datos en el espacio generado por las dos primeras componentes. En este caso, la primera componente separa los datos en tres regiones: los que tienen tendencia positiva, los que tienen tendencia negativa y los que están centrados alrededor del cero. La segunda componente, aún menos visible que la primera, divide a los datos en dos tendencias: los que están centrados cerca del cero con cierta tendencia positiva y los que tienen tendencia negativa.

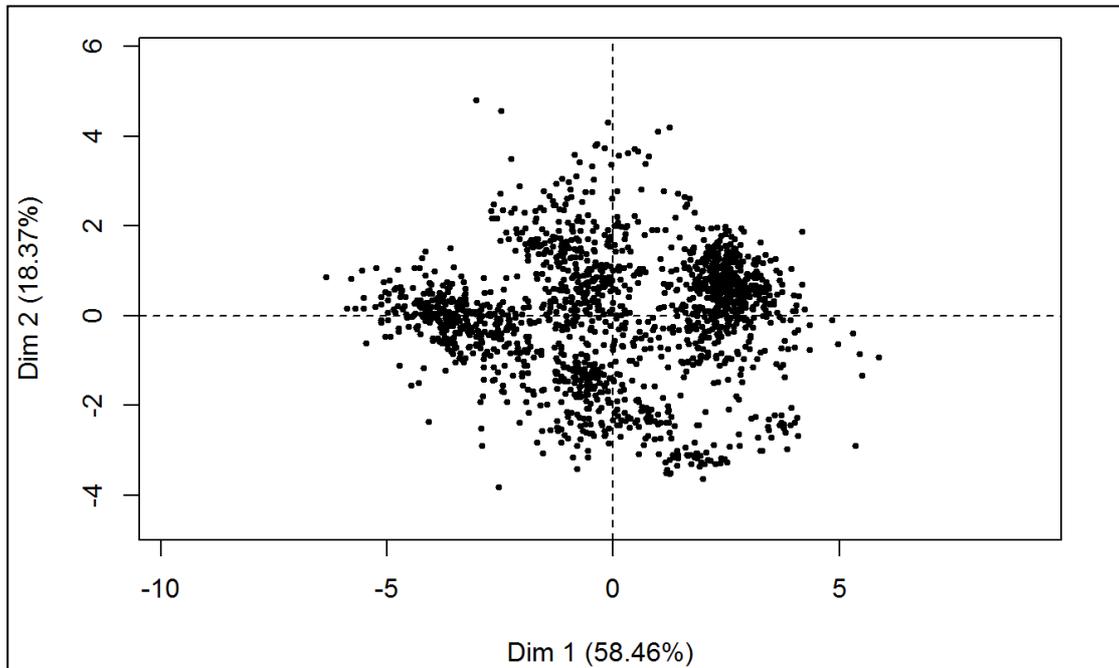


FIGURA 2. Plano principal generado por el ACP de los 5 años del estudio

Por otro lado, en la Figura 3, se muestra el círculo de correlaciones generado por las dos primeras componentes del A.C.P. para las diez variables estudiadas. En esta figura se encuentra que nueve de las diez variables presentan un claro efecto talla, donde la primera componente está fuertemente correlacionada con la magnitud de estas variables; de hecho todas estas variables tiene el mismo signo de correlación con la primera componente principal. La decima variable, el índice de cetano, presenta una correlación cercana a nula con respecto a la primera componente y por ende con el resto de las variables, debido a su ubicación cercana al eje positivo de la segunda componente. Esta segunda componente es la que determina la magnitud del índice de cetano y además hace la separación de los puntos iniciales y finales en la curva de destilación. También se puede observar en el círculo de correlaciones que las variables se concentran en cuatro grupos. El primer grupo reúne a la densidad, la temperatura de destilación al 10 % y la temperatura de destilación al 50 %; el segundo reúne a la temperatura inicial de la destilación, la temperatura de inflamación y la viscosidad, el tercer reúne a la cantidad de azufre, la temperatura de destilación al 90 % y la temperatura de destilación final. Y por último, el cuarto grupo está compuesto por el índice de cetano.

Del estudio de los anteriores grupos se obtiene que existe una fuerte correlación entre la densidad del diésel, la temperatura de destilación al 10 % y al 50 %, es decir se tiene que entre mayor sea la densidad, mayor va a resultar la temperatura de destilación al 10 % y al 50 %. Esto se puede explicar debido a que la densidad, como propiedad física fundamental de los combustibles, caracteriza las fracciones livianas y pesadas del petróleo y de los productos del petróleo. Entre más

denso sea el combustible, este puede estar compuesto por fracciones más pesadas de hidrocarburos, lo cual va a resultar en mayores temperaturas de destilación especialmente en las primeras fracciones de destilación. Conforme sea menor la densidad, ya sea debido a una composición por fracciones más ligeras de hidrocarburos o contaminados con gasolinas, menor será las temperaturas en las primeras fracciones de la destilación.

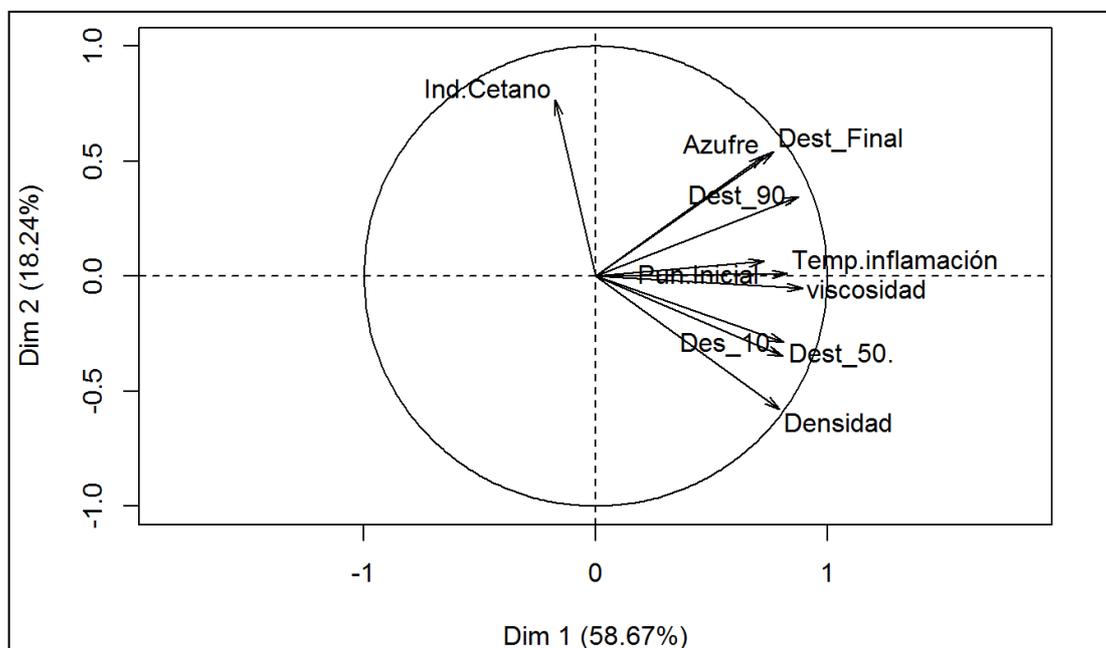


FIGURA 3. Círculo de correlaciones para las dos primeras componentes

Por otro lado, el segundo grupo de variables presenta un comportamiento similar al anterior grupo: se encuentra una correlación cercana a uno entre las variables viscosidad, punto inicial de destilación y la temperatura de inflamación. Así, entre mayor sea la viscosidad del diésel, mayor va a resultar la temperatura del punto inicial de destilación y la temperatura de inflamación. Estos dos primeros grupos presentan una correlación positiva, no tan cercano a la unidad como la correlación entre los atributos, pero lo suficiente para indicar que ambos fenómenos tienen el mismo comportamiento físico. Conforme aumenten la viscosidad o la densidad del diésel, mayor van a resultar las temperaturas de destilación iniciales (inicial, 10 % y 50 %) y la temperatura de inflamación del combustible.

El tercer grupo correlaciona las variables azufre con la temperatura de destilación al 90% y la temperatura de destilación final. Entre mayor sea la concentración de azufre en el diésel, mayor tenderá a ser la temperatura de destilación al 90 % y la temperatura final. Esto se puede deber a que el azufre medido se encuentra ligado de diferentes formas a cadenas de carbono y otros compuesto que componen al diésel, lo que hace que las fracciones con el azufre enlazado sean más pesadas (mayor peso molecular), en comparación con las mismas cadenas de carbón sin el azufre enlazado o ligado. Esto hace que estas fracciones más pesadas requieran mayor energía para lograr su evaporación, destilando inicialmente las fracciones más pequeñas y de menor peso molecular y dejando en los últimos puntos de la curva las fracciones más pesadas. Por lo tanto, entre mayor cantidad de azufre en estas fracciones, mayor será su peso molecular y mayor será la energía necesaria para su ebullición. Resultado altas temperaturas de destilación en los últimos puntos de la curva (ej. 90 % y final).

En la Figura 3 se puede observar una correlación muy débil (cercana a nula) entre el grupo uno y tres, lo cual indica que la densidad en diésel no se ve influenciada por la cantidad de azufre en las concentraciones presentes (de 0.010 % m/m a 0.5000 % m/m), ni que el azufre no afecta los primeros puntos de la curva de destilación.

El cuarto grupo está compuesto únicamente por el índice de cetano (una aproximación del rendimiento de encendido del combustible diésel) y se ubica de forma perpendicular con respecto al resto de las variables. Aunque físicamente el índice de cetano está influenciado por la composición del hidrocarburo y las características físicas del combustible (viscosidad, densidad, temperaturas de recuperación de destilación) [5], se observa una única dependencia con respecto a tres de las variables estudiadas.

Esta ubicación en el círculo de correlaciones se interpreta como una correlación nula con respecto a las variables del segundo y tercer grupo y una correlación negativa con las variables del primer grupo. Esto se debe a que este índice se obtiene a partir de una relación matemática empírica, que es función de la densidad y las mediciones de temperatura de recuperación de destilación. Su dirección opuesta se debe a que la densidad es la variable que más peso tiene en la ecuación y tiene un comportamiento inverso al índice de cetano (conforme aumenta la densidad disminuye el índice de cetano). La influencia de la densidad en la ecuación es tan fuerte que aunque exista una relación proporcional entre el cetano y las temperaturas de destilación al 10 % y al 50 % de volumen, el índice mantiene un comportamiento definido por la densidad, en este caso decreciente con forme se aumenta la densidad. Por lo tanto, las variables del primer grupo mantienen una correlación negativa con respecto al grupo cuatro.

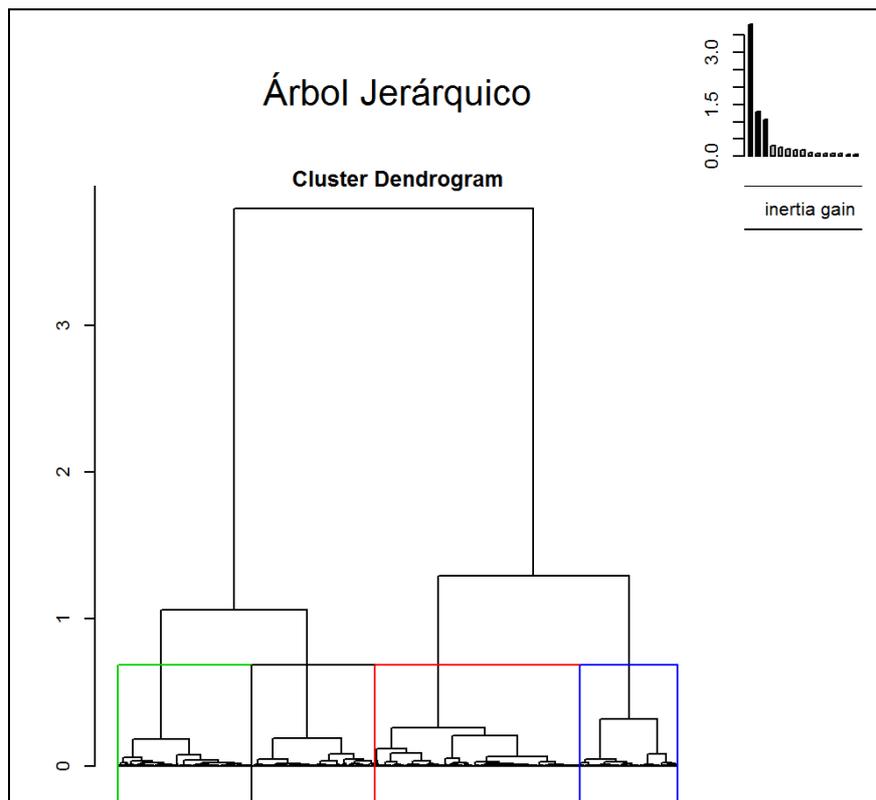


FIGURA 4. Dendrograma con Clasificación Jerárquica de cuatro clúster para los 5 años de estudio

Lo anterior nos proporciona una descripción de cómo se relacionan las variables, pero hasta el momento no conocemos como se comportan los individuos. En la Figura 1 se observan todos los 1660 individuos proyectados sobre el plano principal; al ser tantos puntos es difícil identificar el comportamiento individual, por lo que es necesario identificar grupos de individuos homogéneos, cuyas características iguales o similares hacen que estén próximos los unos a los otros, es decir individuos que se parecen entre sí.

Debido a la relativa poca cantidad de puntos y que es posible realizar una clasificación automática sin mayor dificultad, entonces se decide realizar un análisis de conglomerados a través de una clasificación jerárquica ascendente usando el método de Ward y la distancia Euclídea clásica, la cual se muestra en la Figura 4.

Para este caso, es claro que el árbol se puede “cortar” en tres o cuatro conglomerados, pero observando la Figura 2, donde se aprecia que existen cuatro regiones bien definidas por las dos primeras componentes principales, se decide realizar el corte en cuatro clases (clúster).

En la Figura 5 se representa el plano principal con los cuatro conglomerados. En esta se puede observar, tal como se indicó en la descripción de la Figura 2, que los clústeres están determinados por la dirección de las dos primeras componentes principales, siendo la primera componente la que más influye en el comportamiento y en la separación de los datos, principalmente al dividir los datos en tres regiones. Esto concuerda con el hecho que esta componente contiene cerca del 58 % de la inercia de la nube.

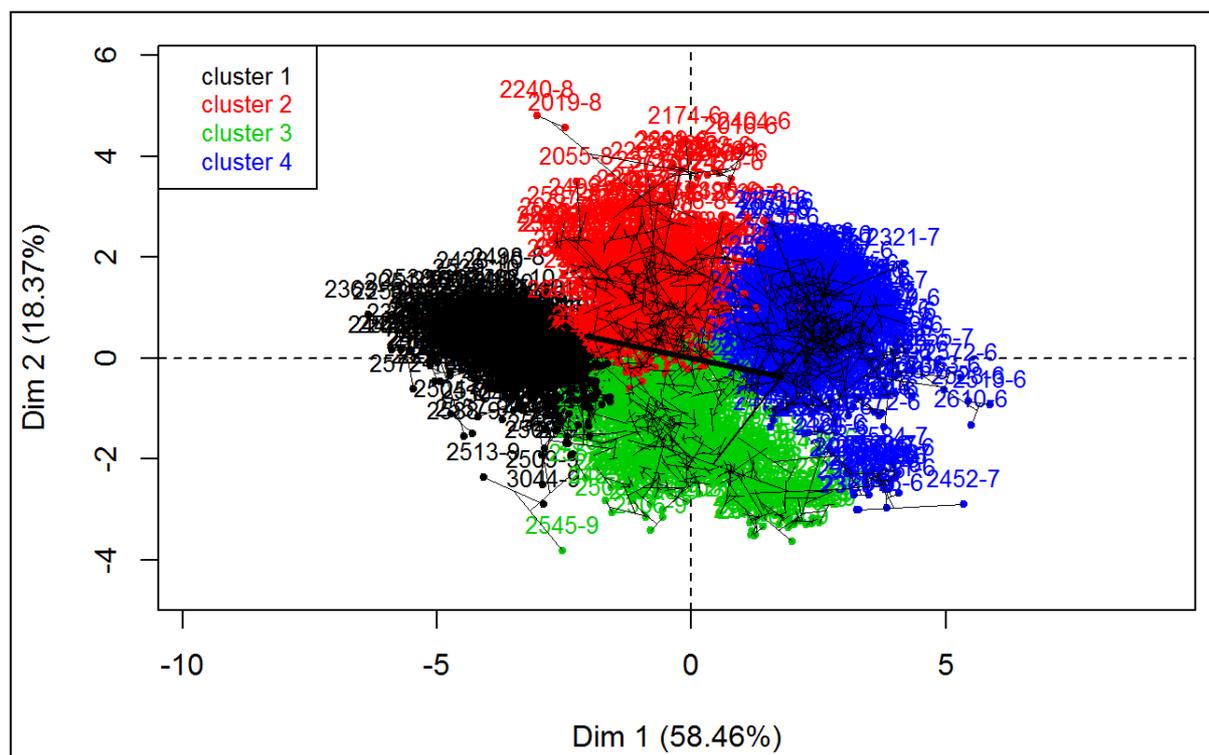


FIGURA 5. Plano principal indicando los cuatro aglomerados encontrados

El Cuadro 1 presenta los centros de gravedad de cada uno de los conglomerados, esto es el valor promedio en cada uno de los clúster. En esta tabla, se muestra que el clúster 1 es el que presenta los valores más bajos en todas las variables, mientras que en el clúster 4, es donde se obtienen los valores más altos en la mayoría de las variables, a excepción de la temperatura al 50 % de la destilación y la densidad del diésel. Los clúster 2 y 3 son los que tienen los valores

CALIDAD DEL DIESEL DE COSTA RICA ENTRE LOS AÑOS 2006-2010

intermedio, a excepción del clúster 3, que presenta los valores más altos en la temperatura al 50 % de la destilación y la densidad del diésel. Es clúster 2 es el que presenta los valores más cercanos al promedio, esto se observa por su cercanía al eje de coordenadas en la Figura 5.

CUADRO 1. Centros de gravedad para las clases encontradas por la clasificación jerárquica ascendente

Clase	Punto Inicial (°C)	Dest al 10% (°C)	Dest al 50% (°C)	Dest al 90% (°C)	Dest Final (°C)	Densidad (kg/m ³)	Temperatura inflamación (°C)	Azufre (%)	Viscosidad (cSt)	Índice de cetano (adim)
Clúster 1	174.0	211.8	269.8	325.9	350.5	850.7	61.93	0.040	2.729	46.629
Clúster 2	175.3	213.8	276.0	335.9	363.7	851.6	64.867	0.161	2.948	47.595
Clúster 3	177.6	222.0	285.1	333.6	354.8	861.1	66.689	0.049	3.069	45.621
Clúster 4	184.3	223.6	282.8	341.5	367.1	859.1	71.201	0.295	3.202	46.630

Al buscar otra característica distintiva, se observa que cada clúster agrupa a cada uno de los años del estudio, a excepción de los años 2006 y 2007 que se agrupan en un solo clúster. En la Cuadro 2 se obtiene la distribución en la frecuencia sobre la coincidencia de los años en los cuatro conglomerados encontrados en la clasificación jerárquica. Cada clúster recibe más del 83 % de los datos por cada año del estudio, por lo que se puede interpretar como el reflejo del comportamiento a través de los años con respecto a las variables del estudio para este conjunto de clúster.

CUADRO 2. Distribución en las frecuencias de coincidencia de los años del estudio en los cuatro clúster encontrados

AÑO	Clúster 1	Clúster 2	Clúster 3	Clúster 4	TOTAL
2006	0	37	4	284	325
2007	0	19	13	305	337
2008	4	307	26	4	341
2009	50	4	270	2	326
2010	325	5	1	0	331
Total	379	372	314	595	1660

En la Figura 6 se observa el dendrograma obtenido a partir de la clasificación jerárquica sobre el plano principal de los individuos. De acuerdo a lo anterior, el grupo de color azul reuniría a los años 2006 y 2007, el grupo rojo reuniría al año 2008, el verde al año 2009 y finalmente el grupo negro reúne al año 2010.

Este comportamiento indica que cada uno de los años del estudio está caracterizado por un cambio en los parámetros de calidad que se miden en el diésel. Por ejemplo: los años 2006 y 2007 (clúster 4) se caracterizan por altos valores en todas las variables del estudio, excepto en la densidad y en la destilación al 50 % (aunque ambos valores se encuentran muy cercanos a valores máximos encontrados en los centros de gravedad), destacando para estos dos años un alto contenido en azufre, viscosidad, índice de cetano, y altas temperaturas en los puntos de destilación y la temperatura de inflamación, situación que coincide con el programa de reducción del contenido de azufre en diésel a nivel Nacional, propuesto en el Decreto Ejecutivo N° 26130-MINAE del 21 de abril de 1997 y su posterior modificación en el Decreto Ejecutivo N° 30690-MINAE del 17 de setiembre de 2002. Este alto contenido de azufre influye en los altos valores encontrados en el resto de las variables, fenómeno que también observa el círculo de correlaciones descrito anteriormente.

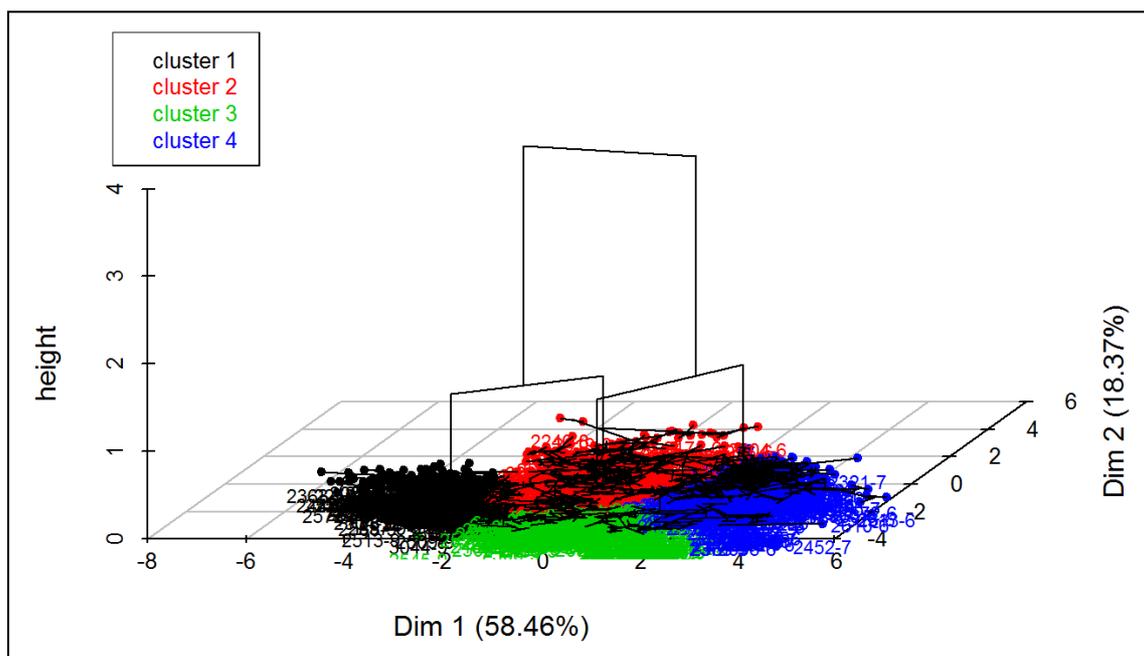


FIGURA 6. Plano principal con árbol jerárquico indicando los cuatro aglomerados encontrados

El año 2008 se considera como el año promedio, debido a que se sitúa muy cerca del origen en el plano principal. Las variables toman valores promedio con respecto al periodo de estudio y este año es cuando se presenta la primera disminución, cerca del un 50 %, de azufre en el diésel. Todas las variables disminuyen sus valores con respecto al periodo anterior, excepto el índice de cetano.

Para el 2009 se presenta un aumento en las temperaturas iniciales y medias de destilación, la densidad, la viscosidad y la temperatura de inflamación con respecto al año anterior, volviendo a valores cercanos a los obtenidos durante el periodo 2006-2007; pero se encuentra una disminución en el índice de cetano y las temperaturas finales de la destilación. Este último relacionado principalmente por la segunda disminución, de un 64 %, en la concentración de azufre en el diésel (sufrida entre el 2008 y el 2009).

Para el periodo del 2010 se registra la mayor disminución en los valores de todas las variables del estudio, excepto por el índice de cetano. Es además, donde se registra el contenido más bajo de azufre en el diésel, disminuyendo por tercera vez aproximadamente 30 % con respecto al año anterior y muy por debajo del indicado en el "Reglamento Técnico Centroamericano RTCA 75.02.17.06 Productos de Petróleo, Aceite Combustible Diésel, Especificaciones" (donde se indica que el contenido de azufre no puede ser mayor a 0,5 % m/m) [6]. En general los índices para este último año son indicativos del acceso a un mejor combustible para la distribución nacional. Por ejemplo: la continua disminución de la concentración de azufre en el combustible, aunque podría afectar el desgaste y los depósitos en los motores, tiene una incidencia directa en el control de emisiones de los vehículos, disminuyendo las emisiones por dióxido de azufre [7]. La temperatura de inflamación aunque ha disminuido, se mantiene por encima del valor normado nacionalmente (52 °C como mínimo).

El índice de cetano oscila entre 45 y 48 unidades en el promedio de los clúster, ya que no se ve afectado por los cambios en la concentración de azufre y se mantiene dentro de lo estipulado en el Reglamento Técnico Centroamericano. La disminución de las temperaturas de destilación ayuda a que los motores en los servicios relacionados con rápidas fluctuaciones de cargas y velocidades,

CALIDAD DEL DIESEL DE COSTA RICA ENTRE LOS AÑOS 2006-2010

como la operación de autobuses y camiones, puedan generar un mejor rendimiento, en particular con respecto al humo y el olor [7].

Al realizar el análisis en componentes principales a cada uno de los años del estudio, Figuras 7 y 8, se encuentran muchos de los comportamientos observados en el ACP de los cinco años, específicamente se conservan varias de las relaciones o grupos entre las variables pero en diferente orden. Por ejemplo, el grupo con las variables punto de inicial de destilación, la temperatura de inflamación y viscosidad del combustible y el grupo de la densidad y temperatura de destilación al 10 % y al 50 % se mantienen como un grupo en común en los cinco años del estudio, mientras que las variables azufre, temperatura de destilación al 90 % de recuperación y destilación final se presentan como un grupo pero solo en los tres primeros años (2005-2007).

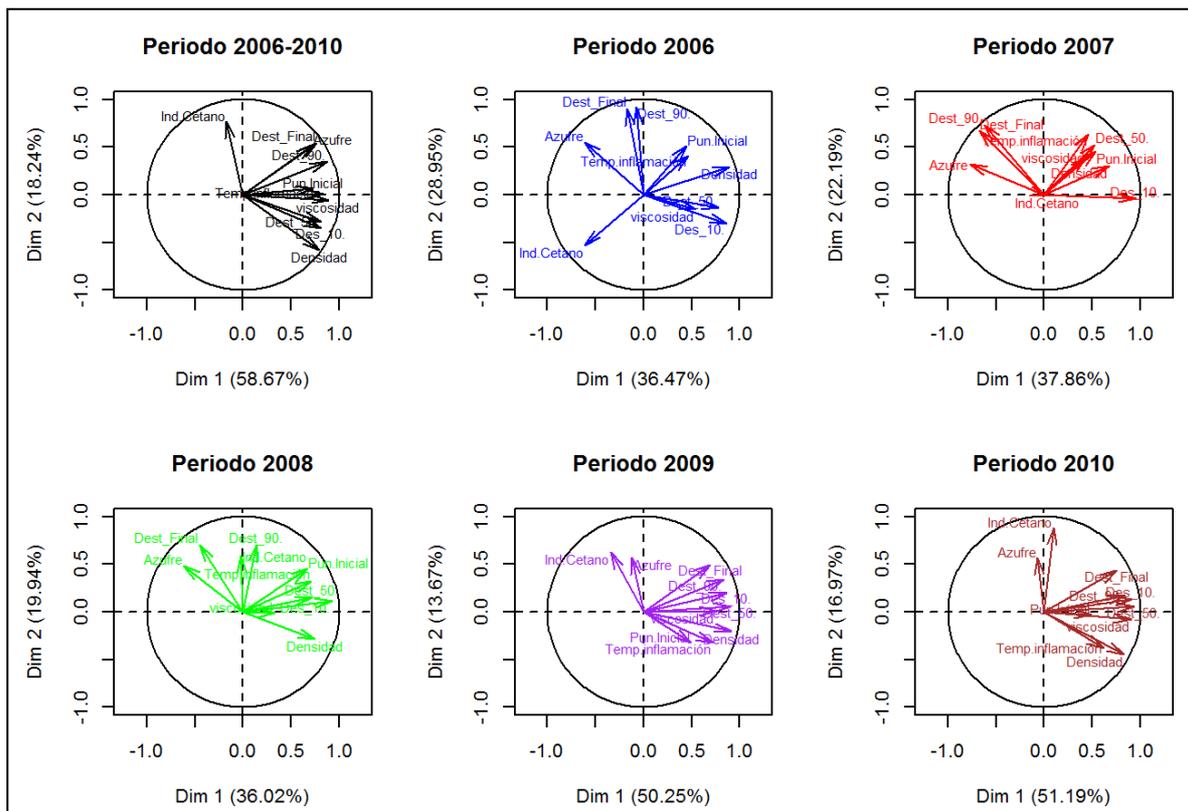


FIGURA 7. Círculos de correlaciones para los años 2006, 2006, 2007, 2008, 2009 y 2010

Es interesante observar como la segunda componente para los primeros tres años (años 2006, 2007 y 2008) es mucho mayor que la inercia de la segunda componente en el ACP inicial, y que en los análisis para los años 2009 y 2010. Es decir, las variables se representan con una mayor dispersión de los puntos (ver Figura 8) y aunque los grupos de cierta forma se mantienen, estas variables se ubican en todos los cuadrantes del plano principal. Esto significa que más de una componente se reparte la inercia que describen a los datos, a diferencia del ACP original y los años 2009 y 2010. Lo anterior se observa en los gráficos de sedimentación en la Figura 9, donde es claro que para los años 2006, 2007 y 2008, la mayoría de la inercia está distribuida en más de dos componentes. Pero esta inercia en la segunda componente sufre una disminución de su magnitud en los años 2009 y 2010, concentrándose en la primera y recuperando en cierta medida el efecto talla sobre las variables.

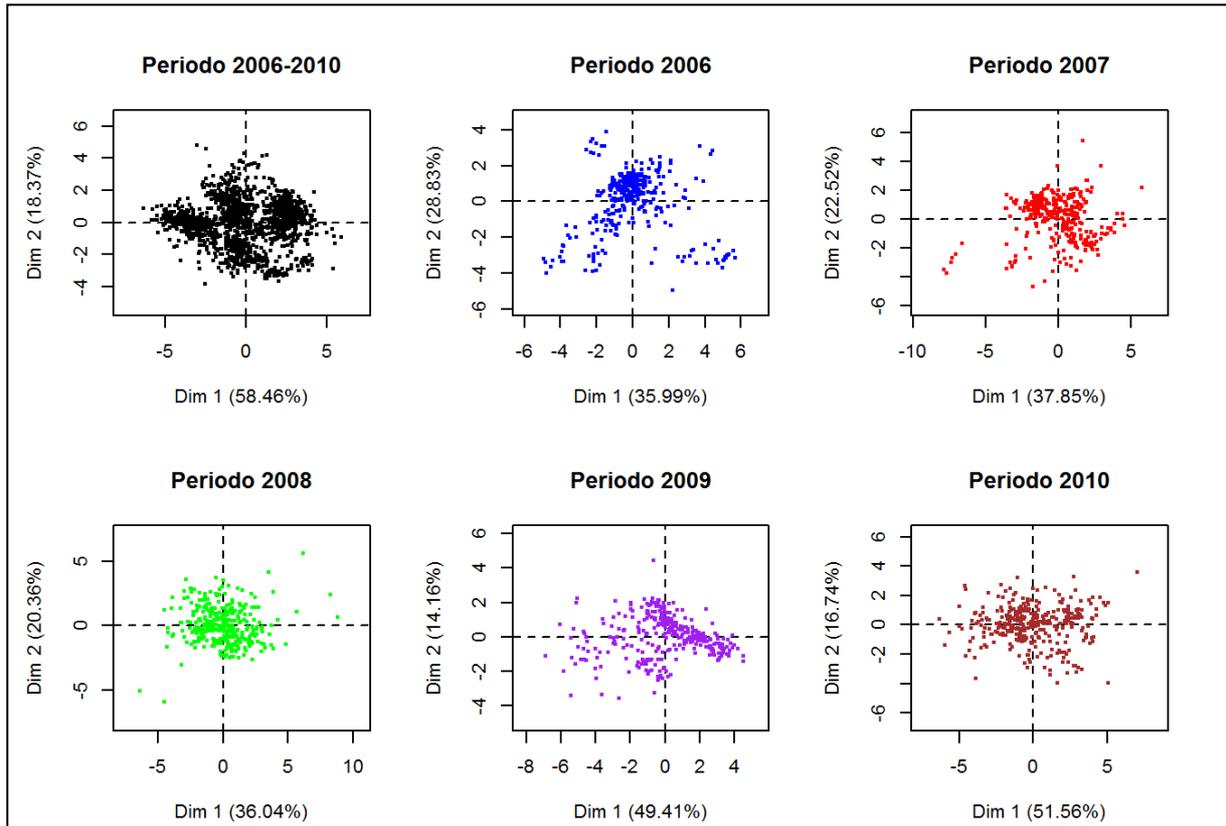


FIGURA 8. Planos principales para los años 2006, 2006, 2007, 2008, 2009 y 2010

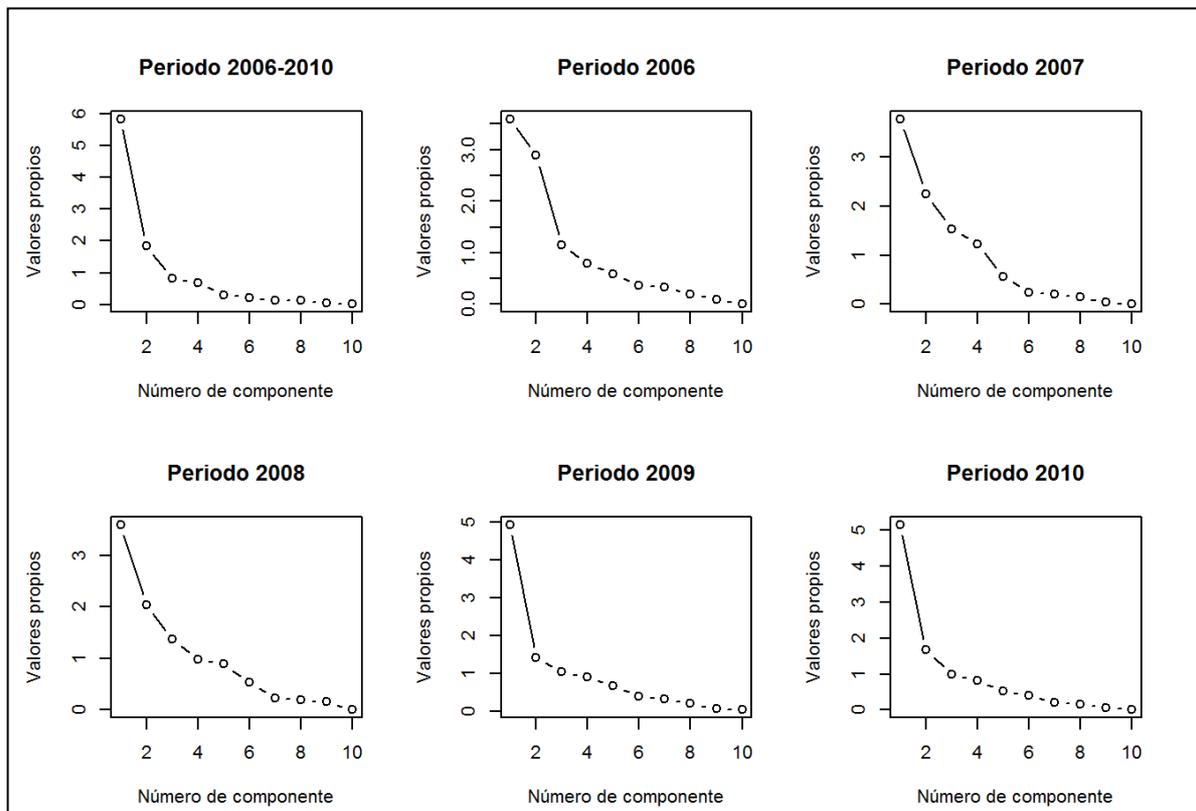


FIGURA 9 Gráficos de sedimentación para los cinco años de estudio

CALIDAD DEL DIESEL DE COSTA RICA ENTRE LOS AÑOS 2006-2010

Lo anterior se ve reflejado en como los grupos se van reuniendo cada vez más hasta crear un efecto talla con la primera componente y deja las variables azufre e índice de cetano de forma perpendicular con respecto a la primera componente, sobre el eje de la segunda. Esto es un indicativo (por lo menos en el caso del azufre) que las concentraciones presentes en el diésel, debido a las reducciones programadas, no son lo suficientemente altas ni presentan la variabilidad necesaria para tener una influencia sobre la matriz del combustible. En los dos últimos años del presente estudio se registra un coeficiente de variación para el azufre del 11 % y 9 % para el 2009 y el 2010, mientras que para los años 2006, 2007 y 2008, donde se registran coeficientes de variación de 19 %, 26 % y 32 % respectivamente.

Esto implica que, durante estos dos últimos años se importó diésel con un contenido bajo y estable de azufre, estos valores podrían estar por debajo del nivel en el cuál este vaya a influenciar el comportamiento del resto de las variables. Hay que tener presente que los valores registrados para el azufre se encuentran cerca del límite de cuantificación del método utilizado en el CELEQ, por lo que esto puede explicar su poca variación en los resultados [1].

Como se determinó que la concentración de azufre y el índice de cetano no tienen influencias significativas sobre las otras variables, en los años 2009 y 2010, se decide eliminarlas para observar si existe alguna modificación en el comportamiento de las restantes variables, ver Figura 10. Al hacer esto, se encuentra que el efecto talla se conserva en las 8 variables que definen el índice de calidad del diesel.

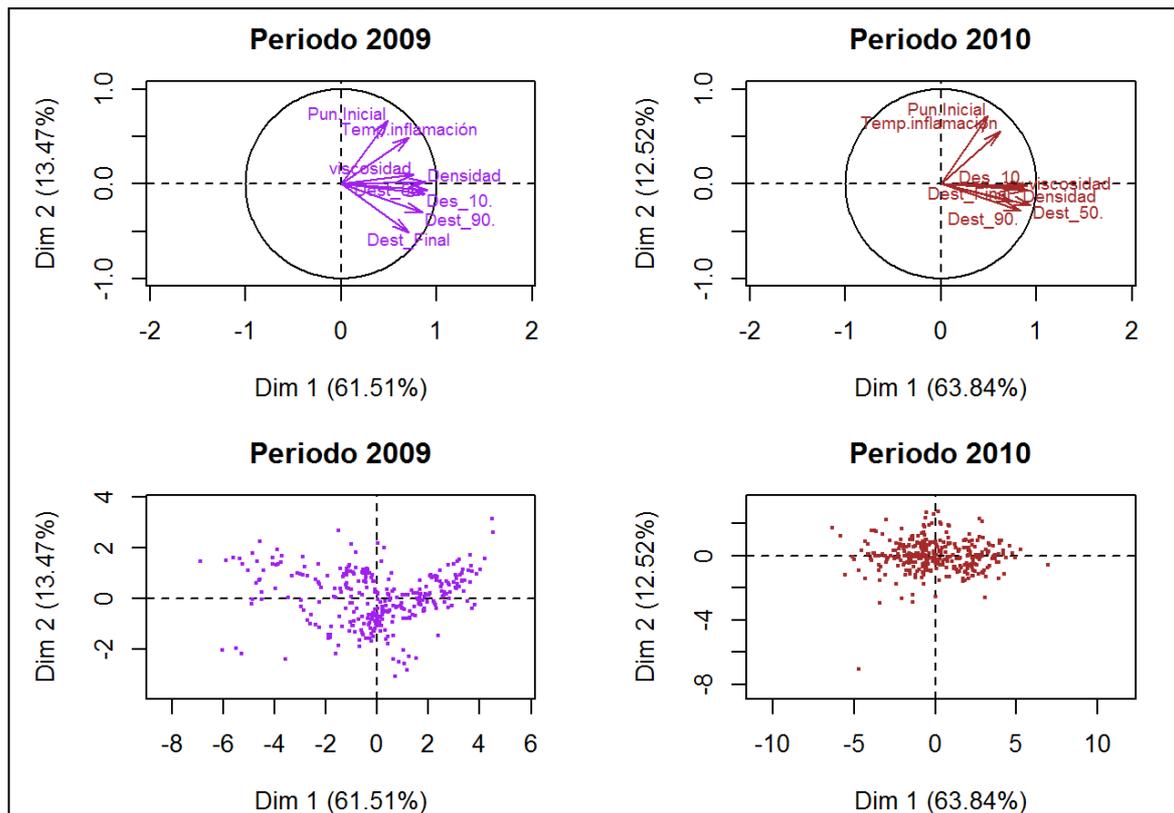


FIGURA 10. Círculos de correlaciones y Planos principales para los años 2009 y 2010

La primera componente correlaciona fuertemente a 6 de los atributos generando un grupo grande de variables mientras que la segunda componente separa a la temperatura del punto inicial de la destilación y la temperatura de inflamación de este grupo. Cabe rescatar que se mantienen

algunos de los comportamientos encontrados anteriormente (Figura 3), por ejemplo las variables se separan en dos subgrupos bien definidos y se mantiene la correlación positiva entre la temperatura inicial de la destilación y la temperatura de inflamación del diésel. Pero contrario a lo encontrado anteriormente este subgrupo (Punto inicial y temperatura de inflamación) presentan una correlación muy pobre (casi nula) con el resto de las variables. Desconectándose del resto de las variables a diferencia de lo presentado en los años 2007 y 2008.

IV. CONCLUSIONES

- En general se encuentra que existe una fuerte correlación entre la concentración de azufre y los puntos finales de destilación, pero esta tiende a anularse cuando la concentración del azufre tiende a disminuir en el combustible que se comercializa.
- Hay una clara correlación entre el punto de inflamación del diésel y la temperatura inicial de destilación, fenómeno que se observa a lo largo de todo el estudio y no se ve afectada por la cantidad de azufre en el diésel ni por el comportamiento que tenga los puntos finales de la destilación. Esto es debido a que la temperatura de destilación inicial y la temperatura de inflamación están dominadas por la presencia de fracciones livianas de combustible y su variación es indicativa de una contaminación con productos livianos, tales como gasolina. Esta situación no influye en los puntos finales de la destilación.
- Se encuentra un efecto interesante en el comportamiento entre las temperaturas de los distintos puntos de destilación y la concentración de azufre en el combustible. Cuando existe relativamente altas concentraciones, por ejemplo durante los años 2006, 2007 y 2008, los puntos de destilación al 10 % y al 50 % y los puntos finales de la destilación presentan una desconexión en su comportamiento, por lo que el ángulo de separación entre sus vectores es aproximadamente recto (Figura 7). Pero en el momento en que se disminuye la concentración de azufre, valores debajo del 0,01 %, en los años 2009 y 2010, esta desconexión desaparece y se correlacionan positivamente, siguiendo la región comportamiento tipo talla (Figura 10).
- Se encuentra que durante el periodo de estudio, existe una evolución positiva en la calidad del diésel distribuido en Costa Rica. Se tiene un combustible de calidad inferior en los tres primeros años del estudio, pero mejora en sus índices en los dos últimos años del estudio. Esta tendencia se da por año de estudio y logra un máximo en el 2010.

V. REFERENCIAS

- [1] León Rojas. C.; Zuñiga, P., ESTUDIO ESTADÍSTICO 2006-Noviembre 2010. Universidad de Costa Rica: Ciudad de la Investigación, 2011.
- [2] Abdi H.; Williams L.J., Wiley Interdiscip. Rev. Comput. Stat., **2010**, 2, 433–459.
- [3] Trejos, J., Castillo, W.; González, W., Analisis Multivariado de datos, Métodos y aplicaciones, Editorial Universidad de Costa Rica, 2014, pp 37-89.
- [4] Jolliffe, I. T., Principal Component Analysis, ed. Springer-Verlag: New York, 2002, pp 10-28.
- [5] Totten, G., Fuels and Lubricants Hand Book: technology, Properties, Performance and Testing. ed. ASTM International: West Conshohocken, 2003, pp 115-139.
- [6] Costa Rica, Reglamento Técnico Centroamericano RTCA 75.02.17:06 Productos de Petróleo. Aceite Combustible Diesel. Especificaciones, La Gaceta, 30 de marzo de 2007, núm 64.
- [7] ASTM D975-15a, Standard Specification for Diesel Fuel Oils, ASTM International, West Conshohocken, PA, 2015, <http://dx.doi.org/10.1520/D0975-15C>.

AGRADECIMIENTO

Un sincero agradecimiento a los colegas del Centro de Electroquímica y Energía Química (CELEQ) quienes con su experiencia proporcionaron una ayuda invaluable a la investigación. A los profesores Dr. Javier Trejos y al Dr. Oldemar Rodríguez de la Escuela de Matemáticas de la UCR, por su colaboración en la interpretación de los resultados del ACP.

Y finalmente al Dr. Carlos León por su apoyo y respaldo en la realización de la presente investigación.