

## CALIDAD DE LA GASOLINA REGULAR DE COSTA RICA ENTRE LOS AÑOS 2006-2010

Víctor Bazán Salazar\*

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

*Recibido abril 2016; aceptado mayo 2017*

### Abstract

This study explores the behavior and evolution on the quality of motor regular gasoline 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 regular gasoline' 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, an automatic classification by an ascending hierarchical classification was performed to determine the behavior of regular gasoline when forming groups with similar characteristics. Based on the analysis, it is found that regular gasoline evolves toward a better fuel within the current regulations, but this evolution it's done in a oscillatory manner during the study period, where the main improvement is the reduction in olefins content in the fuel.

### Resumen

Este estudio explora el comportamiento y evolución de la calidad de la gasolina regular 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.) con los datos que reúnen los cinco años del estudio. Además se realizó una clasificación automática mediante una clasificación jerárquica ascendente para determinar el comportamiento de los datos a la hora de formar grupos con características afines. El análisis de los resultados demuestra que la gasolina regular evoluciona hacia un mejor combustible, con respecto a la reglamentación vigente, pero esta evolución la hace de manera oscilatoria durante el periodo de estudio, donde la principal mejora es en la disminución en el contenido de olefinas en la gasolina.

**Key words:** Regular gasoline, PCA, fuel quality, CELEQ,

**Palabras clave:** Gasolina regular, ACP, calidad de combustibles, CELEQ.

## I. INTRODUCCIÓN

Durante el periodo del 2006 al 2010 las estaciones de servicio movilizaron aproximadamente 60 millones de barriles de producto refinado (gasolina superior, gasolina regular y diesel); la gasolina regular representa el 30 % de este volumen. Esto significa que cerca de 1.46 billones de colones se movilizaron por ventas de este combustible en este periodo de 5 años.[1]

---

\* INVID3.CELEQ@ucr.ac.cr

Debido al alto volumen de hidrocarburos que se distribuyen en el país y a que hay una afectación directa en el sector productivo nacional y en general a todos los consumidores finales de estos productos, el Centro de Investigación en Electroquímica y Energía Química (CELEQ) es el encargado de la ejecución del “Programa de Verificación de la Calidad de los Combustibles Suministrados por las Estaciones de Servicio Del País”, desde el año 2001. Esto significa que el Centro es el responsable de la evaluación de todas las variables fisicoquímicas que afectan la calidad de los combustibles que las estaciones de servicio distribuyen en el país y de verificar la calibración de todos los surtidores (mangueras) que estas empresas utilizan. De esta manera, el Centro genera continuamente información acerca de la calidad de los combustibles expendidos en las estaciones de servicio, formando un registro de cada uno de los parámetros evaluados, los cuales se guarda en su propia base de datos.

Dado que el volumen de información contenida en la base de datos del CELEQ es tan cuantioso, si se limita el estudio a un análisis tradicional de datos donde se aborda de manera unidimensional sobre cada variable, se podría restringir la identificación de interacciones entre las distintas variables y otras relaciones entre individuos o variables. Por ejemplo, dentro del estudio unidimensional tradicional no se consideran los individuos (gasolina regular de las estaciones), ya que se toman promedios anuales entre todas las entradas; de esta forma, la información recolectada se diluye e imposibilita identificar comportamientos entre individuos. Por lo tanto, en esos casos, puede generar mayor información si se aplican técnicas multivariadas para analizar la tabla de datos de forma multidimensional. Uno de estos métodos multivariados es el análisis en componentes principales (A.C.P.) donde el objetivo es encontrar 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 logra visualizar el patrón de similitud de las observaciones y de las variables como los puntos en los mapas [2].

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 reducir la dimensionalidad del conjunto de datos, al tiempo que conserva tanto como sea posible de la variación presente en el conjunto [3]. 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 [2].

El presente estudio es un complemento al estudio estadístico realizado para los primeros cinco años de registro denominado “Estudio Estadístico 2006-Noviembre 2010”[4] y busca extender el análisis de los resultados describiendo tanto las relaciones entre las variables, como el comportamiento multidimensional de la gasolina regular con respecto a todas las variables en su conjunto, mediante el uso del Análisis en Componentes Principales (A.C.P.).

## II. MATERIALES Y MÉTODOS

### Descripción de la base de datos

El CELEQ posee un repositorio de todos los resultados 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 que se obtienen con una frecuencia de tres veces al año.

Para este trabajo, se utiliza la base de datos que contiene los resultados de los parámetros de calidad de la gasolina regular para todas las estaciones visitadas en el periodo del 2006 al 2010 donde se registran las siguientes quince variables: punto inicial de destilación, destilación al 10%, destilación al 50%, destilación al 90%, destilación final, residuo de la destilación, presión vapor Reid, densidad o gravedad API, octanaje, oxigenados, benceno, aromáticos, olefinas, agua y sedimento e índice de manejabilidad ordinario (ODI). Esta última es calculada a partir de los valores obtenidos para la destilación al 10%, destilación al 50%, destilación al 90% y tiene como objetivo proporcionar control de los parámetros de destilación que influyen en el arranque en frío y la facilidad de conducción en caliente. Aunque la ecuación expuesta en la norma ASTM D4814[5] involucra una corrección por el contenido de alcohol en el combustibles, esta no se utiliza para el presente estudio, ya que no se cuenta con este valor en las estaciones donde hubo mezcla con etanol y la intención es observar de forma global cómo se comporta este índice en todo el país. Por lo tanto, para el presente estudio se usa la ecuación expuesta en el Manual para Combustible y Lubricantes de George E. Totten.[6]

### **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. 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, residuo de la destilación, presión vapor Reid, densidad o gravedad API, octanaje, oxigenados, benceno, aromáticos, Olefinas e índice de manejabilidad ordinario (ODI).

En ocasiones, la muestra de combustible analizada puede generar resultados con valores menores a los cuantificables por el método aplicado; cuando esto ocurre, 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, se decide eliminar la inecuación sustituyéndola 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 se repiten en todos los años del estudio o si solamente aparecen 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 presentes en la base de datos de gasolina regular que son candidatos para el A.C.P., son todos aquellos que tengan al menos nueve datos registrados de sus trece variables de estudio. La variable ODI calculada por las otras variables, por lo que esta no se toma en cuenta como presente. Los restantes individuos son descartados, ya que no aportan información suficiente para el estudio.

La selección de individuos genera que los datos faltantes en las variables se reduzcan considerablemente. En el caso de la base de datos, las variables tienen valores superiores al 99 % de los datos, por lo que 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 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[7] versión 3.3.0, junto con los paquetes FactoMiner[8] versión 1.32, rlg[9] versión 0.95.1441, fmsb[10] versión 0.5.2 y ggplot2[11] versión 2.1.0.

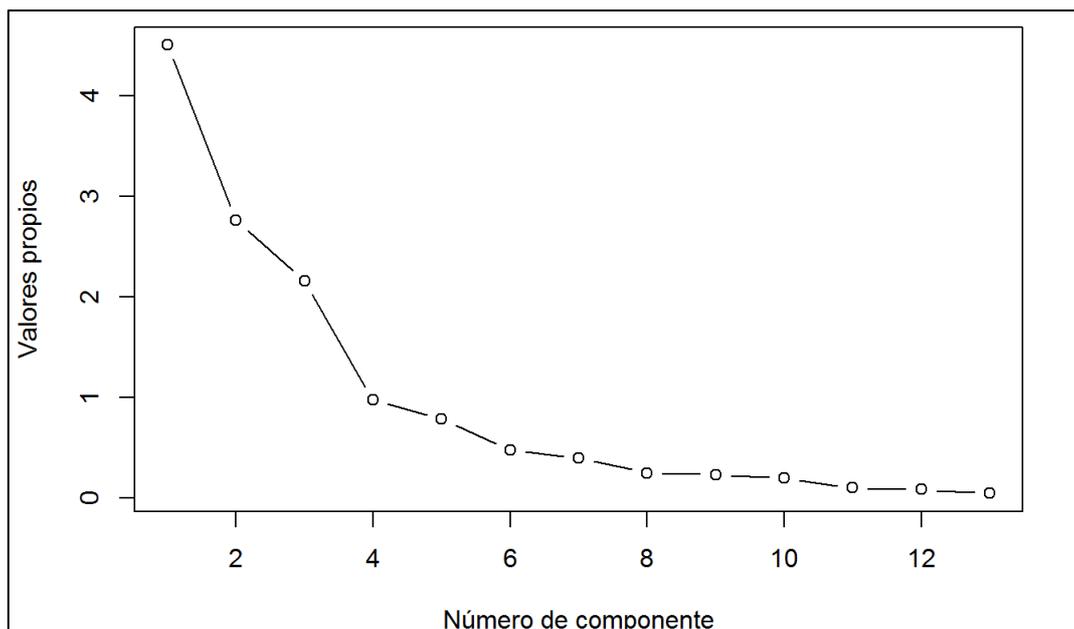
### III. RESULTADOS Y DISCUSIÓN

#### Análisis en Componentes Principales para la gasolina regular entre los años 2006-2010:

Para realizar una descripción adecuada de la tabla de datos para Gasolina Regular se hace uso del Análisis en Componentes Principales como herramienta para la reducción de las dimensiones; de esta forma se puede proyectar tanto los individuos (estaciones de servicio) como las variables sobre espacios de menor dimensión, los cuales son de fácil manipulación e interpretación de los datos. Para esto, después de realizar el cambio de base (el ACP), se debe determinar el número de componentes necesarias para representar la proyección de la nube de datos sobre un espacio de menor dimensión, tratando de maximizar la inercia de la nube de puntos proyectada sobre unas cuantas dimensiones. Por lo tanto, la inercia de cada componente en función del número de la componente se ordena en forma decreciente (ver Tabla 1) y se genera el gráfico de sedimentación para la gasolina regular, que se muestra en la Figura 1.

**TABLA 1.** Valores Propios y porcentajes de inercia para la tabla de gasolina regular

Componente	Valores propios	Inercia explicada (contribución)	Inercia acumulada
Comp. 1	5.757	41.12 %	41.12 %
Comp. 2	2.700	19.28 %	60.41 %
Comp. 3	2.069	14.78 %	75.19 %
Comp. 4	0.863	6.16 %	81.35 %
Comp. 5	0.730	5.22 %	86.57 %
Comp. 6	0.574	4.10 %	90.67 %
Comp. 7	0.359	2.56 %	93.23 %
Comp. 8	0.298	2.13 %	95.36 %
Comp. 9	0.209	1.49 %	96.85 %
Comp. 10	0.170	1.21 %	98.06 %
Comp. 11	0.119	0.85 %	98.91 %
Comp. 12	0.100	0.71 %	99.63 %
Comp. 13	0.052	0.37 %	100.00 %
Comp. 14	0.000	0.00 %	100.00 %



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

De acuerdo a la Figura 1, se observa que entre la tercera y cuarta componente se tiene la suficiente inercia para poder representar los datos; además, el cambio a partir de este punto es tan pequeño, que si se incorpora una componente más, no adicionaría mayor información para la descripción del sistema. A partir de esto y debido a que las tres primeras componentes obtienen valores propios mayores a uno (lo cual representa un 75 % del total de la inercia) se decide hacer el análisis con las tres primeras componentes.[12]

Una vez determinado el número de componente a utilizar, se puede realizar las representaciones graficas que describan el comportamiento de los datos, tal como se muestra en la Figura 2.

En la Figura 2, observa cómo se distribuyen las 14 variables usadas sobre los tres círculos de correlación formados a partir de la combinación de las primeras tres componentes principales. Para el círculo formado por la primera y segunda componente, las variables se distribuyen alrededor de todo el perímetro, pero se puede apreciar un grupo compuesto por las variables ODI, aromáticos, densidad, residuos de destilación y las temperaturas al 50 %, 90 %, y final de destilación, que poseen una fuerte correlación con respecto al eje positivo de la primera componente. En contraste, las variables olefinas y temperatura inicial de la destilación, poseen una fuerte correlación con respecto al eje negativo de la misma componente. Esto indica que un aumento en la cantidad de aromáticos en las gasolinas regular aumenta las temperaturas de destilación a partir del 50 % del volumen y produce un aumento en el residuo de la destilación, en el número de octanos y en la densidad del combustible. Todos estos fenómenos se correlacionan con el hecho de que los compuestos aromáticos son moléculas de alta densidad energética que aumentan el octanaje en la gasolina y que a la vez pueden aumentar los depósitos del motor y las emisiones del tubo de escape.[13]

Por otro lado, un aumento en la cantidad de olefinas, aumenta la temperatura inicial de la destilación, pero no logra tener un efecto significativo en la temperatura de destilación del 10 % del volumen recuperado. Un efecto interesante es la relación inversa que tiene las olefinas con el número de octano del combustible. "Las olefinas son hidrocarburos insaturados y, en muchos casos, también son buenos componentes de octano de gasolina." [13] Pero en el período 2006-2010

se presenta una tendencia decreciente en el contenido de olefinas, disminuyendo aproximadamente de un 20% en el 2006 a un 12% para 2010, opuesto al incremento en la cantidad de aromáticos durante el mismo periodo, de un 22% en el 2006 a un 27% para el 2010, lo cual aumenta el octanaje del combustible superponiéndose sobre el efecto de la reducción del contenido de olefinas. [4]

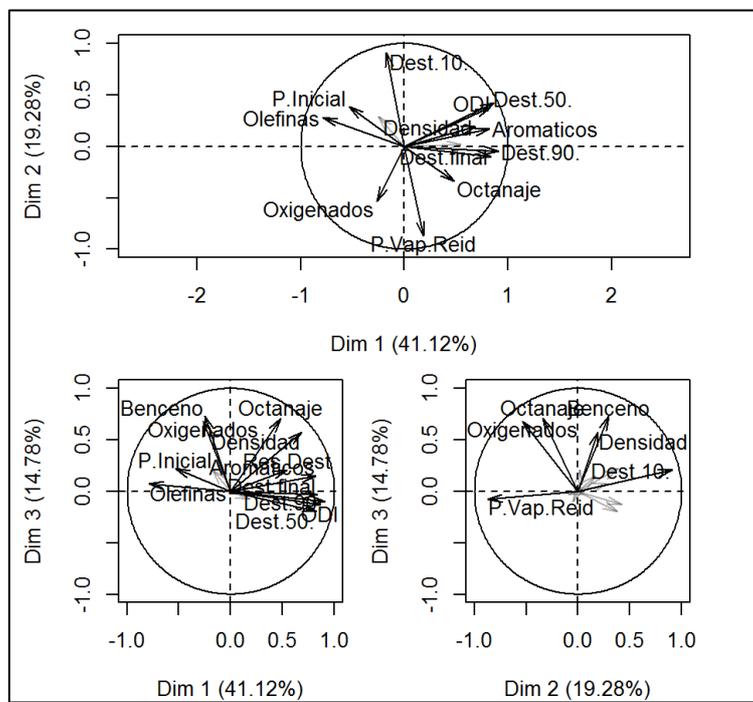


FIGURA 2. Círculos de correlaciones para las componentes 1-2, 1-3 Y 2-3

Ahora, la segunda componente principal tiene altas correlaciones (en valor absoluto) con las variables temperatura de destilación al 10 % del volumen y la Presión de Vapor Reid, pero con direcciones opuestas entre sí. Por lo tanto, la segunda componente principal opone estas dos medidas de volatilidad del combustible. Esto es de gran importancia para la gasolina ya que la temperatura de evaporación al 10% recuperado debe ser lo suficientemente baja como para asegurar el arranque a temperaturas normales pero al mismo tiempo la presión de vapor de combustible debe ser lo suficientemente alta como para asegurar la facilidad de arranque del motor. Esta última debe controlarse de tal forma que no sea tan alta como para contribuir a las emisiones de vapores de bloqueo o evaporación excesiva y las pérdidas en funcionamiento. [5]

Por su parte, la tercera componente principal presenta correlación con las variables benceno, octanaje y oxigenados. Esto, físicamente significa que tanto el benceno como los compuestos oxigenados incrementan el índice de octanos de la gasolina.[13] Un fenómeno interesante que se encuentra es la pobre de correlación (correlación nula) entre la variable oxigenados y las cinco temperaturas de destilación, principalmente porque entre el 2006 y 2010 se usó como compuesto oxigenado el etanol, el cual también es conocido por incidir directamente en las temperaturas de destilación de la gasolina mezclada.[5] Esta correlación nula se puede deber a que el porcentaje de etanol fue poco, inferior al 5 %, y que la mezcla se utilizó solo en ciertas regiones del país y por periodos cortos[4] por lo que su efecto no logró generar una correlación en las variables estudiadas.

Según el comportamiento de las variables y su ubicación en los planos principales descrito anteriormente, es posible etiquetar a los siguientes tres componentes principales de esta forma: la primera componente es “la volatilidad del combustible”, la segunda componente es “la facilidad el arranque en frío” y la tercera componente se puede etiquetar como “el octanaje del combustible”.

En la Figura 3, se representa la proyección de la nube de datos de las estaciones evaluadas durante el periodo 2006-2010 en el espacio generado por las tres primeras componentes principales. En esta figura, se encuentra que la primera componente segrega a las estaciones en dos grupos, separándolos por sus diferencias de volatilidad, mientras que a simple vista no es posible identificar si existe alguna segregación de las estaciones de servicio, de acuerdo a la facilidad de arranque en frío (segunda componente) y el octanaje del combustible (tercera componente).

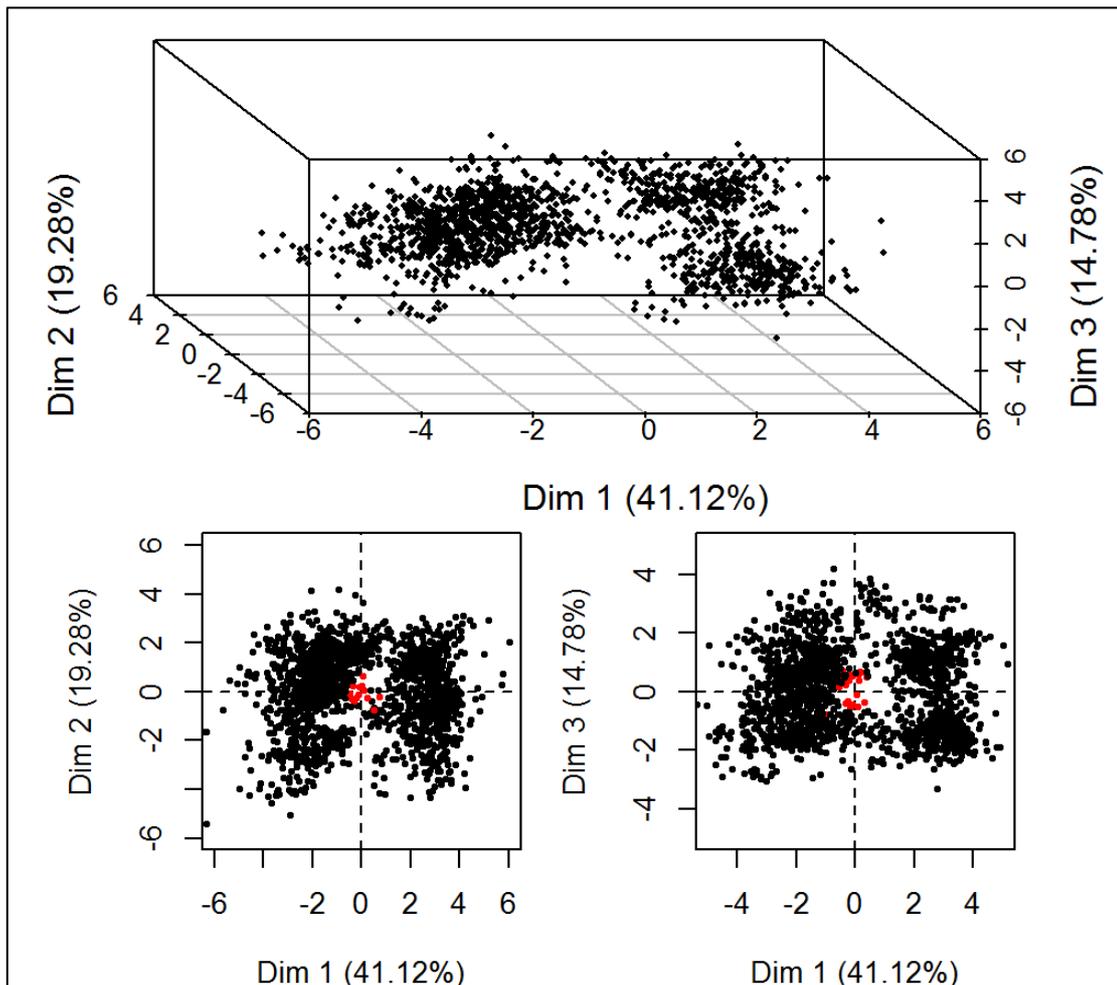
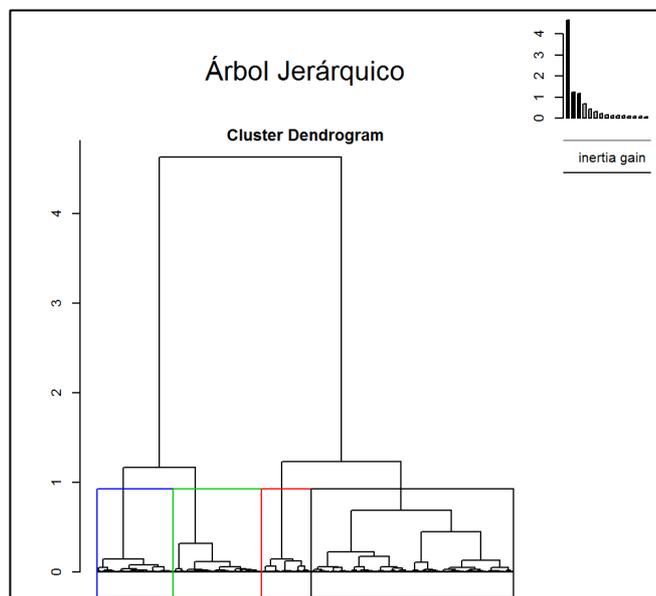


FIGURA 3. Círculo de correlaciones para las dos primeras componentes

Debido al análisis descrito anteriormente y aunado a que el objetivo del estudio es poder describir cómo se comportan las estaciones de servicio y cómo se agrupan, se aplica una clasificación automática de los grupos mediante una clasificación jerárquica ascendente.

La clasificación automática se hace 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. Este árbol jerárquico (dendrograma), generado a partir de las estaciones de servicio, reúne y enlaza por medio de nodos a los individuos más cercanos, creando grupos similares en los niveles inferiores del gráfico. Conforme aumentan las disimilitudes entre los individuos, los nodos que enlazan los

distintos grupos se elevan, hasta que todos quedan enlazados en un último nodo. De esta manera se forma un árbol de fácil interpretación, donde se enlazan y se agrupa a todos los individuos con respecto a su similitud y se separan los grupos respecto a sus disimilitudes, ya que entre más alto queden los nodos que los enlaza, más disimiles son los grupos. En el caso de este árbol, es claro que el "corte" se puede realizar en cuatro conglomerados.



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

Con esta clasificación en cuatro clústeres, se puede representar a los grupos sobre los tres planos principales, tal como se observa en la Figura 5. Lo anterior, junto a la descripción realizada sobre las 14 variables usadas en los círculos de correlación, permite identificar el comportamiento de estos clústeres. Por ejemplo, tal como se describió previamente, la volatilidad del combustible logra separar a las estaciones en dos regiones, pero ahora es evidente que tanto la facilidad de arranque en frío como el octanaje del combustible separan a estas dos regiones en cuatro grupos.

Según la información contenida en la Figura 5, se puede describir lo siguiente:

- El clúster 1 se encuentra ubicado entre el eje negativo de la primera componente y el eje positivo de la segunda, por lo que al compararlo al círculos de correlación correspondiente se encuentra que este grupo de estaciones de servicio presenta niveles más altos que el promedio en la cantidad de olefinas y benceno, y altas temperaturas de destilación inicial y al 10%, mientras que las temperaturas de destilación finales, presión de vapor Reid, oxigenados y octanaje, son menores que el promedio.
- El clúster 2 se ubica en el sector de valores mucho mayores que el promedio en la cantidad de oxigenados y valores altos en la presión de vapor Reid, pero al mismo tiempo, tiene valores más bajos que el promedio en todas las temperaturas de destilación, en la cantidad de aromáticos y en el índice de manejabilidad ordinario (ODI).
- Tanto el clúster 3 como el 4 poseen valores más altos que el promedio las temperaturas al 50 %, 90 %, y final de destilación, el residuo de la destilación, la densidad, el octanaje, los aromáticos y el ODI; el clúster 4 es el grupo que presenta los valores más altos de los clústeres, en la temperatura media y final de la destilación, la densidad, octanaje, aromáticos y el índice de Manejabilidad. Esto también se puede apreciar en el comportamiento de los centros de

gravidad de cada uno de los conglomerados encontrados (valor promedio en cada uno de los clúster). Ver Tabla 2 y Figura 6.

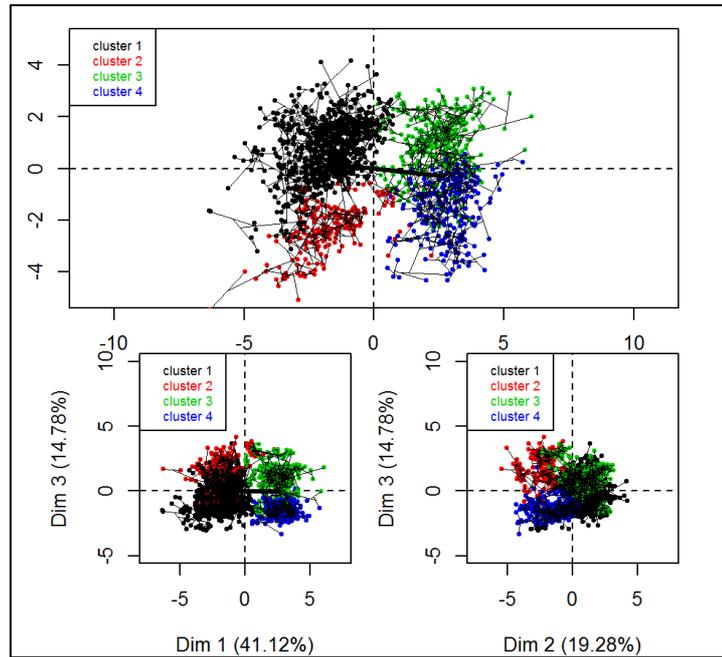


FIGURA 5. Planos principales 1-2, 1-3 y 2-3 con los cuatro aglomerados encontrados

TABLA 2. Centros de gravedad para las clases encontradas por la clasificación jerárquica ascendente

Clúster	P.Inicial	Dest.10.	Dest.50.	Dest.90.	Dest.final	Res.Dest	P.Vap.Reid
clust1	36.15	57.12	101.0	157.1	195.3	0.798	55.77
clust2	36.64	53.47	98.6	156.9	196.5	0.815	60.72
clust3	35.08	57.24	106.3	166.3	204.4	0.945	56.11
clust4	33.19	53.08	106.2	169.5	208.5	0.886	60.55
Clúster	Densidad	Octanaje	Oxigenados	Benceno	Aromáticos	Olefinas	ODI
clust1	739.9	90.18	0.528	1.256	24.18	19.20	546
clust2	743.2	92.15	2.076	1.218	23.76	17.66	533
clust3	748.7	92.36	0.601	1.268	28.77	14.15	571
clust4	742.7	91.43	0.316	0.896	27.62	12.36	568

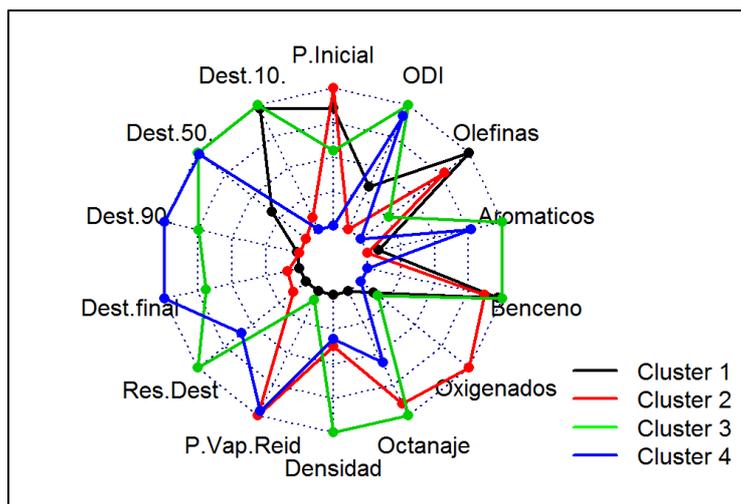


FIGURA 6. Gráfico tipo radar para los centros de gravedad encontrados

Con el comportamiento de los clústeres descrito anteriormente, se busca identificar cuales estaciones componen cada uno de estos grupos y determinar si estos individuos homogéneos poseen otra característica común. De igual forma que para el Diesel[14], se realiza una distribución en la frecuencia sobre la coincidencia de los años con los cuatro conglomerados encontrados en la clasificación jerárquica, como se muestra en la Tabla 3. Se vuelve a encontrar una evolución de la gasolina regular con respecto a cada uno de los años del estudio. En este caso el clúster 1 agrupa a los años 2006 al 2008, el clúster 3 agrupa el 2009 y el clúster 4 agrupa al 2010.

TABLA 3. Distribución en las frecuencias de los años del estudio en los cuatro clúster encontrados usando clasificación jerárquica

Año	Cluster_1	Cluster_2	Cluster_3	Cluster_4
2006	290	36	2	0
2007	250	65	19	1
2008	261	70	7	0
2009	2	23	305	3
2010	1	0	35	304

El clúster 2, línea roja en la figura 6, está compuesto por estaciones de servicio distribuidas en los primeros cuatro años del estudio; su principal característica es el alto contenido de oxigenados en el combustible. Al revisar la ubicación de las estaciones de servicio, se encuentra que el 84 % de estas provienen de las provincias de Guanacaste y Puntarenas, y además el 90.2 % de las estaciones pertenecientes al clúster 2 son abastecidas por el Plantel de RECOPE en Barranca, Puntarenas. Esto coincide con el periodo y la ubicación de Plan Piloto del MINAE y RECOPE para el mezclado de gasolina regular con etanol.[15] Hay que destacar que el etanol en la gasolina aumenta el número de octanos y esto se ve reflejado si hacemos la comparación del número de octanos del clúster 2 con el clúster 1. Esta comparación solo se puede hacer entre estos dos grupos ya que los clúster 3 y 4 reúnen años en los cuales el Plan Piloto se disminuyó a la mitad el contenido de alcohol con que se mezclaba la gasolina regular, finalizando el plan durante el 2011.

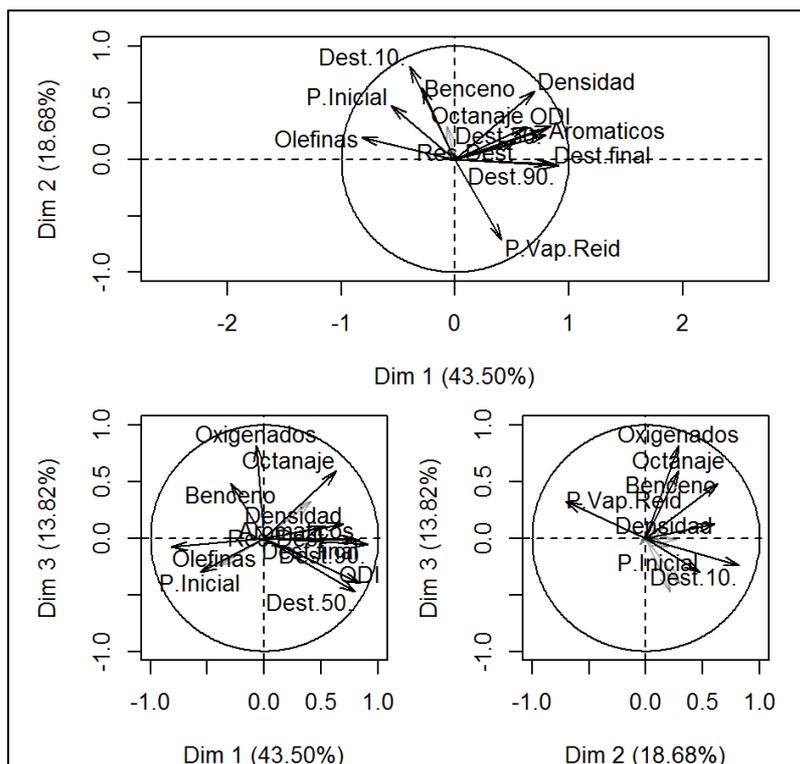
Al buscar otra característica distintiva, se observa que cada clúster agrupa a cada uno de los años del estudio, a excepción de los años 2006 y 2007 que se agrupan en un solo clúster, situación que se repite con respecto a lo visto en el caso del diésel.[14] En la Tabla 3 se obtiene la distribución

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.

Al observar los clúster 1, 3 y 4 se encuentra una evolución de la gasolina regular hacia un producto que mejor se ajusta a los parámetros de calidad conforme pasan los años, principalmente en la reducción de olefinas y benceno, la disminución de las temperaturas iniciales de la destilación y el aumento en el índice de Manejabilidad (Ordinal Driveability Index, ODI), en las temperaturas finales de la destilación y en el número de octanos. El único elemento que evolucionó en detrimento de la calidad del combustible fue el contenido de aromáticos en la gasolina regular, que presenta un aumento durante el periodo del estudio. Este último está relacionado con la promoción en el incremento de los depósitos en la cámara de combustión que aumenta las emisiones de gases en el escape y los niveles de benceno en estas emisiones.[13]

Debido al comportamiento atípico de las estaciones de servicio pertenecientes al clúster número 2 de la clasificación jerárquica, que coinciden con aquellas donde se realizó el Plan Piloto de mezclado de etanol con gasolina regular, se decide excluir los datos correspondientes del conjunto para realizar de nuevo un análisis sobre esta población y describir su comportamiento. Es necesario determinar cuántas componentes principales nuevas se deben utilizar para el análisis: en este caso las tres primeras componen el 76 % de la inercia total y la inercia individual de cada una es superior a la unidad. Por tanto, se utilizan nuevamente las tres primeras componentes principales para analizar los datos.

En la Figura 7 se presenta el círculo de correlaciones para las variable de la calidad de la gasolina regular, tal y como se presentó Figura 2. Al observar ambas figuras, se tiene que las variables presentan el mismo comportamiento, previamente descrito, donde la primera componente describe “la volatilidad del combustible”, la segunda componente representa “la facilitar el arranque en frío” y la tercera componente explica “el octanaje del combustible”. Existe una única variación en el comportamiento de una de las variables: contenido de benceno. Esta se ubica en entre los ejes positivos de la segunda y tercera componente principal, por lo que se entiende que existe una correlación positiva entre el contenido de benceno y la temperatura de destilación al 10 % del volumen recuperado y una correlación negativa entre el contenido de benceno y la presión de vapor Reid. Es decir: a mayor contenido de benceno, la temperatura de destilación al 10 % aumenta y la presión de vapor del combustible disminuye. Adicionalmente, el contenido de benceno se correlaciona (mediante la tercera componente) con el índice de octanaje del combustible. Por lo tanto, sí existe una relación entre el contenido de benceno y el octanaje de la gasolina. Esto es cierto, ya que el benceno es un componente natural del petróleo crudo y el producto del reformado catalítico, que produce flujos de gasolina de alto octanaje.[13]



**FIGURA 7.** Círculos de correlaciones para las componentes 1-2, 1-3 Y 2-3, usando el subconjunto de estaciones resultante al eliminar el clúster 2

Al observar la representación de los individuos sobre las primeras tres componentes principales en la Figura 8, se encuentra nuevamente que existe una fuerte influencia de la volatilidad del combustible en la segregación de las estaciones de servicio en el país y adicionalmente se encuentra que tanto la facilidad del arranque en frío como el octanaje del combustible dividen estos subgrupos, aunque su separación no es tan evidente como con la volatilidad del combustible.

Por lo tanto, se realiza nuevamente una clasificación jerárquica ascendente con el método de Ward y la distancia Euclídea clásica, al subconjunto de las estaciones de servicio resultantes después de eliminar el clúster 2. Con esta nueva clasificación se encuentra que en el nuevo árbol se puede realizar el "corte optimo" en cuatro conglomerados, coincidiendo con el mismo número de clústeres que el corte anterior, ver Figura 9

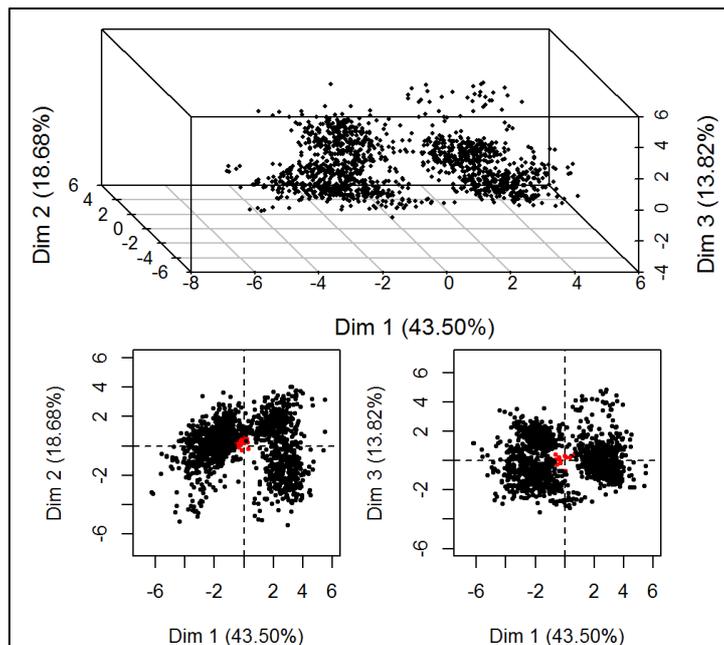


FIGURA 8. Representación de los individuos sobre las tres primeras componentes principales y planos principales 1-2 y 1-3, para el subconjunto de estaciones resultante al eliminar el clúster 2

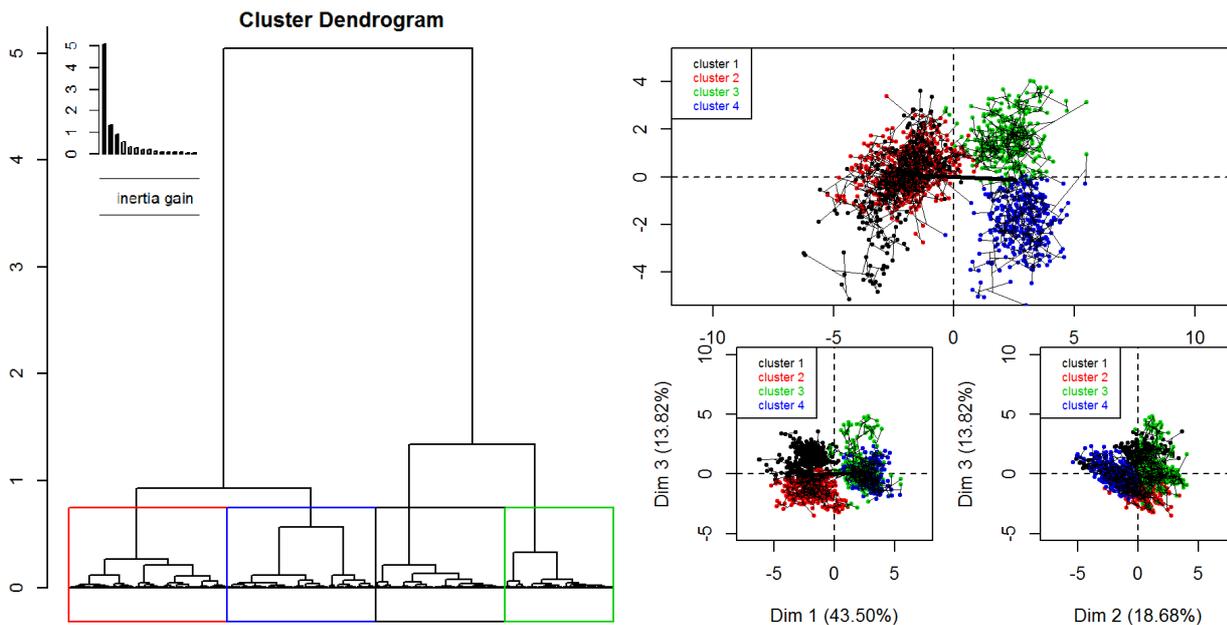


FIGURA 9. Dendrograma para la Clasificación Jerárquica y representación de los individuos sobre las tres primeras componentes principales y planos principales 1-2 y 1-3 para el nuevo subconjunto de datos

Al calcular los centros de gravedad de cada uno de los clúster formados, (ver Tabla 4) y complementando la información con lo indicado en la Figura 9 se puede describir el comportamiento de cada uno de los conglomerados encontrados, obteniéndose lo siguiente:

- El clúster 1 se encuentra ubicado sobre el eje negativo de la primera componente y el eje positivo de la tercera componente, por lo que al superponer el círculos de correlación correspondientes, se observa que este grupo de estaciones de servicio presenta más altos

niveles que el promedio en la cantidad de oxigenados, olefinas y benceno, pero posee puntos más bajos que el promedio en las temperaturas al 50 %, 90 %, y final de destilación, el residuo de la destilación, la densidad, los aromáticos y el ODI.

- El clúster 2 se ubica también sobre el eje negativo de la primera componente pero separado del clúster 1, que se ubica sobre el eje negativo de la tercera componente. Este sector presenta valores elevados en las dos primeras temperaturas de la destilación (inicial y al 10 %) y en la cantidad de olefinas, pero su principal característica es que presenta valores más bajos que el promedio en las dos últimas temperaturas de destilación (90 % y final), en el residuo de la destilación, en la presión de Vapor Reid, la densidad, el índice de octano y la cantidad de aromáticos y oxigenados.
- El clúster 3 es el grupo que presenta la mayoría de las variables por encima del promedio de todos los datos. A excepción del punto inicial de destilación, la presión de Vapor Reid y la cantidad de olefinas, todos los parámetros superan el promedio. Es interesante observar como este grupo posee el parámetro de olefinas más bajo de todos los clústeres encontrados.
- Por último el clúster 4 posee valores más altos que el promedio en las Temperaturas al 50 %, 90 %, y final de destilación, la presión de vapor Reid, los aromáticos y el ODI; pero posee puntos más bajos que el promedio en las dos primeras temperaturas de la destilación, en la cantidad de oxigenados, benceno y olefinas.

**TABLA 4.** Centros de gravedad para las clases encontradas por la clasificación jerárquica ascendente

Clúster	P.Inicial	Dest.10.	Dest.50.	Dest.90.	Dest.final	Res.Dest	P.Vap.Reid
clust1	35.41	56.61	98.7	156.5	194.6	0.809	56.83
clust2	36.8	57.59	103.3	157.8	196.3	0.795	54.78
clust3	35.02	57.24	106.2	166.5	204.4	0.946	56.2
clust4	33.21	53.12	106.2	169.5	208.5	0.885	60.5
Clúster	Densidad	Octanaje	Oxigenados	Benceno	Aromáticos	Olefinas	ODI
clust1	739.5	90.86	0.77	1.378	24.15	18.23	537
clust2	740.4	89.59	0.31	1.15	24.31	19.85	554
clust3	749.1	92.52	0.634	1.271	28.91	14.13	571
clust4	742.8	91.42	0.308	0.904	27.65	12.38	568

En resumen, se puede apreciar en la Figura 10 el comportamiento de los cuatro clúster con respecto a cada una de las variables de calidad medidas en la gasolina regular. El clúster 3 es el que posee la mayor cantidad de variables con valores altos a excepción de la presión de vapor, las olefinas y la temperatura inicial de destilación, mientras que el clúster 1 es el que posee las variables con los valores más bajos en los parámetros de calidad del combustible a excepción del benceno, olefinas y oxigenados.

Con esto, es posible concluir que las estaciones que conforman los clústeres 3 y 4 poseen un mejor combustible al compararlo con las estaciones del clústeres 1 y 2 pues hay un notorio incremento en el octanaje del combustible y una reducción en la cantidad de olefinas y benceno en los clústeres 3 y 4. Aunque hay una relativa mejor entre estos dos grupos de clústeres, existe a la vez un incremento en las temperaturas medias y finales de la curva de destilación que genera un aumento en el índice de manejabilidad ordinario (ODI), dando lugar a un rendimiento más pobre en la capacidad de conducción. Habitualmente, a valores más bajos de ODI se obtiene un mejor arranque en frío y un mejor rendimiento en caliente en el motor, pero una vez que se logra alcanzar un valor adecuado de ODI para las condiciones climáticas de la región, no hay ningún

beneficio para bajar aún más el ODI. Es por esto que este índice queda abierto a investigación con el fin de determinar el valor óptimo a utilizar en Costa Rica.

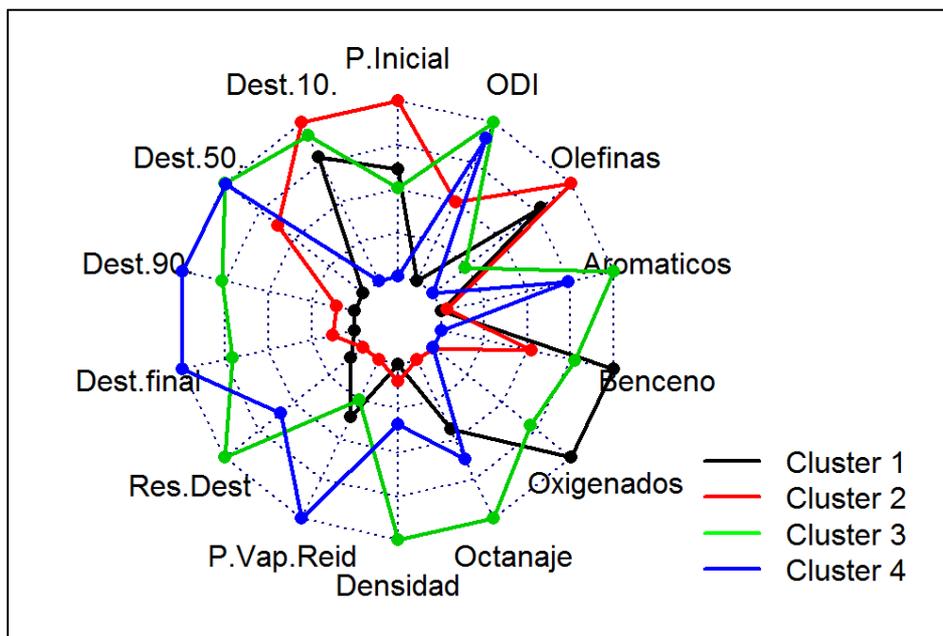


FIGURA 10. Gráfico tipo radar para los centros de gravedad encontrados

Con la premisa de que existe una correspondencia entre los clústeres encontrados con la Clasificación Jerárquica y los años del estudio, se calcula nuevamente la distribución en las frecuencias de los años con respecto a los cuatro grupos encontrados, ver Tabla 5. Esta vez se obtiene una mejor distribución de los años con respecto a los cuatro clústeres encontrados, ya que cada clúster agrupa las estaciones de servicio evaluadas en un periodo determinado; por ejemplo los clústeres 1, 3 y 4 conglomeran las estaciones evaluadas durante los años 2008, 2009 y 2010 respectivamente, y el clúster 2 agrupa a las estaciones evaluadas durante el periodo 2006 al 2007. Este resultado es de especial importancia ya que esta asignación de clústeres con respecto a periodos de estudio puede ser utilizada para determinar una evolución de los combustibles durante el periodo de estudio 2006-2010.

TABLA 5. Distribución en las frecuencias de los años del estudio en los cinco clúster encontrados usando clasificación jerárquica

Año	Cluster_1	Cluster_2	Cluster_3	Cluster_4
2006	56	236	0	0
2007	64	200	5	1
2008	256	7	5	0
2009	4	0	304	2
2010	0	0	28	312

Con respecto al periodo de estudio, la calidad de la gasolina evoluciona hacia un mejor combustible, pero de forma oscilatoria. Es decir, aunque globalmente la tendencia de los parámetros de calidad de la gasolina regular distribuida en Costa Rica es hacia una mejor calidad (al menos durante este periodo) esta mejora no es monótona en todas las variables, sino que tiende a ser oscilatoria; teniéndose por ejemplo que el periodo del 2009 mejores índices en ciertas

variables que los presentados en el años siguiente, ver Tabla 4. Para poder establecer mejor este comportamiento, es necesario ampliar la base de datos.

#### IV. CONCLUSIONES

- Para el periodo de estudio, los datos se clasifican en cuatro grupos que están ligados a una fecha en específico, lo cual sugiere que la calidad del combustible podría estar afectada por políticas anuales internas en Costa Rica o por una disponibilidad de producto terminado a nivel internacional (dependiendo de quien ofrezca el combustible). Esta situación es muy similar a la que se presenta durante el análisis realizado al Diesel, en una publicación anterior.
- El plan piloto de mezcla de etanol con gasolina regular tuvo un impacto visible en la calidad de los combustibles que se distribuyen el país, dado que generó su propio clúster en el primer análisis realizado. Esto causa un impacto directo en el número de oxigenados, la densidad, la presión de vapor y en el octanaje del combustible.
- Una vez eliminados los datos del clúster 2, generados a partir de la primera clasificación automática, de la tabla original; el efecto de la variación de la calidad de la gasolina regular por año es más evidente; esto indica que los datos se están comportando con valores atípicos con respecto al resto de la tabla.
- En general, se encuentra que la gasolina regular evolucionó hacia un mejor combustible, con respecto a la reglamentación vigente, pero esta evolución se hace de manera oscilatoria durante el periodo de estudio, donde la principal mejora es en la disminución en el contenido de olefinas en la gasolina.

#### V. REFERENCIAS

- [1] Estadística de ventas, 2016, Dirección de Ventas, Refinería Costarricense de Petróleo, Costa Rica, Descargada: 19 Julio 2016, <https://www.recope.go.cr/productos/ventas/>
- [2] Trejos, J.; Castillo, W.; González, J., Analisis Multivariado de datos, Métodos y aplicaciones, Editorial Universidad de Costa Rica, 2014, pp 37-89.
- [3] Jolliffe, I. T., Principal Component Analysis, ed. Springer-Verlag, New York, 2002, pp 10-28..
- [4] León, C.; Zuñiga, P., Estudio Estadístico 2006- 2010." Universidad de Costa Rica, Ciudad de la Investigación, p. 55, 2011.
- [5] ASTM D4814-16A, Standard Specification for Automotive Spark-Ignition Engine Fuel, ASTM International, West Conshohocken, PA, 2014.
- [6] Totten, G., Fuels and Lubricants Hand Book: technology, Properties, Performance and Testing. ed. ASTM International: West Conshohocken, 2003, pp 61-85.
- [7] R Core Team, R: A Language and Environment for Statistical Computing. Vienna, Austria, 2016.
- [8] Husson, F.; Josse, J.; Le, S.; Mazet; J., FactoMineR: Multivariate Exploratory Data Analysis and Data Mining. 2016.
- [9] Adler, D.; Murdoch, D., y otros, rgl: 3D Visualization Using OpenGL. 2016.
- [10] Nakazawa, M., fmsb: Functions for Medical Statistics Book with some Demographic Data. 2015.
- [11] Wickham, H., ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

- [12] Williams, L. J.; Abdi, H., Principal Component Analysis, Wiley Interdiscip. Rev. Comput. Stat., 2010, 2 (4), pp. 433–459.
- [13] WWFC Committee, Worldwide Fuel Harmonisation, 2013, Worldwide Fuel Charter Committee, Descargada: 8 diciembre 2015.  
[https://www.acea.be/uploads/publications/Worldwide\\_Fuel\\_Charter\\_5ed\\_2013.pdf](https://www.acea.be/uploads/publications/Worldwide_Fuel_Charter_5ed_2013.pdf)
- [14] Bazán, V., Calidad del Diesel de Costa Rica entre los Años 2006-2010, Rev. Cienc. y Tecnol. , 2015, 31(2), pp. 37–51.
- [15] Secretaria Técnica, Informe Sobre la Auditoría de Carácter Especial Acerca del Avance en el Cumplimiento de las Metas Establecidas en el Plan Nacional de Desarrollo 2011-2014 Referidas a Agregar Etanol y Biodiesel a los Combustibles, 2014, Contraloría General de la República Costa Rica. Descargada: 11 Marzo 2015,  
[https://cgrfiles.cgr.go.cr/publico/docs.../SIGYD\\_D\\_2014022570.pdf](https://cgrfiles.cgr.go.cr/publico/docs.../SIGYD_D_2014022570.pdf)

## 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, al Dr. Carlos León por su apoyo y respaldo en la realización de la presente investigación y a la Tecnóloga Paula Valverde por valiosas correcciones en la escritura del presente artículo.