

EVALUACIÓN DE ELECTROFACIES EN UN YACIMIENTO DE LA FRANJA PETROLERA NORTE CUBANA

ELECTROFACIES EVALUATION IN A DEPOSIT OF CUBAN OIL NORTH BELT

Rosa María VALCARCE, Willy Roberto RODRÍGUEZ MIRANDA

Departamento de Geociencias de la Universidad Tecnológica de La Habana "José Antonio Echeverría", CUJAE. 3H3M+XJ6, La Habana, Cuba. E-mails: rvalcarce1959@gmail.com; wrodriguez1955@gmail.com

Introducción
Características geológicas del área de estudio
Materiales y métodos
 Aspectos teóricos de los métodos estadísticos – matemáticos aplicados
 Análisis de componentes principales
 Análisis clúster no jerárquico K medias
 Registros geofísicos utilizados
 Procedimiento empleado
Resultados y Discusión
Conclusiones
Referencias

RESUMO - Este artigo mostra os resultados obtidos após a aplicação de técnicas de classificação estatística multivariada em registros geofísicos de radioatividade gama natural, coeficiente de absorção fotoelétrica, porosidade total, densidade, resistividade elétrica e calibre em poços localizados no campo Habana del Este, a oeste do Faixa Petrolífera do Norte de Cuba, com o objetivo de identificar eletrofácies correspondentes a selos e reservatórios. O procedimento aplicado integrou análise estatística descritiva, análise de componentes principais e a técnica de classificação estatística multivariada não supervisionada K-média, conseguindo distinguir quatro eletrofácies principais associadas a selos e reservatórios. As eletrofácies que constituem os selos estão relacionadas com a Formação Vega Alta e são caracterizadas por sua alta argilosidade. As eletrofácies que representam os coletores estão associadas ao Grupo Veloz e são constituídas por sequências com predominância de carbonatos ou arenitos, pouco argilosos, fraturados e porosos. Destaca-se uma eletrofácies com presença de matéria orgânica, onde predominam os carbonatos e que constitui o reservatório principal. O desenvolvimento e a aplicação desses procedimentos contribuem para reduzir os riscos da prospecção de petróleo.

Palavras-chave: Eletrofácies. Registros geofísicos de poços. Análise estatística multivariada. Faixa Petrolífera Norte de Cuba.

RESUMEN - Este artículo muestra los resultados obtenidos después de aplicar técnicas de clasificación estadística multivariada sobre registros geofísicos de radiactividad gamma natural, coeficiente de absorción fotoeléctrica, porosidad total, densidad, resistividad eléctrica y calibre en pozos ubicados en el yacimiento Habana del Este, al oeste de la Franja Petrolera Norte Cubana, con el objetivo de identificar electrofacies correspondientes a sellos y reservorios. El procedimiento aplicado integró el análisis estadístico descriptivo, el análisis de componentes principales y la técnica de clasificación estadística multivariada no supervisada K-medias, logrando distinguir cuatro electrofacies principales asociadas a sellos y reservorios. Las electrofacies que constituyen sellos se relacionan con la Formación Vega Alta y se caracterizan por su alta arcillosidad. Las electrofacies que representan los colectores se asocian al Grupo Veloz y están constituídas por secuencias con predominio de carbonatos o de areniscas, poco arcillosas, fracturadas y porosas. Se destaca una electrofacies con presencia de materia orgánica, donde predominan los carbonatos y que constituye el reservorio principal. El desarrollo y aplicación de estos procedimientos contribuye a reducir los riesgos de la prospección petrolera.

Palabras claves: Electrofacies. Registros geofísicos de pozo. Análisis estadístico multivariado. Franja Petrolera Norte Cubana.

ABSTRACT - This paper shows the results obtained after applied multivariate statistical classification techniques on geophysical logs in wells located in Cuban Oil North Belt, in Habana del Este field. The well logs geophysics employed were: natural gamma radioactivity, photoelectric absorption coefficient, total porosity, density, electrical resistivity and caliper. The aim was identifying electrofacies corresponding to seals and reservoirs. The applied procedure integrated the statistical descriptive techniques, main components and K-means algorithm like unsupervised multivariate statistical classification. It was possible to distinguish four main electrofacies associated with seals and reservoirs. The electrofacies that constitute the seals are related to Vega Alta Formation and are characterized by their high clay content. The electrofacies that represent the collectors are associated with Veloz Group and are made up of sequences with a predominance of carbonates or sandstones, fractured and porous, little clayey. An electrofacies stands out with organic matter presence, where carbonates predominate and which constitutes the main reservoir. The application of these procedures contributes to reduce oil prospecting risks.

Keywords: Electrofacies. Well logs geophysics. Multivariate statistical analysis. Cuban Oil North Belt.

INTRODUCCIÓN

La Franja Petrolera Norte Cubana (FPNC) se localiza en la faja costera de las provincias de La Habana, Artemisa, Mayabeque y Matanzas, incluyendo sus aguas marinas adyacentes (Figura 1). Su extensión es de 200 km de largo y 6 km de

ancho aproximadamente, y está compuesta por el cinturón de yacimientos y prospectos que se encuentran a lo largo de la faja costera, tanto en tierra como en el mar. En la actualidad representa el mayor potencial gasopetrolífero del país,

donde los reservorios carbonatados, caracterizados por su gran espesor y alta complejidad, constituyen los principales yacimientos de petróleo y gas en Cuba (Álvarez et al., 2002; Otero et al., 2009).

Identificar las facies geológicas es una tarea esencial en la prospección y caracterización de los yacimientos gasopetrolíferos, porque permite caracterizar la litología de las formaciones geológicas e identificar las rocas que pueden constituir sellos y reservorios. Cuando las facies geológicas son definidas a partir del comportamiento de los registros geofísicos de pozo se denominan electrofacies. El término electrofacies fue introducido por Serra y Abbot (1982), quienes lo definieron como “conjunto de respuestas de los registros geofísicos que caracterizan un estrato y permiten distinguirlo de otro”. Es importante destacar que una electrofacies puede estar asociada a una o varias litologías, porque la respuesta de los registros geofísicos depende de las propiedades físicas de las rocas y, en no pocas ocasiones, diferentes litologías presentan similitud en sus propiedades físicas. Convencionalmente la determinación de las electrofacies se realiza empleando gráficos cruzados y su correlación con datos de núcleos, sin embargo, con el desarrollo de la tecnología han surgido metodologías basadas en técnicas de minería de datos e inteligencia artificial para dar solución eficiente a esta importante tarea.

Wolff y Pelissier (1982) señalaron que los registros geofísicos contienen mucha de la información necesaria para caracterizar las facies geológicas, los que, además, tienen la ventaja de contener esta información de forma continua a lo largo del pozo.

Desarrollaron el algoritmo FACIOLOG para dividir las electrofacies presentes en el pozo y definir su correspondencia con las facies geológicas de pozos o intervalos patrones empleando técnicas estadísticas multivariadas. Sus investigaciones se han continuado desarrollando y

aplicando hasta la actualidad. Lim et al. (1997) empleó el análisis de varianza, componentes principales y análisis de agrupaciones jerárquicas para definir las electrofacies en el estudio de un yacimiento de petróleo y logró la mejor definición y caracterización de los reservorios. Lee et al. (2002) identificaron las electrofacies integrando el análisis de componentes principales con modelos de clasificación basados en análisis clúster y análisis discriminante. Kumar y Kishore (2006) definieron las electrofacies combinando una técnica de clasificación no supervisada (algoritmo K-media) y una técnica de clasificación supervisada (redes neuronales artificiales). También Davis (2018) y Puskarczyk (2020) lograron definir con alta precisión las electrofacies en yacimientos gasopetrolíferos aplicando componentes principales, análisis clúster, análisis discriminante y redes neuronales.

El algoritmo K-media, como análisis de agrupamiento no supervisado aplicado a los registros de densidad y velocidad de la onda P, permitió identificar las litologías presentes en un yacimiento gaso petrolífero (Alí y Sheng-Chang, 2020). A partir del centroide de los clústeres los autores identificaron las litologías presentes en los pozos e identificaron las zonas de mayor interés para la explotación de hidrocarburos.

En un pozo petrolífero de la cuenca Ceará en Brasil, fueron identificadas cinco electrofacies aplicando análisis de correspondencias y análisis discriminante a los registros geofísicos radiactivos, sísmico y densidad (Normando et al., 2021). Los resultados permitieron identificar las electrofacies que constituyen los reservorios de mayor y menor calidad en función de su arcillosidad y porosidad.

El objetivo de la presente investigación es identificar las electrofacies presentes en un yacimiento de petróleo y gas en la zona oeste de la Franja Petrolera Norte Cubana para la mejor caracterización de los sellos y reservorios, contribuyendo a reducir el riesgo de la prospección petrolera.

CARACTERÍSTICAS GEOLÓGICAS DEL ÁREA DE ESTUDIO

El área de estudio se ubica en el sector oeste de la FPNC, al norte de la provincia La Habana (Figura 1). En esta región se han realizado estudios estratigráficos desde la segunda mitad del pasado siglo, como lo demuestran las investigaciones de Brönnimann y Rigassi (1963) y Albear e Iturralde (1985). Desde el punto de vista geológico el área presenta la misma

constitución que el resto de los yacimientos de la FPNC.

En las investigaciones de Morales et al. (2013) se destaca a la formación Vega Alta como el sello regional de la FPNC al constituir la cobertura de la Unidad Tectónico Estructural Placetas donde se localizan los principales reservorios en Cuba, y se describen dos grandes miembros en esta formación,

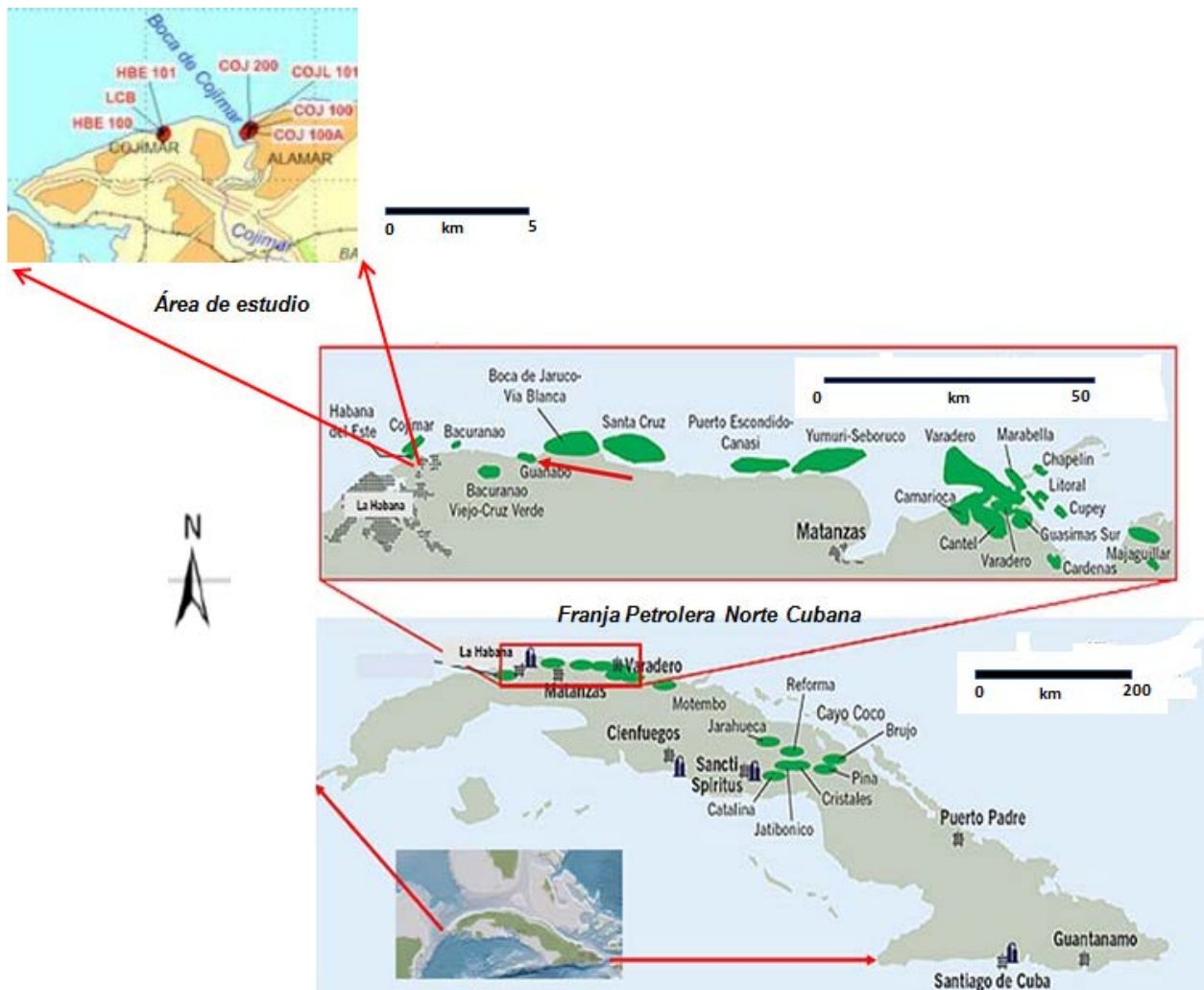


Figura 1 – Ubicación geográfica de la Franja Petrolera Norte Cubana y del área de estudio. Modificado de Gómez (2021)

uno de características olistostrómicas y otro caracterizado por intercalaciones de sedimentos silíceos arcillosos carbonatados.

Brey del Rey (2013) realizó la división estratigráfica del área de estudio identificando la presencia de la formación Vega Alta e intervalos colectores que definió como reservorio Canasí. Señaló que en la formación Vega Alta, de edad Paleoceno-Eoceno inferior, predomina una secuencia caótica policomponente de matriz arcillosa y arcilloso arenosa, que incluye bloques de diferentes dimensiones hasta olistolitos de calizas, areniscas, brechas, silicitas, serpentinitas, rocas volcánicas y metamórficas. Se intercalan paquetes de rocas siliciclásticas, como conglomerados, areniscas, limolitas y lutitas. Se encuentran además rocas del complejo ofiolítico.

Describió la presencia de reservorio en todos los pozos que analizó, algunos de los cuales cortaron dos mantos de reservorio separados por un sello intramanto. Estos reservorios se caracterizan litológicamente por una facies carbonatada subordinadamente silícea – arcillosa que presentan hasta un 85% de carbonatos o incluso

valores mucho más altos. También se describen pedernales entre un 10 a un 15% y caliza compacta del tipo *claystone* hasta un 5%. En pequeños intervalos se describe un alto contenido de marga.

Las investigaciones de Almaguer et al. (2016) señalan que las rocas del cinturón ofiolítico se evidencian en todos los sondeos realizados en el área de estudio, entre 680 y 1 000 m de profundidad, fundamentalmente con presencia de serpentinitas muy fracturadas, en ocasiones alteradas a minerales arcillosos y con elementos de otras litologías subordinados. Los mismos autores reconocen por debajo de las serpentinitas una secuencia principalmente conglomerática probablemente hasta olistostrómica con varios cuerpos menores u olistolitos de serpentinitas, típica de la formación Vega Alta. También identifican el reservorio principal del Grupo Veloz como en el resto de la FPNC, el mismo reservorio que Brey del Rey (2013) identificara como reservorio Canasí, teniendo en cuenta que la formación Canasí es una unidad informal que no está reconocida en el léxico estratigráfico de

Cuba (IGP, 2016).

Reconocen que este reservorio está constituido por una secuencia carbonatada con predominio de *mudstones* calcáreos de edad Jurásico superior Tithoniano superior – Cretácico inferior,

con intercalaciones arcillosas, que se clasifica como colector carbonatado fracturado – poroso.

En la tabla 1 se presentan los rangos de profundidad de las estructuras identificadas en el área de estudio para los pozos analizados.

Tabla 1- Rangos de profundidad del sello y reservorio en el área de estudio.

Estructura geológica	Rangos de profundidad
Formación Vega Alta (sello regional)	770 m a 3230 m
Reservorio Principal	1798 m a 4325 m
Sello Interno Vega Alta	1813 m a 3310 m
Reservorio	1863 m a 2780 m

Fuente: Modificado de Brey del Rey (2013)

MATERIALES Y MÉTODOS

Para dar cumplimiento al objetivo planteado en esta investigación, sobre los registros geofísicos de pozo disponibles se aplicó el análisis de componentes principales y, como método de clasificación no supervisada, el análisis clúster no jerárquico K medias.

Aspectos Teóricos de los Métodos Estadísticos-Matemáticos Aplicados *Análisis de componentes principales*

Es muy común que en el conjunto de atributos disponibles algunos estén correlacionados en mayor o menor grado y por tanto, sean redundantes, es decir, sean portadores de la misma información. Por ello se han desarrollado técnicas estadístico matemáticas como el análisis de componentes principales (ACP) para describir el objeto de estudio con un nuevo sistema de atributos no correlacionados. Este nuevo sistema de atributos es calculado a partir de las variables originales de tal manera que, si existen atributos redundantes entre los p atributos originales, el número de variables que integren el nuevo sistema de atributos deberá ser menor que p .

Las componentes principales (CP) son las variables que constituyen el nuevo sistema de atributos. Se calculan a lo largo de las direcciones principales definidas por los vectores propios asociados a los valores propios de la matriz de correlación lineal de los atributos originales. La primera CP presentará la mayor varianza y la última será la que contribuye con la menor variabilidad. Esto significa que la primera CP contiene la mayor información útil portada por los atributos originales, lo que equivale a decir que en la primera CP (o en las primeras CP) se sintetiza gran parte de la información que aporta el sistema de atributos bajo estudio.

Las CP se obtienen a partir de una relación

lineal homogénea de los atributos originales según la ecuación siguiente (Alfonso, 1989):

$$CP_i = \sum_{j=1}^p a_{ij}x_j \quad i = 1, 2, 3, \dots, q \quad (1)$$

Donde

CP_i: atributo nuevo i ésimo denominado componente principal i ésimo.

x_j : atributo original j ésimo normalizado.

a_{ij} : vectores propios unitarios asociados a los valores propios de la matriz de correlación lineal de los atributos originales. Representan la carga factorial de las componentes principales, es decir, el peso o importancia de cada variable en el cálculo de cada CP. Permiten identificar qué tipo de información porta cada una de las CP.

Análisis clúster no jerárquico K medias

El análisis clúster no jerárquico K-medias es una técnica de reconocimiento de patrones no supervisada, que permite obtener grupos a partir de gran cantidad de datos, de tal manera que los elementos de cada grupo sean muy similares entre sí y, a la vez, sean muy diferentes a los elementos de los otros grupos. Este algoritmo permite asignar cada objeto al grupo más cercano, lo que se logra calculando una medida de similitud entre el objeto y el centroide de cada clúster. La selección adecuada de la medida de similitud juega un rol importante porque define que los resultados finales tengan la mayor confiabilidad posible. En esta investigación se empleó como medida de similitud el coeficiente de distancia euclidiano, recomendado cuando las variables empleadas se expresan de forma cuantitativa y existe baja correlación estadística entre ellas (Hamdan y Emad, 2017).

El número de grupos y sus centroides se calculan inicialmente de forma aleatoria, y

después de la primera asignación de objetos a cada grupo, se recalculan los centroides como la media de las variables de los puntos que le fueron asignados. Una vez actualizado el centroide de cada clúster se vuelven a reasignar los objetos al grupo más cercano. Este procedimiento se repite hasta lograr la convergencia, o sea, hasta que las asignaciones de los puntos no cambien, o hasta alcanzar el número de iteraciones prefijado. Este resultado final representa el ajuste que maximiza la distancia entre los distintos grupos y minimiza la distancia intragrupo. La principal ventaja del método es su sencillez y rapidez, pero es un algoritmo significativamente sensible a los centroides que se seleccionan inicialmente de manera aleatoria. Este efecto se puede reducir incrementando el número de iteraciones del procedimiento

(Hernández–Orallo et al., 2004; Imron et al., 2020).

El algoritmo es más eficiente en la medida que las variables empleadas no sean redundantes, y es muy sensible al hecho de que las variables posean diferente rango de variación, por lo que se recomienda estandarizarlas entre 0 y 1 sustrayendo a cada variable su valor mínimo y dividiendo por su rango. El algoritmo siempre separará los objetos en grupos y solamente el conocimiento del investigador hará posible identificar qué grupos son significativos y cuáles no.

Registros Geofísicos Utilizados

Para identificar las electrofacies fueron utilizados los registros geofísicos medidos en tres pozos, en el intervalo de profundidad entre 2160 y 3215 m. La tabla 2 presenta los registros geofísicos disponibles y sus unidades de medida.

Tabla 2 - Registros geofísicos de pozos disponibles.

Método	Registro geofísico	Variable medida	Unidad de medida
Control técnico del pozo	Caliper (HCAL)	Diámetro del pozo	cm
	Gamma Natural Integral (SGR)	Radiactividad natural integral (U + Th + K)	API
Radioactivos	Gamma Natural Computado (CGR)	Radiactividad natural integral (Th + K)	API
	Contenido de potasio (POTA)	Contenido de potasio	V/V, %
	Contenido de torio (THOR)	Contenido de torio	PPM
	Contenido de uranio (URAN)	Contenido de uranio	PPM
	Efecto Fotoeléctrico (PEF)	Efecto Fotoeléctrico	B/E (bernios/electrón)
	Densidad (RHOB)	Densidad	σ/cm^3
	Porosidad Neutrónica (NPHI)	Porosidad Neutrónica	V/V, %
Eléctricos	Resistividad de corriente enfocada profunda (LLD)	Resistividad eléctrica	Ohm-m
	Resistividad de corriente enfocada somera (LLS)	Resistividad eléctrica	Ohm-m

Estos registros contienen información útil sobre la litología de las formaciones geológicas atravesadas por el pozo y sobre determinados parámetros de los reservorios, por lo que resultan efectivos para identificar las electrofacies.

El registro HCAL mide las variaciones del diámetro del pozo. Las rocas arcillosas, fracturadas, cavernosas, generalmente presentan diámetros del pozo aumentados, a diferencia de las rocas duras y compactas. Por tanto, este registro contribuye a inferir algunas de las características de las formaciones geológicas.

El registro SGR aporta información sobre la radiactividad natural de las rocas y constituye un indicador litológico por excelencia. Los elementos radiactivos como potasio y torio tienden a concentrarse en los minerales arcillosos. Las rocas ultrabásicas se caracterizan por emitir valores

muy bajos de radiactividad natural. También los sedimentos arenosos presentan baja radiactividad. El registro CGR es un registro de radiactividad gamma natural al que se ha extraído la influencia del uranio, solo contiene la radiactividad gamma natural aportada por el torio y el potasio presentes en las rocas. Los carbonatos de origen químico, sin materia orgánica, precipitados en un ambiente oxidante, son rocas no radiactivas. Sin embargo, cuando los sedimentos carbonatados se depositan en un ambiente reductor, generalmente favorable a la conservación de materia orgánica y a su transformación en hidrocarburos, ello va acompañado por un contenido variable de uranio. Es por ello que el registro CGR es muy útil en la separación de electrofacies frente a formaciones carbonatadas ricas en materia orgánica como las que están presentes en el área de estudio, al

presentar valores muy inferiores de radiactividad gamma natural con respecto al SGR. En rocas muy arcillosas son elevados los valores de torio y potasio, y los registros de CGR y SGR presentan valores muy similares (Castro, 2017).

El registro fotoeléctrico es también un indicador litológico al ser un registro que responde a la composición mineralógica de las rocas. Las areniscas presentan bajos valores del coeficiente de absorción fotoeléctrica (PEF), mientras que las calizas y dolomías presentan mayores valores de esta propiedad física. Las rocas que contienen hierro o minerales pesados presentan PEF aún mayor. Este registro es afectado por la presencia de elementos pesados en el lodo de perforación, como la barita, resultando un método efectivo para determinar la presencia de rocas fracturadas en las cuales ha ocurrido la invasión del filtrado del lodo y se manifiestan valores anómalamente altos de PEF.

El registro NPHI responde a la cantidad de hidrógeno en las rocas, por lo que se asocia a la presencia de fluidos en sus poros. En rocas que no tienen hidrógeno en su red cristalina este registro es un indicador de la porosidad total. El registro RHOB mide la densidad de las rocas y es usado principalmente para evaluar su porosidad y, en combinación con otros registros, también se emplea para identificar formaciones arcillosas y de litología compleja. (Darwin y Singer, 2008).

Los métodos eléctricos miden la resistividad eléctrica de las rocas atravesadas por el pozo. El registro LLS y el registro LLD miden la resistividad eléctrica de la roca en la zona invadida por el filtrado del lodo y en la zona virgen respectivamente. Estos registros permiten evaluar la saturación de las rocas y apoyar en la identificación de las litologías.

Procedimiento Empleado

Antes de comenzar el procesamiento de la información fue realizado el control de calidad de las mediciones. Los registros geofísicos fueron

digitalizados cada 1,5 cm obteniéndose un total de 18 611 puntos. Las técnicas estadístico matemáticas empleadas para identificar las electrofacies fueron aplicadas sobre los registros HCAL, SGR, CGR, PEF, RHOB, NPHI, LLD, LLS; medidos en los tres pozos disponibles. Se realizó el análisis exploratorio de datos con la finalidad de estudiar el comportamiento de las variables de manera individual y las relaciones entre ellas.

Todas las variables se normalizaron entre 0 y 1 para evitar la influencia de sus valores absolutos en los resultados. Para ello fue aplicada la siguiente ecuación:

$$X_{normalizada} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

Donde:

X : valor de la variable a normalizar

X_{min} : valor mínimo de la variable X

$X_{máx}$: valor máximo de la variable X

Teniendo en cuenta la alta variabilidad de los registros LLD y LLS, fue normalizada la variable logaritmo de la resistividad eléctrica.

Fueron calculadas las componentes principales sobre el sistema de variables normalizadas ya señalado. Se analizó la importancia de cada CP para explicar la variabilidad total del sistema bajo estudio y fue analizada la influencia de cada variable en el cálculo de las CP. De esta manera fue posible comprender la información que porta cada CP.

El análisis de agrupaciones se realizó aplicando el algoritmo K medias sobre el nuevo sistema de variables constituido por las CP más informativas. Conociendo la información que aporta cada CP desde el punto de vista geofísico y analizando los centroides de cada clúster, fue posible definir las electrofacies que se asocian a cada grupo. La figura 2 resume el esquema del procedimiento empleado.

RESULTADOS Y DISCUSIÓN

La estimación puntual de los estadígrafos media, moda, desviación típica, valor mínimo y máximo para cada una de las magnitudes físicas se presenta en la tabla 3. Los histogramas para visualizar la distribución de frecuencia de las diferentes propiedades físicas y su variabilidad se presentan en la figura 3.

En la tabla 3 se puede apreciar la elevada variabilidad que presentan los registros de resis-

tividad eléctrica, radiactividad natural y porosidad de las rocas. En general, los histogramas muestran elevada asimetría, excepto el correspondiente a la variable densidad (RHOB) que presenta forma aproximadamente gaussiana, con moda 2,5 g/cm³.

Se destaca también el comportamiento bimodal que presenta la distribución de frecuencias de la radiactividad gamma natural integrada (SGR) y

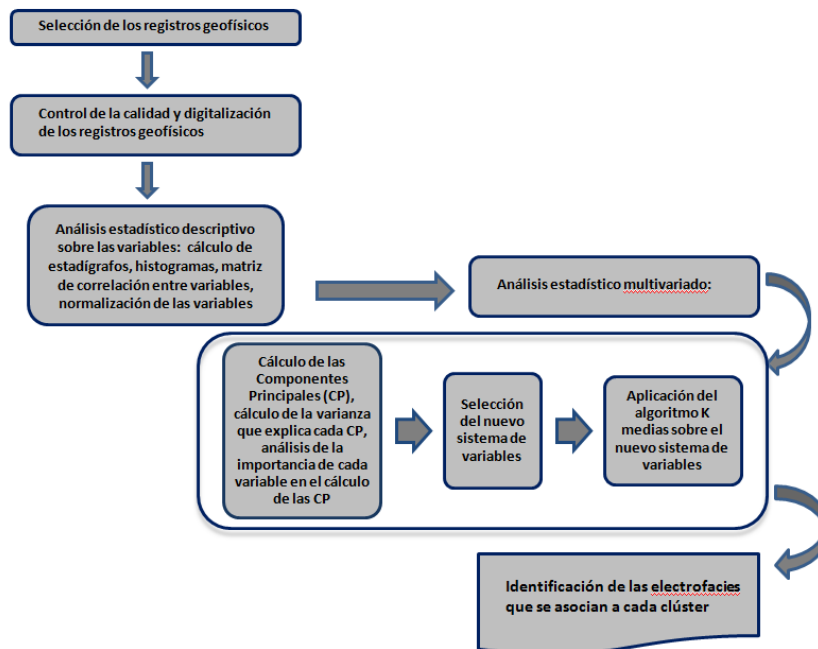


Figura 2 – Procedimiento de trabajo empleado para identificar las electrofacies.

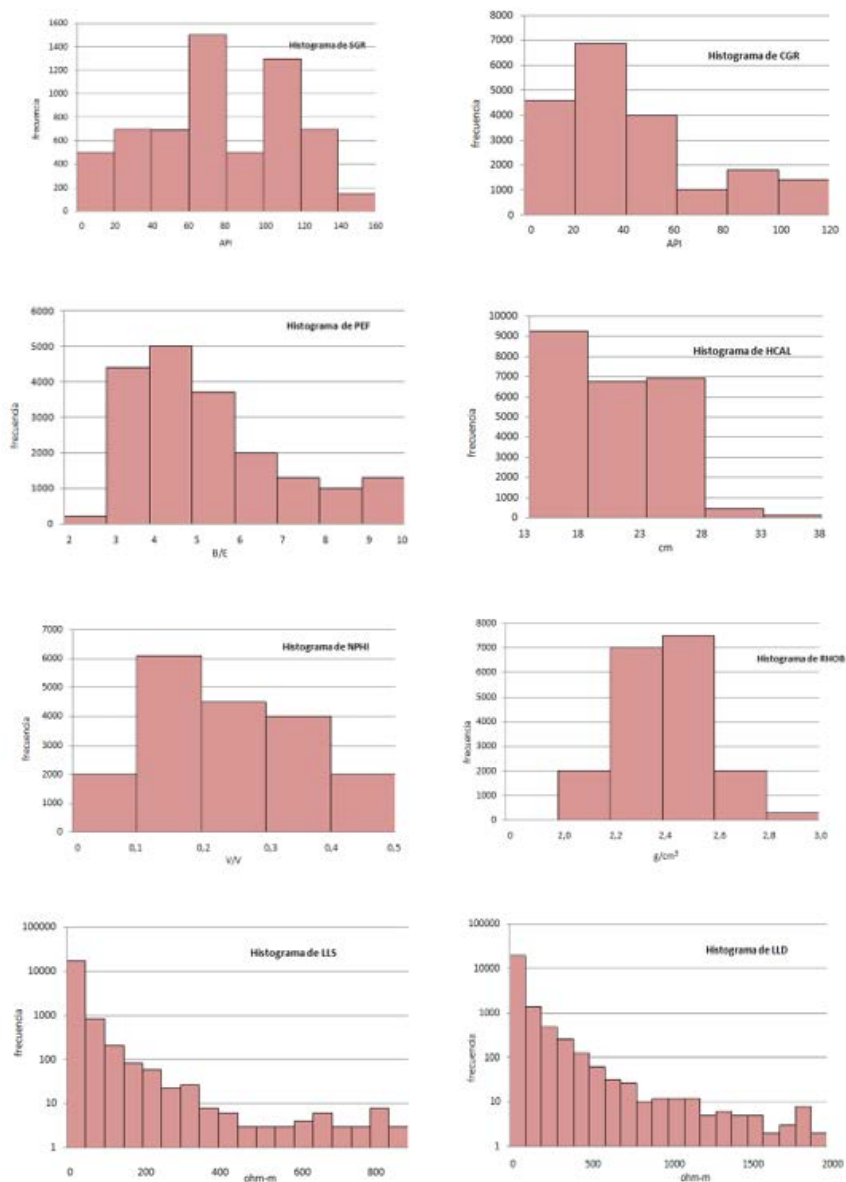


Figura 3 – Histogramas de las variables medidas en los tres pozos empleados en la investigación

Tabla 3 – Estimación puntual de los estadígrafos calculados a cada registro geofísico.

Variable	Unidad de medida	Media	Moda	Desviación típica	Valor máximo	Valor mínimo
HCAL	cm	21,0	15,5	4,5	38,0	13,0
SGR	API	49,2	70,0 y 110,0	28,6	151,0	2,0
CGR	API	39,7	30,0	24,2	135,0	1,0
PEF	b/e	5,4	4,5	1,74	10,0	2,4
RHO	g/cm ³	2,4	2,5	0,17	3,0	2,0
NPHI	v/v	0,24	0,15	0,12	0,46	0,0
LLD	Ohm-m	49,0	50,0	118,3	1982,0	0,2
LLS	Ohm-m	14,3	25,0	39,3	880,0	0,2

computada (CGR).

Las características anteriormente señaladas evidencian la mezcla de poblaciones estadísticas, o sea, la presencia de rocas que son producto de

procesos geológicos diversos y que, en su conjunto, conforman diferentes electrofacies. La tabla 4 presenta la matriz de correlación lineal entre las variables medidas

Tabla 4 – Matriz de correlación entre las variables (N=18 611)

Variable	CGR	SGR	PEF	LLD	LLS	NPHI	RHOB	HCAL
CGR	1,00							
SGR	0,85	1,00						
PEF	-0,27	-0,10	1,00					
LLD	-0,18	-0,15	0,19	1,00				
LLS	-0,36	-0,21	0,24	0,52	1,00			
NPHI	0,54	0,43	-0,40	-0,22	-0,30	1,00		
RHOB	-0,19	-0,23	0,41	0,24	0,10	-0,54	1,00	
HCAL	0,31	0,20	-0,25	-0,11	-0,07	0,29	-0,10	1,00

Se destaca la existencia de correlación lineal entre los registros geofísicos SGR y CGR; NPHI y CGR; LLS y LLD; NPHI y RHOB, para un nivel de significación del 5%.

La elevada correlación positiva entre los registros SGR y CGR refleja que fundamentalmente la radiactividad natural responde al contenido de Th y K en las rocas, lo que es expresión de la arcillosidad del corte geológico.

También el grado de correlación positiva que se observa entre la porosidad neutrónica y el registro gamma computado expresa la relación de estos registros con la arcillosidad de las formaciones geológicas presentes. Por último, se evidencia correlación negativa entre la poro-

sidad total y la densidad de las rocas, como es de esperar en este corte geológico donde predominan rocas sedimentarias arenó arcillosas.

El análisis de correlación estadística indica que existe redundancia en la información de las variables y que el análisis de componentes principales permitirá reducir el número de variables necesarias para explicar la variabilidad total del sistema bajo estudio.

La tabla 5 muestra el porcentaje de la variabilidad del sistema que explica cada componente principal. Se aprecia que las tres primeras componentes principales son capaces de explicar el 75,7% de la variabilidad total del sistema.

Tabla 5 – Porcentaje de variabilidad que explica cada componente principal.

Componente Principal	% de variabilidad que explica	% de variabilidad acumulada que explica
1	48,0	48,0
2	16,5	64,5
3	11,2	75,7
4	10,2	85,9
5	7,4	93,3
6	4,0	97,3
7	1,8	99,1
8	0,9	100,0

En la tabla 6 se refleja la carga factorial (a_{ij}) de cada una de las componentes principales, lo que permite identificar la importancia de cada variable en el cálculo de las componentes principales para poder comprender la información geólogo geofísica que porta el nuevo sistema de atributos calculado.

La CP 1 explica el 48% de la variabilidad total del sistema de variables bajo estudio y está determinada de forma directamente proporcional con la porosidad neutrónica y la radiactividad (lo que resulta aún más evidente con el registro CGR, o sea, cuando se elimina la influencia del uranio), y de forma inversamente proporcional con la resistividad eléctrica, el PEF y la densidad. Todo ello permite asegurar que los mayores

valores de la CP1 se asocian a incrementos de la arcillosidad de las rocas.

El análisis de la correlación estadística entre la primera componente principal y la arcillosidad de las rocas (Varc) obtenida a partir de la interpretación de los registros geofísicos de pozo disponibles (Castro, 2017) se presenta en la figura 4.

El coeficiente de correlación es de 0,71 para un total de 18.611 puntos, lo cual indica existencia de correlación estadística estable para un 99% de probabilidad. Este resultado demuestra la alta dependencia exponencial positiva entre ambas magnitudes y ratifica que la variabilidad de la CP1 responde a la arcillosidad del corte geológico.

Tabla 6 – Carga factorial asociada a las componentes principales.

Variable	a_{1j}	a_{2j}	a_{3j}	a_{4j}	a_{5j}	a_{6j}	a_{7j}	a_{8j}
CGR	0.76	-0.58	0.02	0.00	-0.17	0.01	-0.13	-0.17
SGR	0.62	-0.70	0.18	-0.16	-0.06	-0.13	0.17	0.12
PEF	-0.59	-0.47	0.07	0.15	0.63	-0.05	-0.05	-0.02
LLD	-0.79	-0.34	0.11	-0.34	-0.17	0.22	-0.21	0.09
LLS	-0.83	-0.07	0.02	-0.50	-0.02	0.05	0.21	-0.12
NPHI	0.83	0.10	0.14	-0.06	0.24	0.46	0.07	0.01
RHOB	-0.55	-0.38	-0.45	0.50	-0.20	0.20	0.11	0.01
CALI	0.45	-0.03	-0.79	-0.38	0.17	-0.03	-0.04	0.03

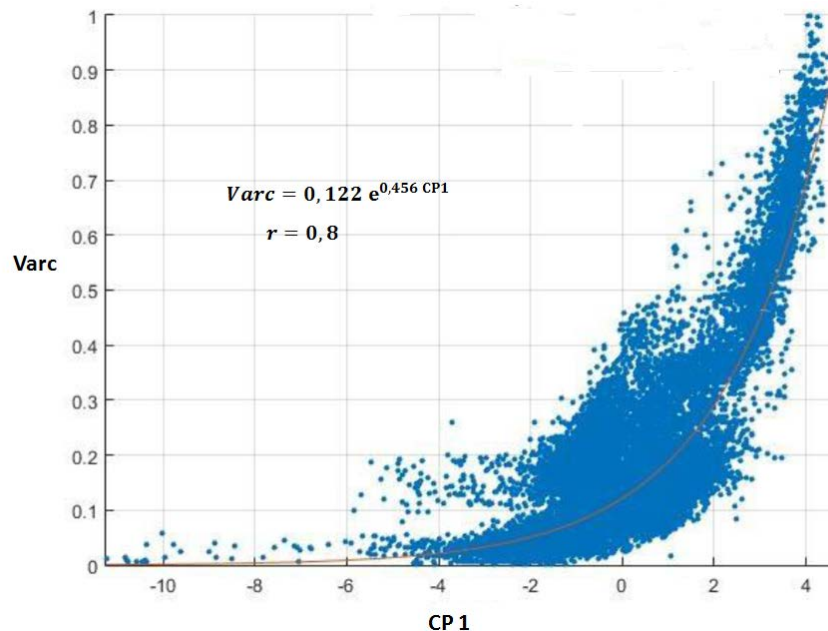


Figura 4 – Análisis de correlación estadística entre la primera componente principal (CP1) y el contenido de arcilla de las rocas (Varc).

La CP 2 explica el 16.5% de la variabilidad del sistema y varía de forma inversamente proporcional con la radiactividad y el PEF, es decir, incrementos de esta CP se asocian a rocas no radiactivas y con bajos valores del coeficiente de absorción fotoeléctrica. Este comportamiento indica que las electrofacies con predominio de

areniscas presentarán mayores valores de la CP 2.

La CP 3 explica el 11,2% de la variabilidad total del sistema y se asocia fuertemente y de manera inversa con el diámetro del pozo, es decir, bajos valores de esta CP reflejan altos valores del diámetro del pozo.

Se realizó el análisis de agrupaciones aplicando el algoritmo k medias sobre las tres primeras componentes principales. Fueron obtenidos diferentes modelos de clasificación y para seleccionar el número óptimo de grupos se

aplicó el Método del Codo (Moya, 2016), resultando óptimo el modelo de agrupación que considera cuatro clústeres. La tabla 7 presenta los valores medios de las tres primeras componentes principales para cada clúster.

Tabla 7 – Valores medios de las tres primeras componentes principales para cada clúster.

Clúster	N	CP1	CP2	CP3
1	3385	1.77	-0.43	-1.10
2	4898	-1.16	-0.78	-0.32
3	5252	-1.62	1.04	-0.07
4	5076	1.61	-0.04	1.12

Los clústeres 1 y 4 presentan altos valores de la CP 1 indicando la alta arcillosidad de estas rocas. Estos grupos logran separarse solo por el comportamiento que exhibe la CP 3, que para el clúster 1 presenta los menores valores del conjunto reflejando mayores valores del diámetro del pozo. Este comportamiento permite asociar los clústeres 1 y 4 a electrofacies principalmente compuestas por arcillas.

Los clústeres 2 y 3 presentan rocas con bajos valores de la CP 1 y valores medios de la CP 3 por tanto, menor arcillosidad. Estos grupos logran separarse por el comportamiento de la segunda componente principal, que se presenta con los menores valores para el clúster 2, reflejando mayor radiactividad y PEF, indicando la presencia de electrofacies con predominio de rocas carbonatadas enriquecidas con materia

orgánica, apreciable en la separación de los registros SGR y CGR, y que constituyen el reservorio principal de petróleo y gas en la zona bajo estudio.

El clúster 3 presenta mayores valores de la CP 2 reflejando baja radiactividad y coeficiente de absorción fotoeléctrica, asociado a electrofacies con predominio de areniscas.

Este comportamiento permite asociar los clústeres 1 y 4 a la formación Vega Alta, que constituye el sello regional de la FPNC compuesto por rocas de matriz arcillosa; y los clústeres 2 y 3 a los reservorios pertenecientes al Grupo Veloz, donde se describen rocas carbonatadas, fracturadas y porosas con intercalaciones de areniscas y arcillas.

La figura 5 presenta los gráficos cruzados de las componentes principales en el plano.

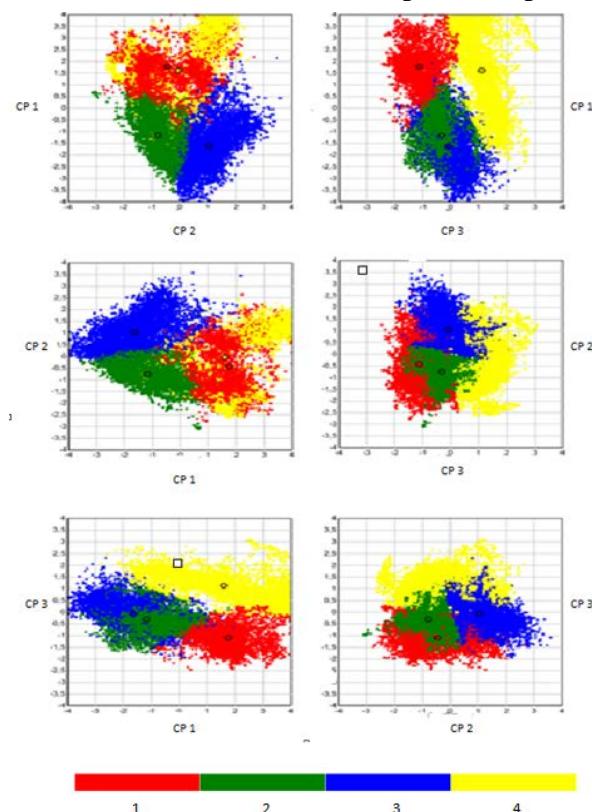


Figura 5 – Gráficos cruzados de las componentes principales en el plano. Muestran el poder resolutivo de las componentes principales para separar las cuatro electrofacies.

Se puede observar el poder resolutivo de las tres componentes principales para separar las cuatro electrofacies principales presentes en el

corte geológico. La figura 6 presenta las principales electrofacies obtenidas al aplicar el procedimiento descrito en el Pozo 1.

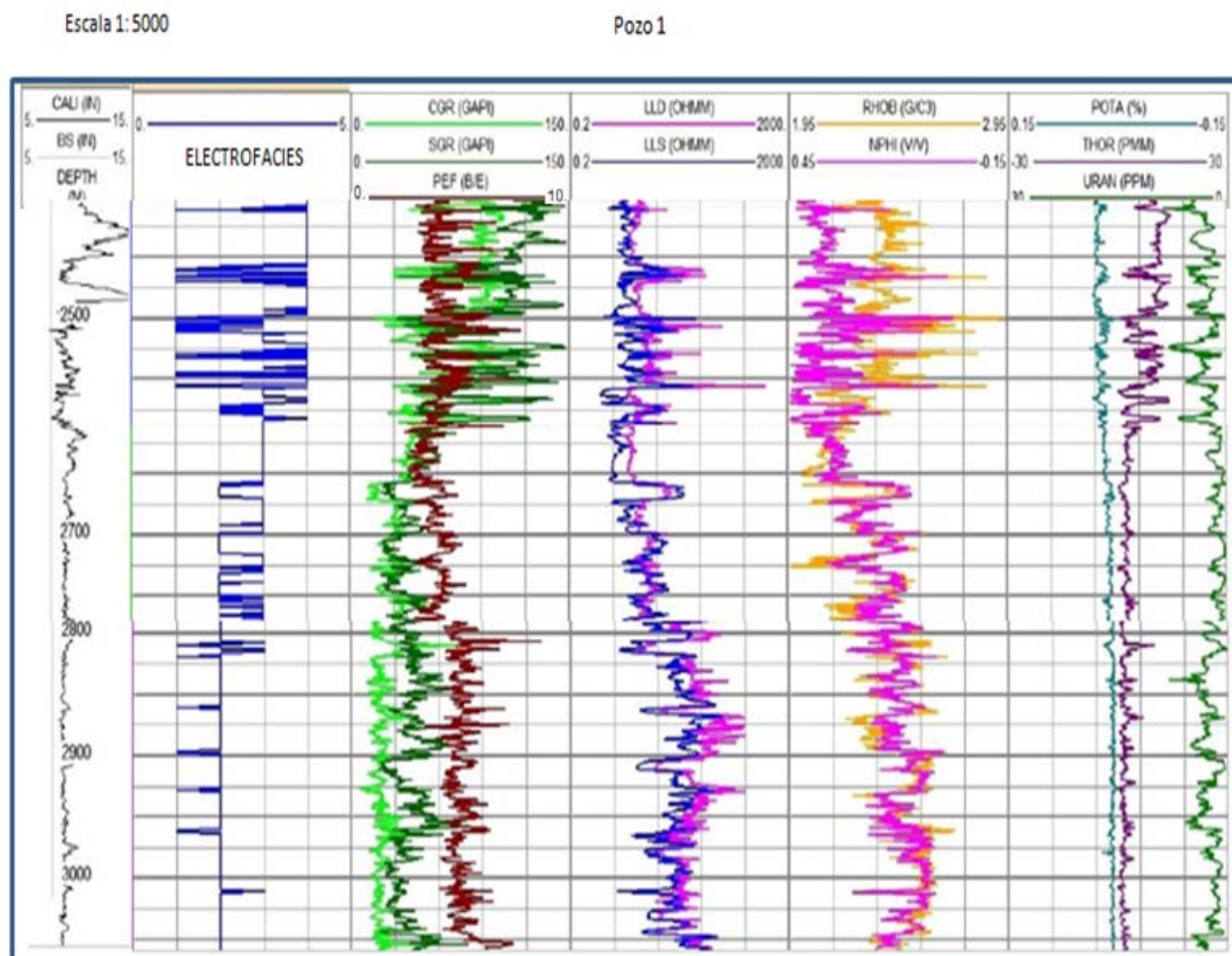


Figura 6 – Electrofacies principales obtenidas a partir del procedimiento descrito y la información geológica disponible en el Pozo 1.

CONCLUSIONES

La aplicación integrada del análisis de componentes principales y el análisis de clasificación no supervisada k-medias sobre los registros geofísicos de pozo, ha demostrado su poder resolutivo para identificar las electrofacies presentes en el corte geológico. La proyección en el plano de las componentes principales calculadas permite visualizar su efectividad en la separación de las electrofacies. Ha sido posible identificar dos electrofacies que constituyen

sellos, asociadas a la Formación Vega Alta; y 2 electrofacies asociadas a los reservorios pertenecientes al Grupo Veloz. Se destaca una electrofacies que constituye el reservorio principal en el área estudiada.

Se recomienda la aplicación de este procedimiento para evaluar las electrofacies en otros yacimientos de la Franja Petrolera Norte Cubana y en yacimientos de hidrocarburos de otras regiones del mundo.

REFERENCIAS

- ALBEAR, J.F. & ITURRALDE, M. Estratigrafía de las provincias de La Habana. En: Contribución a la Geología de las provincias de La Habana y Ciudad de La Habana. **Científico Técnico**, La Habana, p. 12-54, 1985.
- ALFONSO, J.R. **Estadísticas en las Ciencias Geológicas**, Tomo 2. ISPJAE, La Habana, 308 p., 1989.
- ALÍ, A. & SHENG-CHANG, C. Characterization of well logs using K-mean cluster analysis. **Journal of Petroleum Exploration and Production Technology**. 2020. <https://doi.org/10.1007/s13202-020-00895-4>.
- ALMAGUER, Y.; TOYRAC, R.; GIL, S. Estratigrafía del sector Morro-Bacuranao, provincia de La Habana, Cuba. **Minería y Geología**, v. 32, n. 2, p. 35-52, 2016.
- ÁLVAREZ, J.; VALLADARES, S.; RÍOS, J. Características de la Franja Norte de Crudos Pesados y su extensión hacia La Zona Exclusiva de Cuba en el Golfo de México. MEMORIAS IV CONFERENCIA LATINOAMERICANA DE GEOFÍSICA Y II CONGRESO CUBANO DE GEOFÍSICA. La Habana, Cuba. 2002. **Actas...** La Habana: Sociedad Cubana de Geología, 2002, 15 p.

- BREY DEL REY, D. Caracterización del reservorio en el área Habana del Este – La Cabaña. *Memorias V Convención Cubana de Ciencias de la Tierra, GEOCIENCIAS'2013*, Sociedad Cubana de Geología. La Habana, Cuba, 2013. *Actas...* La Habana: Sociedad Cubana de Geología, 2013, 15 p.
- BRÖNNIMANN, P. & RIGASSI, D. Contribution to the geology and paleontology of the area of the city of La Habana, Cuba and its surroundings. *Eclogae Geologicae Helvetiae*, v. 56, n. 1, p. 193-430, 1963.
- CASTRO, O. **Evaluación petrofísica de formaciones gasopetrólicas cubanas**. Centro Nacional de Información Geológica, La Habana, Cuba. 224 p., 2017.
- DARWIN, V.E. & SINGER, J.M. **Well Logging for Earth Scientists**. Springer, The Netherlands. 699 p., 2008.
- DAVIS, J. Electrofacies in reservoir characterization. p. 211-224, 2018. In: DAYA, B.S.; CHENG, Q.; AGTERBERG, F. (Eds.). **Handbook of Mathematical Geosciences**. Editora Springer, Disp. en; https://doi.org/10.1007/978-3-319-78999-6_11 Consultado el 5 de junio 2022.
- GÓMEZ, J. E. Exploración petrolera en Cuba. **Nuevo enfoque a la metodología del análisis de riesgo**, v. 11, n. 2, 2021.
- HAMDAN, H. & EMAD, L. K-means clustering algorithm applications in data mining and pattern recognition. **International Journal of Science and Research**, v. 6, n. 8, p. 1577-1584. 2017.
- HERNÁNDEZ-ORALLO, J.; RAMIREZ, M.J.; FERRI, C. **Introducción a la Minería de datos**. Pearson Educación, 680 p. 2004.
- IMRON, M.; HASANAH, U.; HUMAIDI, B. Analysis of DataMining Using K-Means Clustering Algorithm for Product Grouping. **International Journal of Informatics and Information System**, v. 3, n. 1, p. 12 - 22, 2020.
- INSTITUTO DE GEOLOGÍA Y PALEONTOLOGÍA, IGP. **Mapa Geológico de la República de Cuba a escala 1:100 000**. La Habana, Servicio Geológico de Cuba, 2016.
- KUMAR, B. & KISHORE, M. Electrofacies Classification – A Critical Approach. 6TH INTERNATIONAL CONFERENCE & EXPOSITION ON PETROLEUM GEOPHYSICS “KOLKATA 2006”. 2006. *Actas...* La Habana: Sociedad Cubana de Geología, p. 822-825, 2006.
- LEE, S.H.; ARUN, K.; AKHIL, D.G. Electrofacies characterization and permeability predictions in complex reservoirs. **Society of Petroleum Engineers**, v. 5, n. 3, p. 237-248, 2002.
- LIM, J.S.; KANG, J.M.; KIM, J. Multivariate Statistical Analysis for Automatic Electrofacies Determination from Well Log Measurements. **Society of Petroleum Engineers**. SPE 38028, p. 109-113, 1997.
- MORALES, C.; GARCÍA, D.; HERNÁNDEZ, J.; BREY DEL REY, D.; TAMAYO, Y.; ROSELL, Y.; CASTRO, O. Caracterización litológica de la Formación Vega Alta en la Franja Norte de Crudos Pesados. Propuesta para un nuevo enfoque geológico. *MEMORIAS V CONVENCION CUBANA DE CIENCIAS DE LA TIERRA, GEOCIENCIAS' 2013*, La Habana. 2013. *Actas...* LaHabana: Sociedad Cubana de Geología, 15 p., 2013.
- MOYA, R. **Selección del número óptimo de clusters**. 2016. Disp. en <https://jrrob.com/seleccion-del-numero-optimo-clusters>. Consultado el 4 de octubre 2021.
- NORMANDO, M.; MAIA DE ALMEIDA, N.; RODRIGUES DO NASCIMENTO, D.; SALGUEIRO, A. R. Classificação de eletrofácies de um poço petrolífero descobridor utilizando estatística multivariada: um estudo de caso no campo Curimã (Bacia do Ceará). **Revista Geociencias**, v. 40, n. 4, p. 911-924, 2021.
- OTERO, R.; PÉREZ, Y.; LÓPEZ, J.G.; ECHEVERRÍA, G.; RODRÍGUEZ, R. La exploración petrolera en la Franja Norte de Crudos Pesados Cubana. *PETROGAS 2009. MEMORIAS III CONVENCION CUBANA DE CIENCIAS DE LA TIERRA, GEOCIENCIAS'2009*, La Habana. *Actas...* LaHabana: Sociedad Cubana de Geología. La Habana, Cuba. 7 p., 2009.
- PUSKARCZYK, E. Application of Multivariate Statistical Methods and Artificial Neural Network for Facies Analysis from Well Logs Data: an Example of Miocene Deposits. **Energies**, v. 13, n. 1548, 2020.
- SERRA, O. & ABOTT, H.T. The contributions of loggings data to sedimentology and stratigraphy. **Society of Petroleum Engineers**, v. 22, n. 1, p. 117-131, 1982.
- WOLFF, M. & PELISSIER-COMBESCURE, J. FACI-OLOG – automatic electrofacies determination, Transactions of the SPWLA 23rd Annual Logging Symposium, Paper FF, 22 p., 1982

*Submetido em 28 de novembro de 2022
Aceito para publicação em 14 de abril de 2023*