

**PREDICCIÓN DE PROPIEDADES PETROFÍSICAS A TRAVES DE COMPONENTES  
PRINCIPALES Y REGRESIÓN LINEAL MÚLTIPLE**

Yesid Carreño Peñaloza

Proyecto de grado presentado como requisito para optar al título de  
Especialista en Estadística Aplicada

Asesor:

Sébastien Lozano Forero



Fundación Universitaria Los Libertadores

Departamento de Ciencias Básicas

Especialización en estadística aplicada

Bogotá D.C.

2017

**Nota de Aceptación**

---

---

---

---

---

---

---

Firma del presidente del jurado

---

Firma del Jurado

---

Firma del Jurado

Bogotá, D.C. 31 de julio de 2017

Las Directivas de la Universidad de  
Los Libertadores, los jurados calificadores y el cuerpo  
Docente no son responsables por los  
Criterios e ideas expuestas En el presente documento.  
Estos corresponden únicamente a los autores.

## Tabla de contenido

<b>Capítulo 1. Introducción</b> .....	8
Formulación del problema.....	9
Objetivos.....	10
<b>Capítulo 2. Marco de Referencia</b> .....	11
Yacimiento de petróleo .....	<b>11</b>
<b>mamar</b> Caracterización de yacimientos .....	<b>12</b>
Registros de Pozo.....	<b>12</b>
Gamma Ray (GR).....	12
Densidad de la Roca (RHOZ).....	12
Porosidad Neutrón (TNPH) .....	13
Perfil Sónico (DT) .....	13
Registro de Absorción Foto Eléctrica. (PEFZ).....	14
Registros de Resistividad (RT).....	14
Propiedades Petrofísicas de un Yacimiento .....	<b>14</b>
Volumen de Arcillosidad (VCLAY).....	14
Porosidad (PHIE).....	15
<b>Capítulo 3. Marco Teórico</b> .....	18
<b>Capítulo 4. Marco Metodológico</b> .....	22
Análisis de Componentes Principales (ACP) .....	<b>22</b>
Estadísticos de Adecuación Muestral .....	<b>23</b>
<b>Capítulo 5. Resultados</b> .....	33
<b>Capítulo 6. Conclusiones</b> .....	36
<b>Capítulo 7. Bibliografía</b> .....	37

## Tabla de Figuras

<b>Figura 1</b>	Configuración geológica de un yacimiento de petróleo. Recuperado de <a href="http://www.aggep.org">http://www.aggep.org</a> Asociación de Geólogos y Geofísicos Españoles del Petróleo. ....	11
<b>Figura 2</b>	Template del Set básicos de registros o variables y su respectivo modelamiento de las propiedades petrofísicas VCLAY, PHIE y Sw. De izquierda a derecha primer track profundidad; 2 Calidad del hueco; 3.GR 5. RT 6. RHOZ-TNPH 7.PHIE; 11 SW .....	17
<b>Figura 3</b>	Template del Set básicos de registros o variables. De izquierda a derecha primer track 1.GR, 2.RHOZ-TNPH-3.PEFZ-4.DT-5. RT .....	19
<b>Figura 4</b>	Variabes denotadas por $X_i$ ( $i = 1, 2, \dots, m$ ) en donde X es un vector que consta de N elementos y esta a su vez es mayor que m. ....	20
<b>Figura 5</b>	Gráfico de Sedimentación .....	27
<b>Figura 6</b>	Gráfico de Componentes Principales .....	28
<b>Figura 7</b>	Se encuentra la distribución en planta o espacial de los pozos PE-PEN y PPN para el campo. ....	30
<b>Figura 8</b>	Estructura de la regresión múltiple. ....	31
<b>Figura 9</b>	Estructura de aprendizaje y aplicación de la regresión lineal múltiple. En el aprendizaje se utilizan 1 pozo corazonado y en la aplicación lo mismo. ....	32
<b>Figura 10</b>	Template de registros de pozo. De izquierda a derecha: 1.Profundidad, 2.GR, 3.RHOZ-TNPH, 4.PEFZ-DT, 5. Comparación de los modelos VCL 6. Comparación de los modelos PHIE.....	34
<b>Figura 11</b>	Crossplot entre el VCLAY realizado por modelos petrofísicos y el VCLAY obtenido con la regresión lineal múltiple con componentes principales. ....	35
<b>Figura 12</b>	Crossplot entre la porosidad efectiva realizado por modelos petrofísicos y la porosidad efectiva obtenida con la regresión lineal múltiple con componentes principales.....	35

## **Lista de Tablas**

<b>Tabla 1.</b>	Matriz de Correlaciones.....	24
<b>Tabla 2</b>	Estadísticos de adecuación de la muestra .....	24
<b>Tabla 3.</b>	Tabla de Comunalidades.....	25
<b>Tabla 4.</b>	Varianza Total Explicada.....	27
<b>Tabla 5.</b>	Matriz de componentes o pesos factoriales de las variables en cada componente.....	28

## **PREDICCIÓN DE PROPIEDADES PETROFÍSICAS A TRAVÉS DE ANÁLISIS DE COMPONENTES PRINCIPALES Y REGRESIÓN LINEAL MÚLTIPLE.**

### **RESUMEN**

Los modelos petrofísicos, son las técnicas más usadas en la industria del petróleo a la hora de predecir propiedades de Volumen de Arcilla (VCL) y Porosidad Efectiva (PHIE), son realizados a partir de la aplicación de ecuaciones petrofísicas a las variables Gamma Ray (GR), Densidad (RHOZ) Neutrón-Porosidad (TNPH), Absorción Fotoeléctrico (PEFZ), Sónico (DT) y Resistivos (RT), estos modelos petrofísicos son aprobados o descartados mediante pruebas de laboratorio en un núcleo de roca y de producción de crudo en un pozo petrolero, la cuales demandan tiempo y dinero.

Este trabajo usó técnicas estadísticas como lo son el Análisis de Componentes Principales (ACP) y regresiones lineales múltiples para la predicción de 2 propiedades petrofísicas muy usadas en el momento de realizar el completamiento de un pozo de petróleo, Volumen de Arcillosidad (VCLAY) y la Porosidad (PHIE).

El análisis de componentes principales y la regresión lineal múltiple, se aplicaron a las variables mencionadas anteriormente; las cuales se toman de una base de datos conformada por la información de registros de cada pozo perforado y puesto en producción en un campo petrolero con características geológicas y estructurales homogéneas.

Los resultados de VCLAY, PHIE; obtenidos por el método convencional de un modelo petrofísico y por las técnicas estadísticas fueron comparados uno a uno, caracterizando litologías como las rocas reservorio y las rocas sellos. Esto demuestra que mediante la aplicación de técnicas estadísticas se pueden tomar decisiones en el completamiento de un pozo.

**Palabras Claves:** Perfiles de Pozos, Modelos Petrofísicos, Análisis de Componentes Principales, Volumen de Arcilla, Porosidad y Saturación de Agua.

## **PREDICTION OF PETROPHYSICAL PROPERTIES THROUGH PRINCIPAL COMPONENTS AND MULTIPLE LINEAR REGRESSIONS**

### **ABSTRACT**

Petrophysical models, is the most widely used technique in the oil industry to predicting these properties, they are made from the application of Petrophysical equations to the follow variables Gamma Ray (GR), density (RHOZ) Neutron (TNPH), photoelectric factor (PEFZ) Sonic (DT) and resistive (RT), these petrophysical models are approved or discarded by physical testing laboratory and production of crude oil on well, which require time and money.

This study used statistical techniques such as principal components and multiple linear regressions for the prediction of the 2 Petrophysical properties most commonly used in performing of oil well completion, clay volume (VCLAY) and porosity (PHIE).

Principal component analysis (PCA) and multiple linear regressions were applied to the variables mentioned above; which are taken from a database formed by information of each logs oil well drilled and put into the production in a same oil field with geological and structural homogeneous features.

The results of VCLAY and PHIE; obtained by the conventional method of a petrophysical model and statistical techniques were compared one by one. This proves that by applying statistical techniques can make decisions on well completion.

**Key words:** well profiles, Petrophysical models, Principal component analysis, clay volume, porosity, and water saturation.

## **Capítulo 1. Introducción**

Los modelos de las propiedades petrofísicas como lo son el Volumen de Arcillosidad (VCLAY) y Porosidad Efectiva (PHIE), son de gran importancia para caracterizar yacimientos petrolíferos.

Estas propiedades petrofísicas fueron calculadas aplicando ecuaciones de VCLAY y PHIE, a las variables Gamma Ray (GR), Densidad (RHOZ) Porosidad Neutrón (TNPH), Absorción Fotoeléctrico (PEFZ), Sónico (DT) y Resistivos (RT), y son aprobadas o descartadas mediante pruebas físicas como Difracción de Rayos X (DRX) y Microscopia de Barrido Electrónico (SEM) para arcillosidad; pruebas de porosímetro de Boyle, Mercurio y Helio para PHIE. Estas pruebas son realizadas en laboratorio a núcleos de roca extraídos en pozos anteriormente perforados aumentando significativamente los costos y los tiempos en la toma de decisiones, creando la necesidad de aplicar modelos más rápidos, económicos y más confiables que los Quick Log, donde solo se observan las variables y se definen los intervalos.

El Análisis de Componentes Principales y la Regresión Lineal Múltiple han demostrado ser una nueva, rápida, económica y acertada herramienta Moss y Seheult (1987), Willhite, y Doveton, (2012), para caracterizar yacimientos con ecuaciones que predicen propiedades petrofísicas VCLAY y PHIE, realizando el completamiento de los pozos en menor tiempo que lo convencional y disminuyendo costos en esta etapa.

## **Formulación del problema**

Luego de ser perforado un pozo de petróleo, se debe realizar la caracterización del yacimiento, se identifican los intervalos en profundidad con presencia de hidrocarburos los cuales quedaran abiertos a producción y los intervalos con agua que deben quedar cerrados, esta caracterización se realiza en la industria con los siguientes métodos:

- **Quick Log:** Es la interpretación empírica de los registros de pozo para caracterizar las zonas donde pudieran existir acumulación de hidrocarburos. Se realiza de manera rápida sin tener en consideración todas las variables del caso. Alta probabilidad de error.
- **Modelo Petrofísico:** Para realizar la caracterización del yacimiento es necesario realizar el cálculo de las Propiedades Petrofísicas mediante la integración de registros de pozo con pruebas realizadas a rocas en los laboratorios. Demanda tiempo y dinero.

## **Justificación**

Se propone reducir significativamente el intervalo de tiempo de la caracterización de un yacimiento aplicando una metodología de predicción de propiedades petrofísicas basada en la combinación de las técnicas estadísticas Análisis de Componentes Principales (ACP) y Regresión Lineal Múltiple.

## **Objetivo General**

Reducir de manera significativa y con alto grado de confiabilidad, el tiempo de la caracterización de un yacimiento en la etapa de completamiento de pozos petroleros.

### **Objetivos Específicos**

Reducir el número de variables a las componentes principales necesarias que expliquen en mayor parte la información y combinarlos con regresión lineal múltiple con el fin de predecir propiedades petrofísicas.

## Capítulo 2. Marco de Referencia

### Yacimiento de petróleo

“Los yacimientos son rocas, en su mayoría areniscas y carbonatos, empapadas de petróleo. Están llenas de poros interconectados que actúan absorbiendo el crudo como una esponja”

Se forman por acumulación durante millones de años de microorganismos muertos. “El plancton del mar va muriendo y se acumula en el fondo marino mezclado con sedimento con el paso del tiempo otros sedimentos los cubren, formando sucesivas capas que van aumentando la presión sobre la primera capa. En estas condiciones se producen una serie de reacciones químicas de descomposición que dan lugar a los hidrocarburos.

Pero para que se forme un yacimiento se necesita algo más, que caiga en una trampa. El crudo se desplaza desde donde se ha formado hasta la roca porosa moviéndose por micro fracturas de las rocas, que son como pequeños canales, hasta llegar a la mencionada roca porosa donde se almacena. Si ésta recubierta de un material impermeable, como arcillas (Rocas sello), el petróleo queda atrapado, Sánchez V. (2014) (Figura 1).



**Figura 1.** Configuración geológica de un yacimiento de petróleo. *Recuperado de <http://www.aggep.org> Asociación de Geólogos y Geofísicos Españoles del Petróleo.*

## **Caracterización de yacimientos**

Es un proceso en el cual son aplicados diversos conocimientos sobre ingeniería para Interpretar lógicamente todos los datos y características del yacimiento mediante diferentes herramientas y técnicas modernas orientados a la definición y al estudio de características geológicas, petrofísicas y dinámicas que controlan la capacidad de almacenamiento y de producción de los yacimientos petroleros.

A continuación, se describirá brevemente los registros de pozo convencionales, que son las variables usadas como entrada en modelos petrofísicos para el cálculo de propiedades petrofísicas Volumen de Arcillosidad (VCLAY) y porosidad (PHIE).

### **Registros de Pozo**

#### *GAMMA RAY (GR)*

El perfil de rayos gamma registra la radioactividad natural de los elementos Torio (Th) Uranio (U) y Potasio (K), estos se encuentran en las rocas arcillosas, por esta razón pueden usarse para identificar litologías. Las areniscas libres de arcillas y los carbonatos tienen concentraciones bajas de material radiactivo, y dan bajas lecturas de Gamma Ray (GR). Cuando el contenido de arcilla aumenta, la respuesta del perfil GR aumenta debido a la concentración de material radiactivo (Schlumberger,2000).

#### *DENSIDAD DE LA ROCA (RHOZ)*

El perfil de densidad de formación es la suma ponderada de las densidades de la matriz de la roca y los fluidos presentes dentro de esta. (Schlumberger, 2000).

El dispositivo, que registra la densidad, es una herramienta de contacto que consiste en una fuente de rayos gamma de energía media, los rayos chocan con los electrones en la formación; las colisiones causan una pérdida de energía de la partícula de rayos gamma. Tittman y Wahl (2005) llamaron a la interacción entre las partículas

de rayos de gamma entrantes y los electrones en la formación, Efecto Compton. Los rayos de gamma esparcidos que alcanzan el detector se localizan a una distancia fija de la fuente de rayo de gamma, se cuentan como un indicador de densidad de la formación

El número de colisiones por efecto Compton son una función directa del número de electrones en una formación (densidad de electrones). Por consiguiente, como este efecto está relacionado con la densidad de electrones se puede calcular la densidad ( $\rho_d$ ) de una formación en gm/cc.

#### *POROSIDAD NEUTRÓN (TNPH)*

Los perfiles neutrónicos son perfiles de porosidad que miden la concentración de iones de hidrógeno en una formación. Se bombean rayos gamma desde una fuente que chocan con los electrones del hidrogeno que se encuentran en la formación, estas colisiones causan una pérdida de energía. En las formaciones limpias (% bajo de arcillas) donde el espacio poroso se encuentra ocupados por hidrocarburos o agua, el neutrón detecta bajas cuentas lo cual es relacionado a altas porosidades (Schlumberger, 2000).

#### *PERFIL SÓNICO (DT)*

El perfil sónico es un perfil de porosidad que mide tiempo de tránsito de intervalo ( $\Delta T$ ) de una onda de sonido compresional que viaja a través de un pie de formación. El tiempo de tránsito del intervalo ( $\Delta T$ ), medido en microsegundos por pie, es el recíproco de la velocidad (medida en pies por segundo) de una onda compresional de sonido (Schlumberger, 2000).

### *REGISTRO DE ABSORCIÓN FOTO ELÉCTRICA. (PEFZ)*

El registro mide el factor de absorción fotoeléctrica,  $Pe$ , que se define como  $(Z/10)$ , donde  $Z$  es el número atómico promedio de la formación.  $Pe$  no tiene unidades, pero dado que es proporcional a la sección transversal fotoeléctrica por electrón, a veces se expresa en barnios/electrón. Dado que los fluidos tienen números atómicos muy bajos, su influencia es muy escasa, de modo que  $Pe$  es una medida de las propiedades de la matriz de roca. Las areniscas tienen un  $Pe$  bajo, en tanto que las dolomías y las calizas poseen un  $Pe$  alto. Las arcillas, los minerales pesados y los minerales que contienen hierro poseen un  $Pe$  alto. Por consiguiente, el registro es de gran utilidad para la determinación de la mineralogía Moake, Y Schultz, (1989).

### *REGISTROS DE RESISTIVIDAD (RT)*

Los perfiles de Resistividad son los perfiles eléctricos utilizados para determinar las zonas productivas de hidrocarburo vs las productoras de agua, puesto que la matriz de las rocas o los granos, no son conductores, la habilidad de la roca de transmitir una corriente es una función casi completamente condicionada por la presencia de fluidos en los poros.

Los hidrocarburos, como la roca y la matriz, no son conductores; por consiguiente, cuando la saturación de hidrocarburo de los poros aumenta, la resistividad de la roca también aumenta (Schlumberger, 2000).

### **Propiedades Petrofísicas de un Yacimiento**

#### *VOLUMEN DE ARCILLOSIDAD (VCLAY)*

La presencia de Arcillosidad en la roca yacimiento, es un factor altamente sensible a la hora de realizar evaluaciones de formaciones, afecta las habilidades del yacimiento para almacenar fluidos y complica la estimación del volumen de hidrocarburos.

Se determina tanto por métodos directos como indirectos:

**Directos.** Difracción de rayos X (DRX) (Directo prueba de laboratorio), Microscopia electrónica de barrido (SEM) (Directo prueba de laboratorio)

**Indirectos.** El volumen de arcilla también puede estimarse mediante la utilización de correlaciones empíricas publicadas en la literatura, una vez determinando el indicador de arcillosidad a partir del GR (Archie, G.E, 1942).

Ecuación 1. Volumen de Arcillosidad

$$IGR = \frac{GRLog - GRMin}{GRMax - GRMin} \quad Ec. 1$$

#### *POROSIDAD (PHIE)*

Es la medida del espacio intersticial, espacio existente entre grano y grano, se define como la relación entre el volumen poroso y el volumen total de la roca, entendiéndose como volumen poroso, al volumen total menos el volumen de los granos contenidos en dicha roca.

La porosidad representa la capacidad que tiene una roca de almacenar los fluidos, por esto es de gran importancia ya que una formación rocosa que sea productora tiene valores de porosidades altos.

Se determina tanto por métodos directos como indirectos.

**Métodos directos.** En la determinación de la porosidad en el laboratorio es necesario conocer o evaluar los dos parámetros que la definen, el volumen total (Vt) y el volumen sólido (Vs) o el volumen poroso (Vp). Para ello se usan muestras o núcleos obtenidos en los pozos, previamente tratados y preparados.

**Métodos indirectos.** A pesar de no existir un registro que mida directamente la porosidad de una roca, los investigadores han podido definir y utilizar diferentes métodos para calcular esta propiedad a partir de la data proporcionada por algunos registros, Fertl, (1986).

Si las consideraciones sobre las cuales se basan dichos métodos se cumplen, los resultados obtenidos son usualmente correctos. Sin embargo, aun cuando se consideren todas las correcciones recomendadas en los cálculos, se recomienda validar y calibrar los valores de porosidad resultantes del análisis de perfiles, comparándolos con los resultados obtenidos a través de los análisis de núcleos, mediante la realización de una correlación núcleo – perfil Babadagli, T; Alsalmi, S, (2002).

A continuación, se muestra las ecuaciones aplicadas para la determinación de la porosidad a través de perfiles.

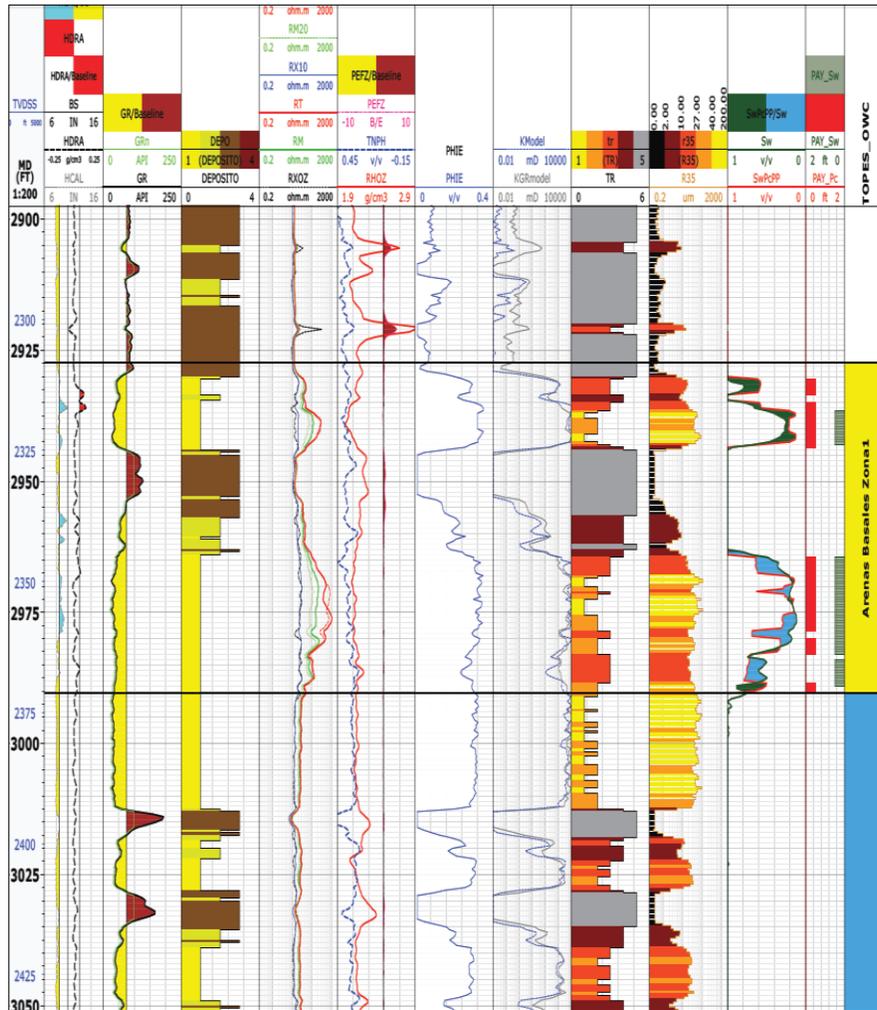
Ecuación 2. Porosidad calculada con el registro DT:

$$PHI = \frac{\Delta t_{log} - \Delta t_{mat}}{\Delta t_{fl} - \Delta t_{mat}} \quad Ec. 2$$

Ecuación 3. Porosidad calculada a partir del RHOZ

$$PHI = \frac{RHOB_{ma} - RHOB_{log}}{RHOB_{ma} - RHOB_f} \quad Ec. 3$$

En la figura 2 podemos observar el conjunto de registros de pozos o variables y el modelo petrofísico realizado por los métodos convencionales para el cálculo de las propiedades petrofísicas promedio para la zona de interés.



**Figura 2** Template del Set básicos de registros o variables y su respectivo modelamiento de las propiedades petrofísicas VCLAY, PHIE y Sw. De izquierda a derecha primer track profundidad; 2 Calidad del hueco; 3.GR 5. RT 6. RHOZ-TNPH 7.PHIE; 11 SW

### Capítulo 3. Marco Teórico

El Análisis de Componentes Principales (ACP), estudia las relaciones que se presentan entre un conjunto de variables correlacionadas (*Miden información común*) y la transformación de estas variables en otro conjunto de nuevas variables incorreladas entre sí (*sin repetición o redundancia en la información*) llamado conjunto de componentes principales. Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra.

El Análisis de Componentes Principales, se ha aplicado con éxito para una variedad de aplicaciones de datos de registros de pozos tales como identificación y caracterización de intervalos con baja permeabilidad y caracterización de unidades litoestratigráficas (Kassenaar, 1989) se puede utilizar en conjunto con Regresión Lineal Múltiple para obtener ecuaciones que predicen las propiedades petrofísicas de las rocas como el Volumen de Arcilla (VCLAY) y la Porosidad Efectiva (PHIE).

Supongamos que tenemos  $m$  variables estadísticas representadas por una muestra de  $N$  elementos. ( $N$  es generalmente mucho mayor que  $m$ ). En el caso de análisis de registros de pozo, estas variables  $m$  serán (GR, RHOZ, TNPH, DT, PEFZ y RT) medidos en el mismo intervalo y en los mismos  $N$  sitios de profundidad (Figura 3 flechas rojas). Denotados por  $X_i$  ( $i = 1, 2, \dots, m$ ) en donde  $X$  es un vector que consta de  $N$  elementos (Figura 4).

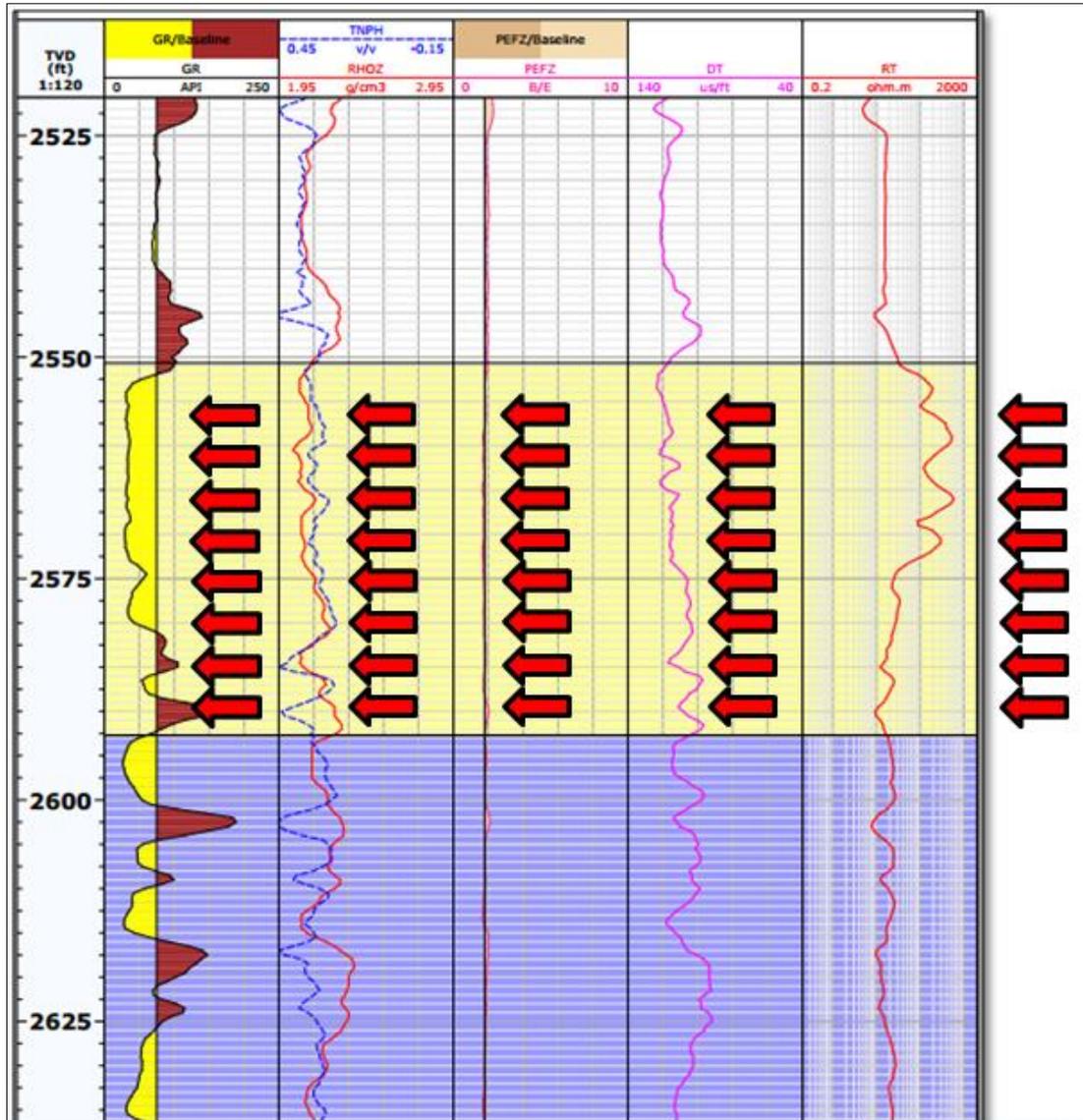


Figura 3 Set básicos de registros o variables. De izquierda a derecha primer track 1.GR, 2.RHOZ-TNPH-3.PEFZ-4.DT-5. RT.

Xi	1	2	..	..	..	....	m
	DEPT	GR	TNPH	RHOZ	PEFZ	DT	RT
1	2525.5	72.9	0.33	2.18	1.89	113.67	17.08
2	2526	72.26	0.33	2.15	1.86	115.99	16.98
.	2526.5	73.08	0.35	2.13	1.84	117.25	16.59
.	2527	71.59	0.37	2.11	1.85	117.19	16.33
.	2527.5	73.87	0.38	2.11	1.85	116.7	16.09
.	2528	77.71	0.37	2.12	1.86	116.08	16.13
.	2528.5	77.41	0.37	2.13	1.87	116.54	16.34
.	2529	75.64	0.36	2.12	1.89	117.22	15.94
.	2529.5	77.14	0.37	2.1	1.92	118.25	15.34
.	2530	79.69	0.36	2.1	1.95	119.29	15.37
.	2530.5	77.91	0.37	2.1	1.97	119.91	15.69
.	2531	77.03	0.38	2.1	2	120.01	15.85
.	2531.5	73.33	0.38	2.11	1.98	121.5	15.98
.	2532	74.31	0.37	2.11	1.96	121.15	16.24
.	2532.5	73.05	0.36	2.11	1.94	120.22	16.03
.	2533	73.94	0.37	2.1	1.98	119.66	15.85
N	2533.5	76.02	0.37	2.09	2.04	119.93	15.82

**Figura 4** Variables denotadas por  $X_i$  ( $i = 1, 2, \dots, m$ ) en donde  $X$  es un vector que consta de  $N$  elementos y esta a su vez es mayor que  $m$ .

El concepto básico de análisis de componentes principales es el siguiente: Se suponen las variables observadas  $m$ , se determinan por algunos factores de fondo “invisibles” (su número,  $n$  es generalmente menor que  $m$ ), en el caso de análisis de registros, los factores pueden ser entre otros, el volumen de arcilla y porosidad efectiva de la roca.

Estos factores no se pueden medir directamente, pero influyen en las mediciones de registros; el objetivo del análisis de los registros de pozo es de hecho la estimación de estos factores de los registros medidos. Denotemos estos factores por  $V_j$  ( $j = 1, 2, \dots, n$ ). Suponemos que hay una relación lineal entre las variables observadas  $X_i$  y los factores de fondo  $V_j$ :

Ecuación 4. Relación lineal Variables-Factores

$$x_i = \sum_{j=1}^n b_{i,j} * V_j + \epsilon_i \quad Ec. 4$$

Donde  $\epsilon_i$  es la parte de la variable  $X_i$  que no se explica por los factores  $V_j$ ; lo consideramos como "error" o "ruido" que no contiene información útil.

Nuestra tarea es determinar los factores ocultos  $V_j$ ,  $b_i$  y  $j$  coeficientes de las variables  $X_i$  observados. Hacemos algunas suposiciones antes de la formulación matemática exacta de un proceso de optimización.

Todas las relaciones entre las variables originales  $X_i$  y entre las variables  $X_i$  y los factores desconocidos  $V_j$  son lineales. Suponiendo la linealidad se deduce que queremos construir los factores como combinaciones lineales de las variables estadísticas originales.

Los factores son estadísticamente independientes es decir, su correlación es cero, incorreladas entre sí, cuyas varianzas vayan decreciendo progresivamente.

Ecuación 5. Factores a partir de las variables

$$v_j = \sum_{i=1}^m a_{i,j} * x_i \quad Ec. 5$$

Ecuación 6. Varianza de las variables

$$var(x_i) = \sum_{j=1}^n b_{i,j}^2 * var(V_j) + var(\epsilon_i) \quad Ec. 6$$

## **Capítulo 4. Marco Metodológico**

### **Análisis de Componentes Principales (ACP)**

La Porosidad Efectiva (PHIE), es la propiedad petrofísica de la roca que tiene la mayor influencia en los registros de pozos convencionales escogidos para este estudio. Esto es cierto para los RHOZ, TNPH y DT, los cuales miden los espacios o el volumen poroso.

Resistividad (RT) es un indicador no directo de la porosidad, ya que las corrientes aplicadas en la formación rocosa se transmiten por los fluidos que esta almacena, a medida que esta presenta mayor acumulación fluidos, tendrá valores de porosidades mayores.

Registros Gamma Ray (GR) muestran una relación inversamente proporcional con la porosidad efectiva, a mayor radioactividad (valores de GR altos) significa mayor Volumen de Arcilla lo cual disminuye las porosidades efectivas en las rocas.

Se puede concluir de lo anterior, que el factor  $V_i$  más correlacionable en las mediciones de registro  $X_i$  en las rocas del yacimiento en el mismo intervalo y en los mismos  $N$  puntos de profundidad es la porosidad (Figura 3).

Lo anterior indica que los  $V_i$  formados a partir del Análisis de Componentes Principales contendrán la mayor información de Porosidad Efectiva en su primera componente (Crains Petrohysical Handbook).

Por estas razones, el análisis por componentes principales, se pueden utilizar para la estimación de la porosidad efectiva y otro tipo de variables como el volumen de arcilla.

## **Estadísticos de Adecuación Muestral**

Para que se pueda realizar el análisis de componentes principales, es necesario que las variables presenten factores comunes, es decir, que estén muy correlacionadas entre sí. Los coeficientes de la matriz de correlaciones deben ser grandes en valor absoluto.

Para comprobar que las correlaciones entre las variables son distintas de cero aplicamos el Test de Bartlett donde se prueba que si el determinante de la matriz es distinto de uno, la matriz de correlaciones es distinta de la matriz identidad (Tabla 1.)

$$H_0 = |R| = 1$$

$$H_1 = |R| \neq 1$$

En este caso el determinante es 0.008, el cual es próximo a cero, indica que una o más variables podrían ser expresadas como una combinación lineal de otras variables (Tabla 1).

Los valores altos obtenidos en la distribución chi cuadrado y los valores de significancias críticas  $< 0.05$  en este caso 0 llevan a rechazar  $H_0$  (Tabla 2).

Se observa que el valor del  $KMO = 0.583$  está en el límite de los valores recomendados por Kaiser, Meyer y Olkin, pero los resultados del test de Bartlett nos permiten aplicar la técnica de análisis de componentes principales a este grupo de variables (Tabla 2) .

**Tabla 1** Matriz de Correlaciones (Tomado de software IBM SPSS)

**Matriz de correlaciones<sup>a</sup>**

		DT	GRn	PEFZ	RHOZ	RT	TNPH
Correlación	DT	1,000	-,025	-,433	-,763	-,142	,702
	GRn	-,025	1,000	,576	,214	-,276	,348
	PEFZ	-,433	,576	1,000	,618	-,023	-,267
	RHOZ	-,763	,214	,618	1,000	,345	-,801
	RT	-,142	-,276	-,023	,345	1,000	-,473
	TNPH	,702	,348	-,267	-,801	-,473	1,000
Sig. (Unilateral)	DT		,115	,000	,000	,000	,000
	GRn	,115		,000	,000	,000	,000
	PEFZ	,000	,000		,000	,139	,000
	RHOZ	,000	,000	,000		,000	,000
	RT	,000	,000	,139	,000		,000
	TNPH	,000	,000	,000	,000	,000	

a. Determinante = ,008

**Tabla 2** Estadísticos de adecuación de la muestra (Tomado de software IBM SPSS)

**KMO y prueba de Bartlett**

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,583
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	11116,778
	gl	15
	Sig.	,000

### Número de factores a extraer.

Las denominadas comunalidades en el inicio de una técnica ACP, siempre son iguales a uno, cobrando sentido en los componentes obtenidos al final. Al considerar tan sólo un número reducido de factores entre todas las variables observadas, la varianza total no queda totalmente explicada; por tanto, sus valores oscilarán entre cero y uno, es decir, entre la posibilidad de que los factores comunes no expliquen nada de la variabilidad de una variable o que por el contrario ésta quede totalmente explicada por los factores comunes.

**Tabla 3.**Tabla de Comunalidades (Tomado de software IBM SPSS)

Comunalidades		
	Inicial	Extracción
DT	1,000	,864
GRn	1,000	,920
PEFZ	1,000	,835
RHOZ	1,000	,935
RT	1,000	,977
TNPH	1,000	,939

Método de extracción:  
Análisis de Componentes  
principales.

La tabla **¡Error! No se encuentra el origen de la referencia.**3 presenta las comunalidades obtenidas para cada una de las variables, una vez realizada la extracción. Así, podemos comprobar que variables tales como la PEFZ con un 0.835 de comunalidad explica en menor proporción la varianza según su participación en los componentes resultantes en el análisis, mientras que variables TNPH, RHOZ y DT participan en mayor medida con un 0.939, 0.935 y 0.864 respectivamente, en la explicación de dicha varianza.

Teniendo en cuenta la tabla 4 de varianza total explicada podemos determinar con cuantos componentes principales vamos a trabajar. Existen reglas para saber el número más adecuado a conservar, por ejemplo, la que se conoce como criterio de Kaiser que indica que hay que conservar los componentes principales cuyos valores propios son mayores que la unidad; para este caso serían los dos primeros componentes número 1 y 2.

El criterio más utilizado es el de observar el porcentaje de varianza total explicada por cada componente o factor y cuando éste llega a un porcentaje acumulado considerado alto, normalmente superior a 80%, significa que el número de factores es suficiente. En este caso los dos primeros factores explican el 78.8% de la varianza estando en el límite de lo aceptable (Tabla 4).

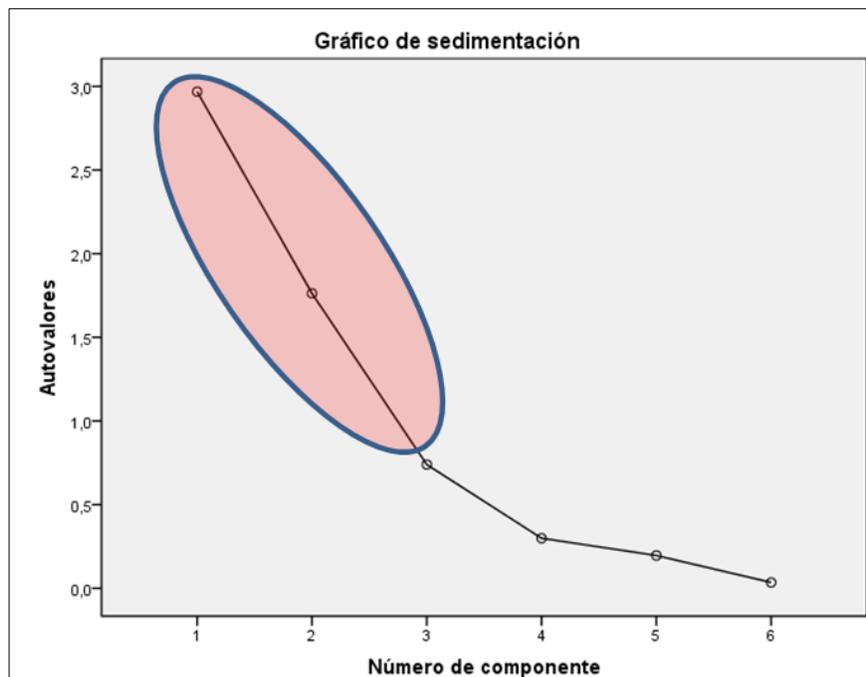
Podemos observar en la Figura 5, que a partir del componente número 3 el valor propio es inferior a la unidad y según el criterio de la figura de sedimentación, se extraen todas las componentes que están situadas previamente a la zona de sedimentación, entendiendo por ésta la parte del gráfico en la que los componentes presentan un cambio de pendientes fuertes, en nuestro caso las tres primeras componentes. Además en la componente número 3 el porcentaje de la varianza explicada acumulada asciende al 91.171 % (Tabla 4).

Por las razones anteriormente explicadas se escogen que el número de componentes ideal luego de realizar el análisis es de 3 (Tabla 5).

**Tabla 4.** Varianza Total Explicada (Tomado de software IBM SPSS)

Varianza total explicada						
Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,968	49,469	49,469	2,968	49,469	49,469
2	1,763	29,378	78,848	1,763	29,378	78,848
3	,739	12,323	91,171	,739	12,323	91,171
4	,299	4,985	96,156	,299	4,985	96,156
5	,196	3,266	99,422	,196	3,266	99,422
6	,035	,578	100,000	,035	,578	100,000

Método de extracción: Análisis de Componentes principales.

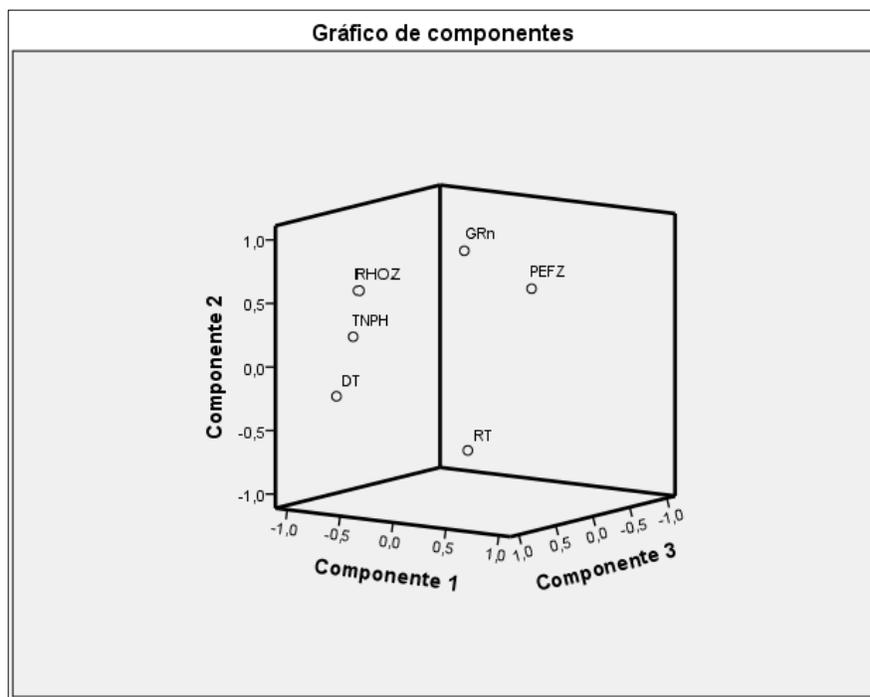


**Figura 5** Gráfico de Sedimentación (Tomado de software IBM SPSS)

**Tabla 5.** Matriz de componentes o pesos factoriales de las variables en cada componente. (Tomado de software IBM SPSS)

	Componente		
	1	2	3
DT	-,850	-,012	,376
GRn	,088	,915	,274
PEFZ	,627	,651	,135
RHOZ	,961	,102	,033
RT	,418	-,561	,698
TNPH	-,864	,419	,131

Método de extracción: Análisis de componentes principales.  
a. 3 componentes extraídos



**Figura 6** Gráfico de Componentes Principales (Tomado de software IBM SPSS)

En la tabla 5 podemos observar, la primera componente principal asocia las variables RHOZ, TNPH y DT las cuales tienen los mayores pesos, esta asociación se debe a que estos registros o variables son usados en la industria mediante la aplicación de ecuaciones petrofísicas para el cálculo de la PHIE, esto nos indica que la primera componente estará más asociada a la porosidad de la roca.

La segunda componente principal asocia las variables GR y PEFZ las cuales tienen los mayores pesos, esta asociación se debe a que estos registros o variables miden la litología del yacimiento (Figura 6).

La tercera componente principal refleja la variabilidad de los datos relacionados con el fluido. El mayor peso lo presenta la variable RT.

### **Análisis de regresión lineal múltiple con las componentes principales y propiedades petrofísicas.**

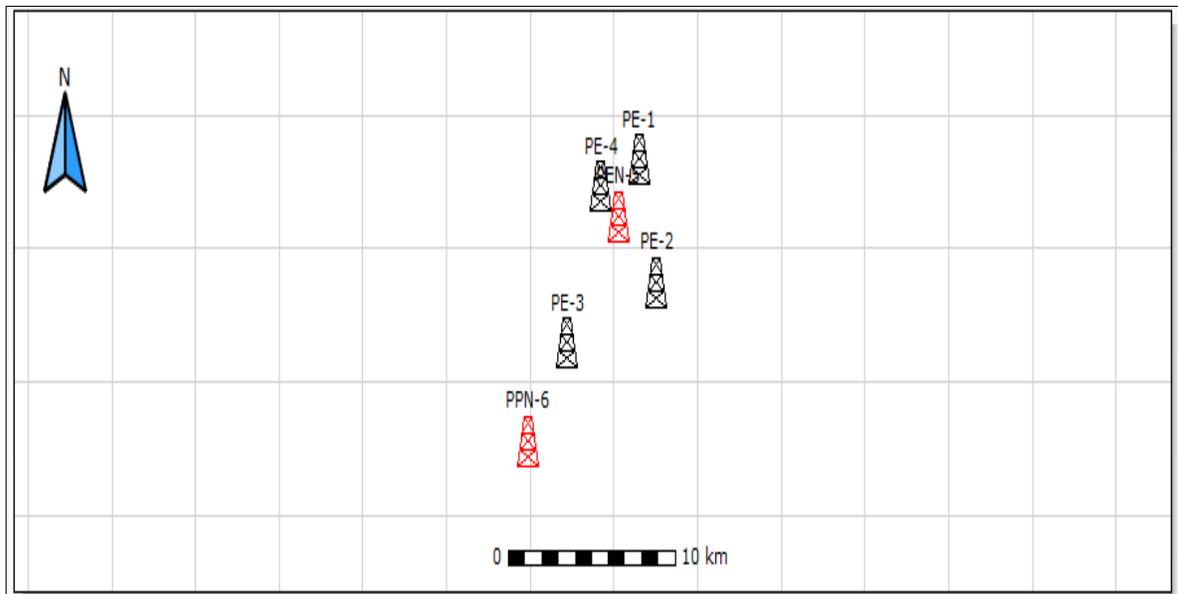
Las propiedades petrofísicas en la industria del petróleo, son calculadas aplicando ecuaciones de arcillosidad y porosidad a las variables GR, RHOZ TNPH, PEFZ, DT y RT, como se explicó en el marco de referencia y son aprobadas o descartadas mediante pruebas físicas realizadas a los núcleos de roca como Difracción de Rayos X (DRX) y Microscopio Eléctrico de Barrido (SEM) para VCLAY y pruebas de porosímetro de Boyle, Mercurio y Helio para la PHIE.

Para este estudio de análisis de componentes principales se contó con la información de registros de 6 pozos, de los cuales 4 cuentan con sus respectivas variables petrofísica VCLAY y PHIE previamente calculadas. Estos pozos se llamarán Pozos de Entrenamiento (PE) (Figura 7).

Los 2 pozos restantes tendrían los cálculos de los PE además se les realizó las pruebas de laboratorio para obtener la información de núcleos de roca, estos son el Pozo de Entrenamiento con Núcleo (PEN) y el Pozo Prueba con Núcleo (PPN).

Los pozos se encuentran perforados en un campo de petróleo colombiano, el cual presenta características geológicas muy homogéneas, no existen cambios fuertes tanto en la estratigrafía como en su geología estructural además de ser un campo de desarrollo somero. Esto nos garantiza que la lectura de los registros no va ser alterada por la complejidad que otros campos pueden presentar en la geología.

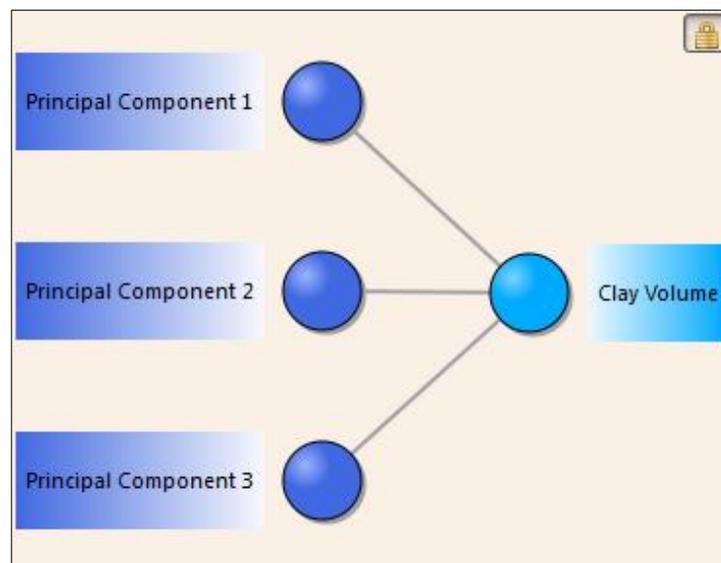
El análisis aquí propuesto pretende reducir el tiempo en la cuantificación de las dos propiedades petrofísicas VCLAY Y PHIE y así reducir considerablemente el tiempo a la hora de tomar la decisión en el completamiento del nuevo pozo PPN.



**Figura 7** Se encuentra la distribución en planta o espacial de los pozos PE-PEN y PPN para el campo.

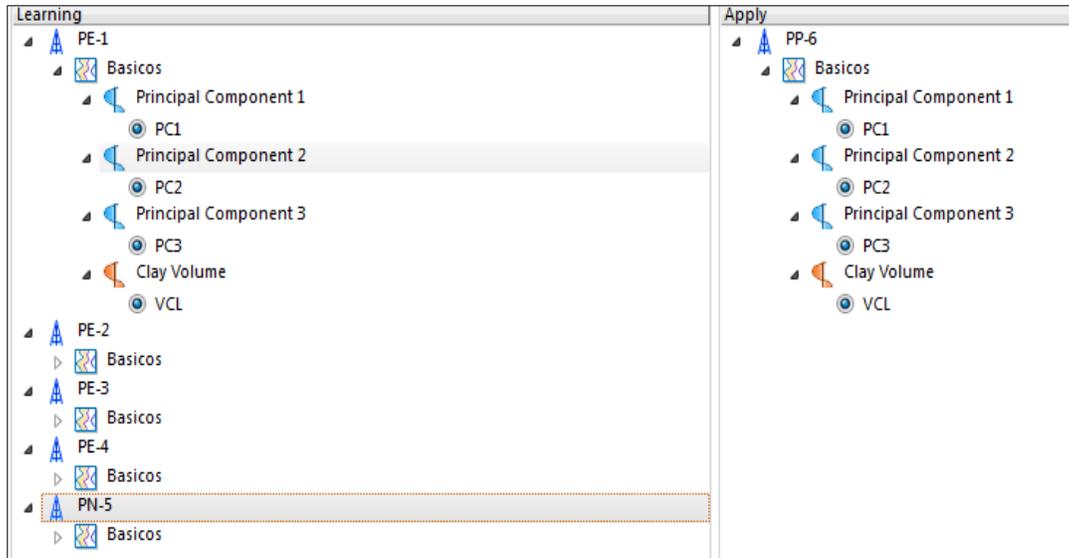
La idea es llevar las componentes principales extraídas (variables independientes), más las propiedades petrofísicas previamente calculadas de los pozos PE y PEN a un análisis de regresión lineal múltiple como lo vemos en la ecuación 7 para predecir VCLAY y PHIE de los pozos PPN (Figura 8).

$$y = m + ax_1 + bx_2 + cx_3 \quad \text{Ec.7}$$



**Figura 8** Estructura de la regresión múltiple.

En la **¡Error! No se encuentra el origen de la referencia.9**, muestra un gráfico de las curvas de componentes principales, las que son ahora las variables de entrada en el Software, usado por la industria para el cálculo de las propiedades petrofísicas VCLAY y PHIE.



**Figura 9** Estructura de aprendizaje y aplicación de la regresión lineal múltiple. En el aprendizaje se utilizan 1 pozo corazonado y en la aplicación lo mismo.

Ecuación 8. Volumen de Arcillosidad

$$VCLAY = + 0.00256247 * PC 1 + 0.009703042 * PC 2 + 0.006262425 * PC 3 - 0.3105766$$

Coefficiente de Correlación: 0.93

RMSE: 0.103

Ecuación 9. Porosidad Efectiva

$$PHIE = - 0.001603913 * PC 1 - 0.003035988 * PC 2 - 0.001460025 * PC 3 + 0.3184466$$

Coefficiente de Correlación: 0.92

RMSE: 0.04

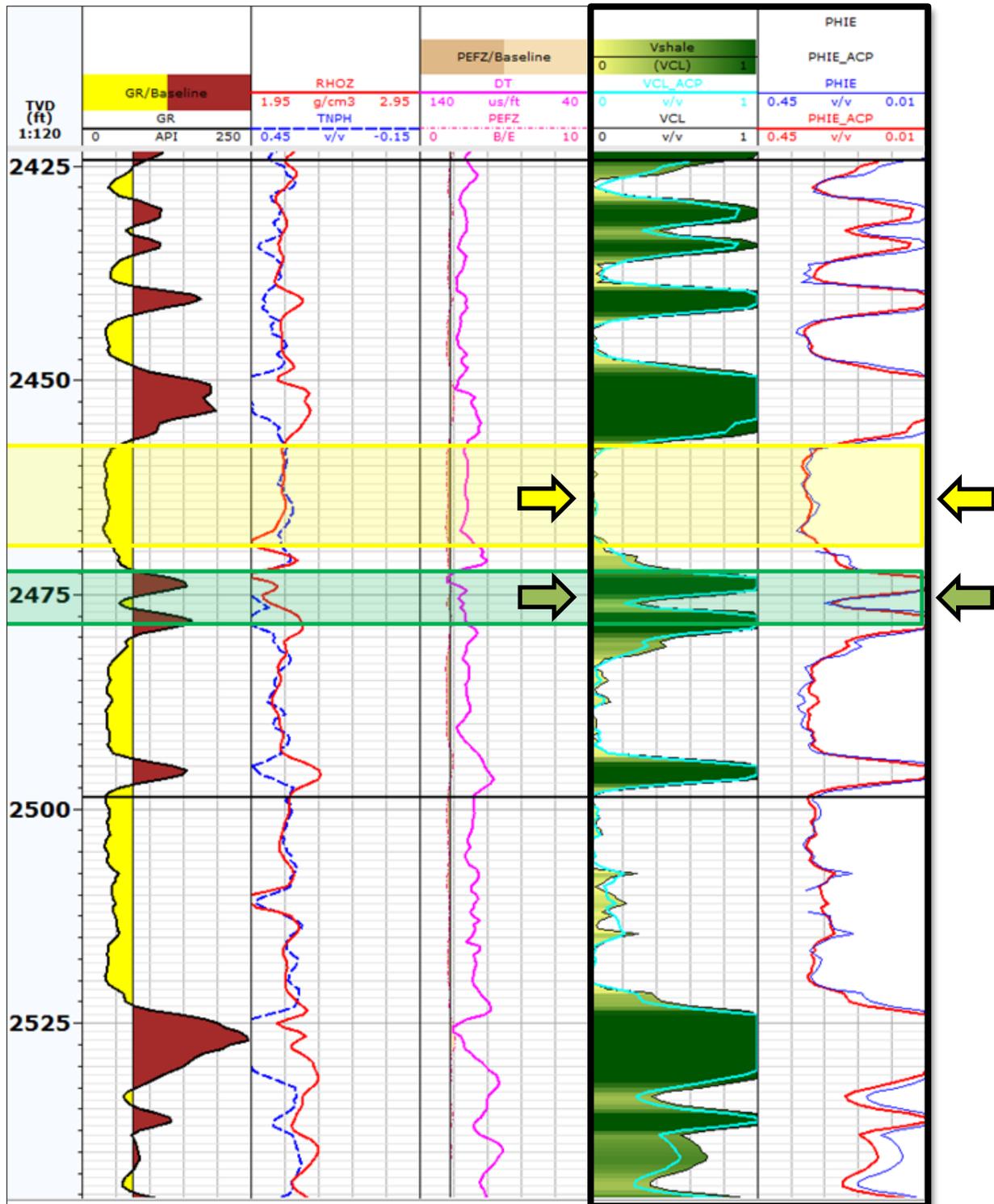
Las ecuaciones presentan excelentes coeficientes de correlación con tendencia a 1 y los RMSE bajos nos indican que las predicciones son ideales.

## **Capítulo 5. Resultados**

Las ecuaciones 8 y 9 obtenidas por los métodos de análisis de componentes principales y regresiones lineales múltiples se aplican al Pozo de Prueba con Núcleo (PPN), se comportaron de manera satisfactoria al caracterizar las principales litologías del yacimiento:

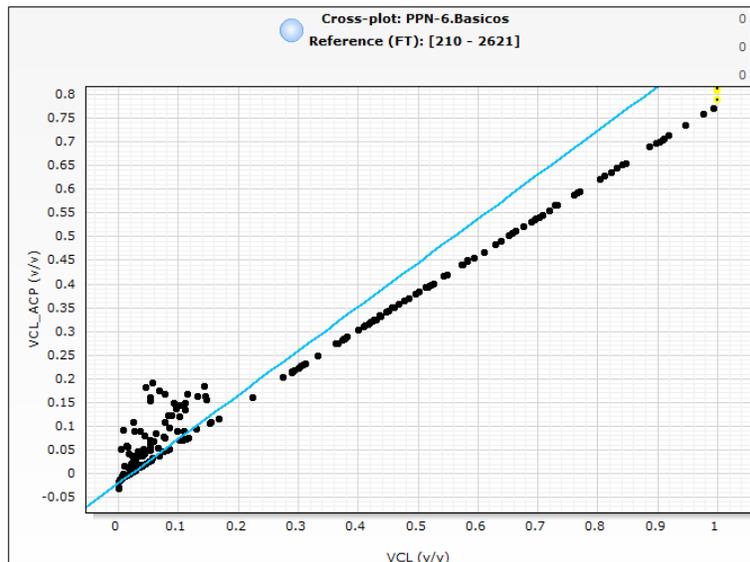
- **Rocas areniscas:** Son las rocas reservorios de fluidos en nuestro caso las rocas con mayor importancia porque pueden contener hidrocarburo. Se caracterizan a nivel de registros por bajas lecturas del GR, unión de las curvas RHOZ y TNPH, Tiempo de Tránsitos DT ligeramente más bajos y modelos petrofísicos con bajos VCLAY (<20%) y PHIE superiores al 30%, siendo concordante con las ecuaciones 8 y 9, donde identifican muy bien las mejores arenas con VCL\_ACP<10% graficadas de color agua marina y PHIE\_>30% de color rojo, en los intervalos de 2456-2476 Ft (Figura 10 flechas amarillas)
- **Rocas sello:** permiten que los fluidos atrapados en las rocas reservorios no fluyan a superficie y se pierdan. Se caracterizan a nivel de registros por altas lecturas del GR, separación de las curvas RHOZ y TNPH, Tiempo de Tránsitos DT más altos a las rocas reservorios y modelos petrofísicos con altos VCLAY (>70%) y PHIE menores al 5%. siendo concordante con las ecuaciones 8 y 9 donde identifican muy bien las mejores rocas sello con VCLAY>80% y PHIE>3% en los intervalos de 2470-2480 FT (Figura 9 flechas verdes)

Esto nos demuestra que las técnicas estadísticas aquí usadas tienen un buen comportamiento en la predicción de las propiedades petrofísicas y pueden ser una herramienta para calcularlas de una manera más rápida disminuyendo costos.

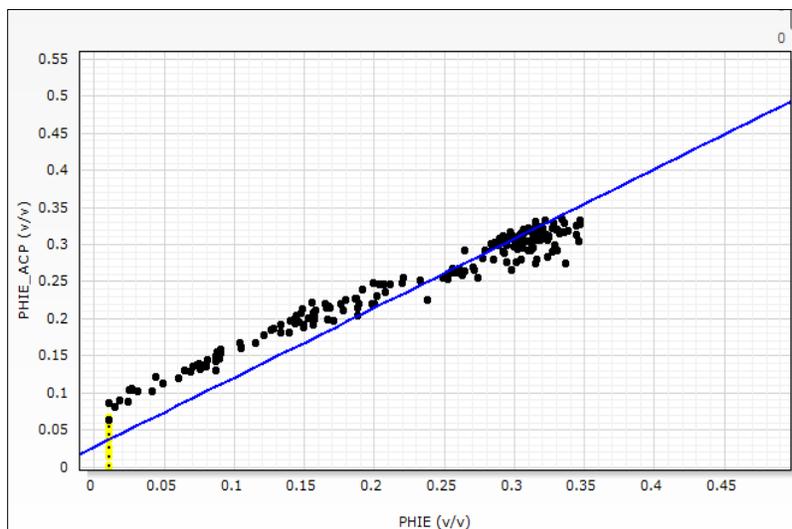


**Figura 10** Template de registros de pozo. De izquierda a derecha: 1. Profundidad, 2. GR, 3. RHOZ-TNPH, 4. PEFZ-DT, 5. Comparación de los modelos VCL 6. Comparación de los modelos PHIE

Se realizó el gráfico cruzado entre las variables del modelo petrofísico calculado y las variables obtenidas de la regresión lineal múltiple para VCLAY y PHIE, donde se puede observar nuevamente el buen grado de predicción de la metodología aquí propuesta (Figura 11-12).



**Figura 11** . Gráfico cruzado entre el VCLAY realizado por modelos petrofísicos y el VCLAY obtenido con la regresión lineal múltiple.



**Figura 12** Gráfico cruzado entre PHIE realizado por modelos petrofísicos y la porosidad efectiva obtenida con la regresión lineal múltiple.

## **Capítulo 6. Conclusiones**

Este estudio llevado a cabo en conjuntos de datos de registros de pozos de un campo de petróleo colombiano sugiere que la combinación de las técnicas estadísticas de Análisis de Componentes Principales (ACP) y Regresión lineal Múltiple puede complementar con éxito los métodos convencionales de evaluación de la formación.

Las ecuaciones 8 y 9 obtenidas de los modelos de VCLAY y PHIE, separan unidades litoestratigráficas como las rocas reservorio y rocas sellos. Estas mismas al ser comparadas con los modelos previamente calculados muestran coeficiente de correlación cercano a 1.

Las pruebas Bartlett, KMO, gráficas de sedimentación y el porcentaje de explicación de la varianza, indican que los componentes extraídos son tres: El primer componente principal responde a la PHIE, asociando registros RHOZ\_TNPH y DT. Por otro lado, la segunda componente explica cambios importantes de litología y variaciones de VCLAY que se ven reflejados por la predominancia del GR. La tercera componente principal refleja la variabilidad de los datos relacionados con el fluido.

Estas técnicas estadísticas al caracterizar yacimientos petrolíferos, permiten tomar decisiones de manera más rápida en la etapa de completamiento disminuyendo costos.

## **Capítulo 7. Bibliografía**

- Archie, G.E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics, Trans. AIME
- Babadagli, T., AlSalmi, S. (2002). Improvement of Permeability Prediction for Carbonate reservoirs Using Well Log Data”, SPE-77889 Paper.
- Fertl, W. H. (1986). Clay Typing Based On Well Logs. Petroleum Society of Canada. doi:10.2118/86-37-18
- Gómez, O. (1981): A Conciliating Porosity Exponent Relationship: Its Application To Practical Well Log, Society of Professional Well Log Analysts.
- Hook, J. R. (2003, May 1). An Introduction to Porosity. Society of Petrophysicists and Well-Log Analysts.
- Kassenaar, J.D.C. (1989) Automated Classification of Geophysical Well Logs; M.Sc. Thesis, University of Waterloo, Waterloo, Ontario, Canada.
- Moake, G. L., & Schultz, W. E. (1989). Improved Density Log Lithology Identification Using a Borehole-Compensated Photoelectric Factor. Society of Petroleum Engineers. doi:10.2118/17961-MS
- Moss, B., & Seheult, A. (1987). Does Principal Components Analysis Have A Role In The Interpretation Of Petrophysical Data? Society of Petrophysicists and Well-Log Analysts.
- Schlumberger, (2000). Effective Porosity and Permeability of Sandstones Investigated Through Nuclear Magnetic Resonance Principles.
- Spec2000.net. (2017). Crain's Petrophysical Handbook - Lithology -- Principal Component Analysis. [online] Available at: <https://www.spec2000.net/13-lithpca.htm> [Accessed 31 Jul. 2017].
- Teh, W., Willhite, G. P., & Doveton, J. H. (2012). Improved Reservoir Characterization in the Ogallah Field using Petrophysical Classifiers within Electrofacies. Society of Petroleum Engineers. doi:10.2118/154341-MS
- Tittman, J, Wahl J.S. (2005) The physical foundations of formation density logging (gamma-gamma). American Geological Institute. DOI: 10.1190/1.1439574