

ESTIMACION DE MATERIA ORGÁNICA TOTAL A PARTIR DE PERFILES DE POZO: COMPARACION DE METODOS DETERMINISTICOS, ESTADISTICOS Y NO LINEALES (REDES NEURONALES)

David Curia¹, Claudio Larriestra¹, Claudio Haring²

1 Larriestra, Curia y Asociados

2 REPSOL-YPF

Resumen

Los perfiles de pozo brindan distintos grados de estimación del contenido de materia orgánica, dependiendo de la tecnología de los mismos y las propiedades petrofísicas de la roca madre entre otros factores. Se compara el contenido de materia orgánica obtenido en recortes de perforación con la respuesta de perfiles de un pozo en la Cuenca de Malvinas. Dicha comparación se efectuó mediante un método determinístico (fórmula matemática que involucra a los perfiles sónico y resistividad), un método estadístico (Análisis Discriminante) y un método de estimación no lineal (Redes Neuronales). Con los dos primeros se alcanza un ajuste regular y limitado. La estimación tiene un carácter cualitativo e indicador (hay o no hay materia orgánica). Con la red neuronal en cambio, y a partir de la convergencia entre los datos continuos (perfiles) y los discretos (contenido de materia orgánica), se consiguió un ajuste mejor que en los restantes métodos y de carácter cuantitativo. La conversión indirecta (usando información que infiere propiedades: los registros eléctricos) de una serie discreta (análisis de TOC) a una serie continua (curva de TOC), permite realizar estimaciones en zonas no analizadas del pozo y a partir de condiciones geológico-ambientales relativamente constantes, se pueden realizar estimaciones en pozos donde no se hayan efectuado ese tipo de análisis.

Introducción

De acuerdo a la literatura especializada el problema de la estimación de TOC se puede resolver con diferentes aproximaciones metodológicas, basadas en principios diferentes, a saber:

1. *Métodos Determinísticos*
2. *Métodos Probabilísticos*

En el presente artículo se pretende introducir otra alternativa para este problema de estimación, que es la aplicación de *Redes Neuronales* para calcular TOC (carbono orgánico total) a partir de los registros eléctricos.

Descripción de las metodologías

1.-Métodos Determinísticos

El más evolucionado de los métodos de este grupo es el propuesto por Passey et al.(1990) y que consiste en la combinación de los registros de Resistividad y Tiempo de Tránsito con el control de curvas como Rayo Gama o Autopotencial, a partir de las cuales es posible cuantificar el contenido total de materia orgánica.

El método de basa en el grado de separación relativo entre las curvas de resistividad y porosidad, a partir de la siguiente fórmula:

$$(1) \text{ Dlog(R)} = \log(R/R_b) + K(\Phi - \Phi_b)$$

Donde: Dlog(R) = separación entre la resistividad y porosidad

R =	resistividad
Rb =	resistividad de línea de base de referencia
K =	factor de escala
Phi =	porosidad
Phib =	porosidad de línea de base de referencia

Cuantitativamente el valor de TOC se calcula con la ecuación:

$$(2) \text{ TOC} = D \log(R) 10^{(2.297 - 0.1688 \times \text{LOM})}$$

Donde: LOM = grado de maduración

Las constantes de la fórmula precedente fueron estimados por Passey et al (op. cit.) para pozos de diversas cuencas. Es recomendable investigar estadísticamente el ajuste de estos parámetros para las zonas donde se pretende aplicar este método.

Tomando Rb y Phib como los valores leídos en arcillas no generadoras, para obtener un *indicador cualitativo* se define:

$$(3) Y_r = \log(R/R_b)$$

$$(4) Y_p = K (\Phi - \Phi_{ib})$$

resulta para zonas sin materia orgánica:

$$(5) Y_r = Y_p = 0$$

Luego calculando para cada muestra los valores Y_r e Y_p, resulta que donde Y_r > Y_p es un indicador de:

- *Materia orgánica en zonas no permeables*
- *Hidrocarburos en reservorios.*

Además, en zonas saturadas con agua R = R_o y de la ecuación (3) se tiene

$$(6) R_o = R_b 10^{Y_r}$$

y, nuevamente donde R > R_o actúa como identificador de materia orgánica o hidrocarburos en zonas de reservorios.

Por último aplicando la ecuación de Archie (1942)

$$(7) S_w = (R_o/R_t)^{(1/n)} \text{ y sustituyendo con (6)}$$

$$(8) S_w = (R_b 10^{Y_r} / R_t)^{(1/n)}$$

Se obtiene un método de cálculo de la saturación sin recurrir al factor de formación (ni a los parámetros de cementación -m- y tortuosidad -a-). Este último resultado representa un aporte adicional de esta metodología.

2.-Métodos probabilísticos

Dentro de la multiplicidad de métodos estadísticos, en la literatura se han aplicado con resultados satisfactorios dos grandes grupos procedimientos estocásticos para la estimación de materia orgánica a partir de perfiles de pozos:

- *Regresión múltiple*
- *Análisis discriminante*

El primero de ellos consiste en "calibrar" los valores de las curvas eléctricas sobre la base de mediciones directas (de propiedades como TOC, S2, TMAX, PIROLISIS, VITRINITA, etc). Básicamente se aplica el método de mínimos cuadrados para calcular los parámetros (a, b, c, ...) de una ecuación del tipo:

$$\text{TOC} = a * \text{TIEMPO DE TRANSITO} + b * \text{RESISTIVIDAD} + c * \text{RAYO GAMA} + \dots$$

y se aplica donde no se dispone de mediciones directas de las propiedades geoquímicas. La elección de los registros a utilizar depende de la correlación que se pueda establecer en aquellos puntos donde se tienen las mediciones directas (Davis, 1986).

Para aplicar el análisis discriminante, se parte de una calibración sobre valores medidos en zonas donde se conoce *a priori*, en forma directa, la "presencia" o "ausencia" de materia orgánica (e incluso su "calidad generadora") para formular (luego de estimar los parámetros) reglas de decisión del tipo:

$$\text{Si: } a * \text{TIEMPO DE TRANSITO} + b * \text{RESISTIVIDAD} + c * \text{RAYO GAMA} > 0$$

ENTONCES la muestra pertenece al conjunto de generación de hidrocarburos.

Al mismo tiempo es posible calcular, para cada punto de los registros eléctricos la probabilidad de pertenecer al grupo mencionado (Davis, op. cit.)

3.-Redes Neuronales (RN)

Una RN es una colección de procesos simples pero altamente relacionados o interconectados. Estos procesos simples o elementales (también llamados nodos) calculan sumas pesadas de las entradas, agregan un sesgo y aplican una función de transferencia antes de obtener una salida. En términos matemáticos podemos escribir:

$$S_i = \sum_j W_{ij} I_j + q_i$$

$$O_i = F(S_i)$$

Donde: S_i =Entrada a Nodo i

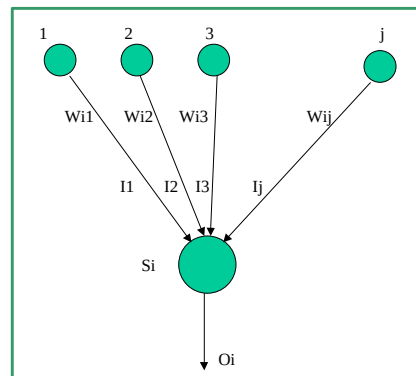
W_{ij} =Peso desde el nodo j al nodo i

I_j = Señal que proviene del nodo j

F =Función de Transferencia

O_i = Salida del nodo i

q_i = sesgo



Los nodos son análogos a las neuronas del cerebro humano (son solamente un modelo simplificado), de aquí el nombre de red neuronal al presente esquema de cálculo.

Las conexiones entre los nodos, tienen asociadas pesos que se modifican en un procedimiento iterativo con el objetivo de minimizar el error de la salida. Este mecanismo es conocido como aprendizaje o entrenamiento de la RN. Estos procedimientos de entrenamiento pueden ser supervisados (donde se usa una salida correcta o medida en forma independiente) o no-supervisados (donde la red determina la salida en función de la estructura interna de los datos). Dentro de cada una de estas grandes categorías existen varios algoritmos diferentes.

Un nodo en particular carece de significado. El comportamiento colectivo de los nodos y sus conexiones (es decir la RN) determinan el resultado del procedimiento.

En general la operación de una RN se divide en dos partes: entrenamiento y procesamiento. En el entrenamiento se utiliza un subconjunto de los datos, una vez que el comportamiento de la salida es satisfactorio se pasa a la fase de procesamiento del total de los datos. Si el subconjunto de datos usados para el entrenamiento no es representativo del total los resultados serán erróneos.

Existen muchos procedimientos para aplicar una RN, en este artículo se ha elegido uno de los mas difundidos: **Multi – Layer Perceptron (MLP)**

El MLP es un algoritmo de RN supervisado que aprende comparando la salida de la RN con la real. En este algoritmo los nodos se agrupan en capas y se conectan con los de la capa anterior y la posterior. No se determinan conexiones dentro de una capa. La primera capa recibe como entrada el datos reales (usualmente cada nodo corresponde a un registro disponible), mientras que la salida de la última capa es el resultado del procedimiento (en general hay un nodo para estimar TOC).

Finalmente se dispone de una o mas capas intermedias u ocultas. Para cada nodo la función de transferencia en una logística o sigmoidea (en general puede ser cualquier función clase c2, es decir que ella y la primera derivada deben ser continuas y biyectivas). El algoritmo de entrenamiento comienza definiendo los pesos iniciales al azar, se calcula la salida de cada nodo de la red hasta la última capa; en este punto se compara la salida con una muestra testigo y se calcula el error (aquí se debe definir un criterio de convergencia, en función del error), se modifican los pesos (en general por gradientes conjugados) para disminuir el error y se vuelve a calcular sobre cada nodo de la red.

Aplicación de los tres métodos al mismo problema

Para efectuar la comparación entre los diferentes métodos, se tomaron los datos de un pozo para calibrar el modelo, utilizándolo luego para la estimación de la propiedad problema en otro pozo. Posteriormente y a los efectos de estudiar el error se compara el resultado de la estimación con datos medidos no empleados en la generación del modelo.

1.-Método PASSEY et. al. (1990)

La figura 1 muestra la aplicación de la fórmula de Passey et. al. (op. cit.), para calcular TOC y su comparación con datos medidos en laboratorio, en un pozo de la Cuenca Malvinas. En la misma se aprecia un ajuste básicamente cualitativo y resulta evidente que el método para este caso adolece de poder de predicción. Este resultado puede deberse a que la respuesta de los perfiles utilizados, en especial el sónico, que puede estar afectado por condiciones de compactación diferentes a las existentes para los casos utilizados en la construcción del modelo de Passey et. al. (op. cit.).

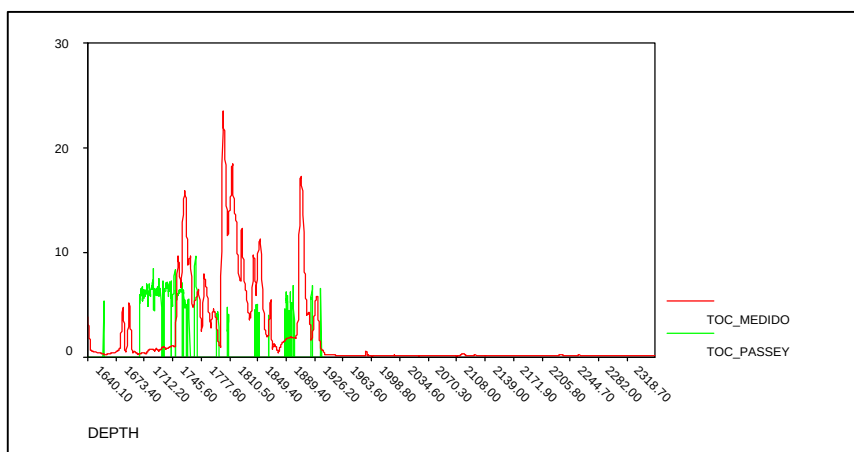


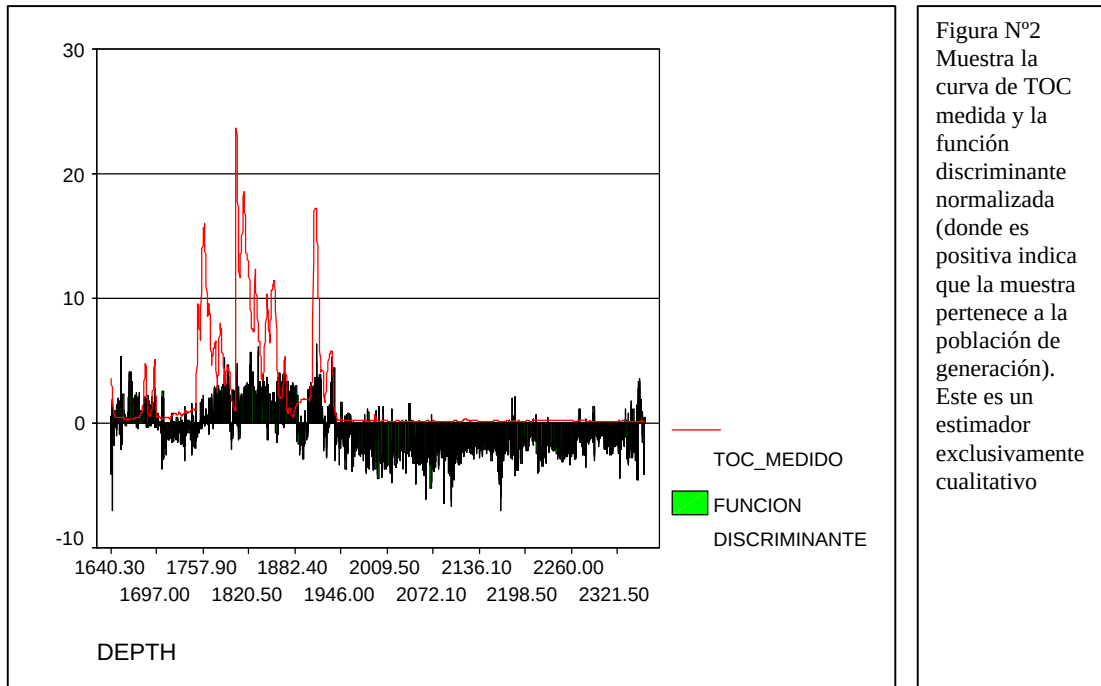
Figura 1:
comparación del
modelo de Passey
et. al. (1990) y la
curva de TOC
medida. La curva
estimada tiene ca-
rácter de indi-
cador cualitativo.
Las medidas de
correlación cuan-
titativas son muy
bajas.

2.-Método Análisis Discriminante

La ecuación discriminante utilizada por Meyer y Nederlof (1984) emplea la resistividad normalizada por temperatura y el sónico compensado, los que integran la fórmula:

$D = a * Rt_{n75} + b * Dt + c$ donde a, b, c son los coeficientes calculados por el análisis discriminante y D es el número discriminante. Si $D > 0$ entonces aumenta la probabilidad de estar en zonas generadoras.

Este trabajo fue realizado sobre un conjunto de pozos de diferentes cuencas, pero que no se puede considerar lo suficientemente representativo como para aceptar la fórmula mencionada como cierta en todos los casos.



En este caso la fórmula mencionada más arriba no brindó resultados satisfactorios, debiéndose calcular una nueva ecuación discriminante incluyendo la curva del Rayos Gamma y cuya forma genérica quedó definida como:

$D = a * Rt_{n75} + b * Dt + c * GR + d$ donde a, b, c, d son los coeficientes calculados por el análisis discriminante y D es el número discriminante. Como en el caso anterior, si $D > 0$ entonces aumenta la probabilidad de estar en zonas generadoras.

En la figura 2 se ilustra la comparación de la curva original de TOC y la función discriminante normalizada, donde se puede apreciar el carácter cualitativo de la estimación y solo muestra probabilidad de encontrar espesores generadores.

1.- Redes Neuronales

Para aplicar esta metodología se realizó un preproceso de promedios móviles de 25 puntos sobre los registros eléctricos para evitar el sobre entrenamiento de la red. En la figura 3 se muestra el

histograma con la medida del error de estimación en todo el rango de profundidad del pozo, donde se observa que la media del error es cero (en unidades de TOC medido), con lo cual no se debe esperar que tienda a sobre o subestimar.

Además y a manera de validación del método puede apreciarse en la figura 4 la correlación entre la curva estimada por la real y la curva de TOC medida. Notese el alto valor de correlación lineal.

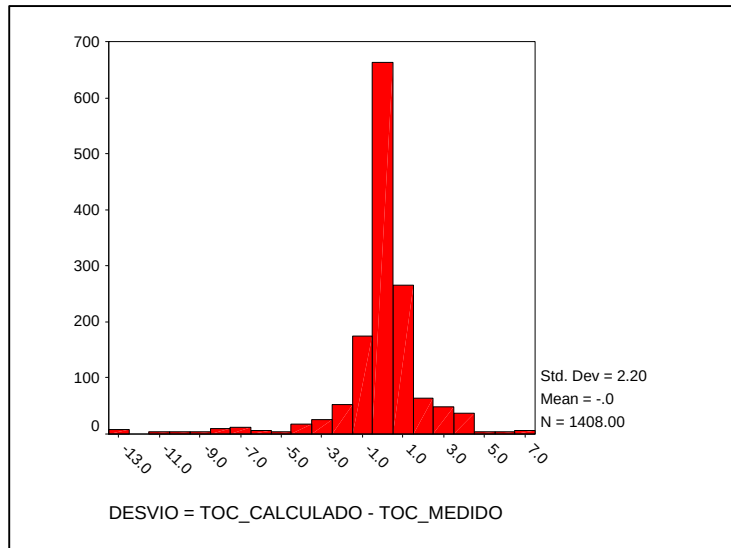


Figura 3: error de estimación medido a lo largo del pozo

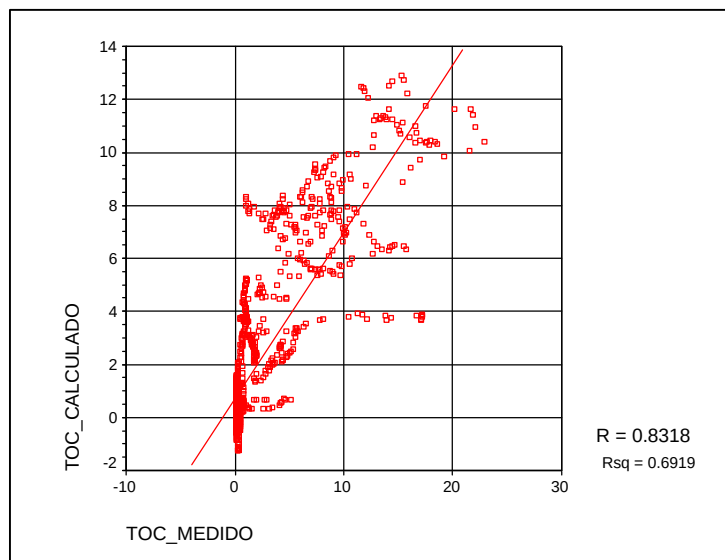


Figura 4: correlación entre la curva estimada y la curva de TOC medida. A la derecha coef. correlación

Por último en la figura 5 se ilustra el resultado del procesamiento de la red, en donde existe una mejor aproximación a los datos medidos de TOC, con respecto a los métodos precedentes. Además, es de notar que los cambios bruscos en el TOC medido han sido suavizados en el estimado. Por último cabe agregar que, a diferencia de los casos anteriores, la estimación tiene un carácter cuantitativo.

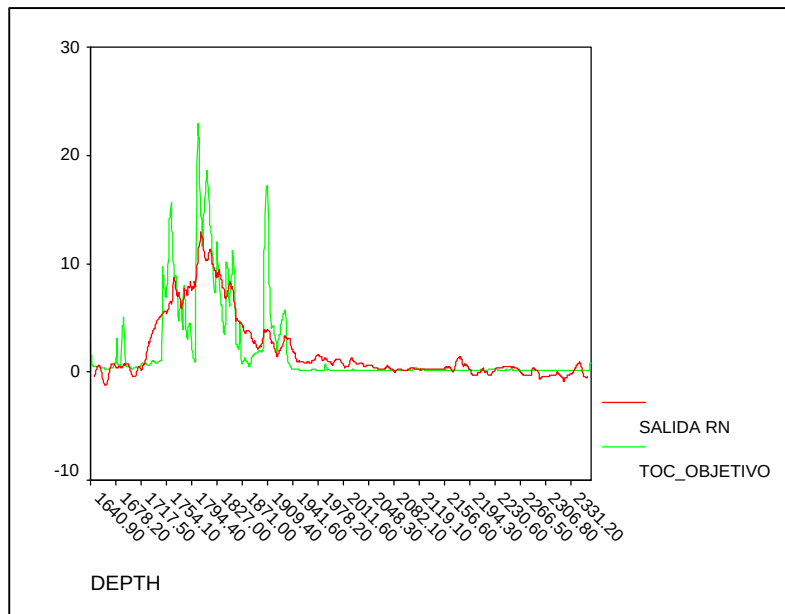


Figura N°3 Muestra la curva de TOC medida y la obtenida como objetivo de la red neuronal. Esta estimación tiene carácter cuantitativo

Conclusiones

La aplicación de los métodos sobre un mismo ejemplo muestran que con la red neuronal se logra un ajuste cuantitativo razonable. Es necesario destacar que la aplicación de un procedimiento de red neuronal en un problema de estimación de este tipo requiere un tiempo de preproceso y de modelado de la red a utilizar. El método propuesto por Passey, et. al. (1990) no es lo suficientemente robusto como para ser aplicado en cualquier situación y es aplicable a casos similares con los que se puso a prueba el mismo. El método estadístico (análisis discriminante) es robusto pero su resultado está en función de la calidad de datos tomados para las poblaciones a comparar, (roca generadora y roca no generadora) siendo su estimación de neto carácter cualitativo. Por último si bien la aplicación de la tecnología de redes neuronales producen estimaciones cuantitativas, su aplicación requiere un estudio particularizado de cada situación.

Bibliografía citada en el texto

- Archie, G., 1942 , "The electrical resistivity log as an aid in determinig some resrvoir characteristics", Trans. of AIME, 146, nro. **1**, p.54-62
- Davis, J., 1986 "Statistics and Data Analysis in Geology", John Wiley & Sons, N.Y., 646 p.
- Doveton, J.H.: Geologic Log Analysis Using Computer Methods. AAPG Computer Applications in Geology, **2**, 1994
- Meyer, B. Y Nederlof, M., 1984 "Identification of source rock on wireline logs by Density-Resistivity and Sonic Transit Time – Resisitivity crossplots", AAPG Bulletin, **68/2**, p.121-129
- Passey, Q.R., Creaney, S., Kulla, J.B., Moretti, F.J., Stroud, J.D. ,1990 "A Practical Model for Organic Richness from Porosity and resistivity Logs" AAPG Bulletin, 74/5, p.1210-1214.