

**IMPLEMENTACION DE METODOS GEOESTADISTICOS PARA EL
MODELAMIENTO DE PROPIEDADES PETROFISICAS EN YACIMIENTOS DE
DEPOSITACION FLUVIAL EN COLOMBIA: CAMPO CASTILLA**

**SILVIO JOSE DIAZGRANADOS DIAZ
FELIPE ANDRES NIETO FERNÁNDEZ**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERIAS FISICOQUÍMICAS
ESCUELA DE INGENIERIA DE PETROLEOS
BUCARAMANGA**

2005

IMPLEMENTACION DE METODOS GEOESTADISTICOS PARA EL
MODELAMIENTO DE PROPIEDADES PETROFISICAS EN YACIMIENTOS DE
DEPOSITACION FLUVIAL EN COLOMBIA: CAMPO CASTILLA

SILVIO JOSE DIAZGRANADOS DIAZ
FELIPE ANDRES NIETO FERNÁNDEZ

**Trabajo de Grado como requisito parcial para optar al título de
Ingeniero de Petróleos**

Director

NICOLAS SANTOS SANTOS

Ingeniero de Petróleos, MsC.

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERIAS FISICOQUÍMICAS
ESCUELA DE INGENIERIA DE PETROLEOS
BUCARAMANGA**

2005

A nuestros padres gracias por sus grandes enseñanzas que hoy en día toman más vigencia que nunca. A Dios por toda una vida llena de oportunidades. Finalmente, a todas las personas que de una u otra colaboraron en el desarrollo de este trabajo.

AGRADECIMIENTOS

Los autores expresan sus agradecimientos a:

A Dios por darnos vida, salud y ganas de seguir adelante sin importar los tropiezos.

A Edwin Pineda Carreño por su supervisión y guía en el desarrollo de este proyecto; además por la elaboración de uno de los programas utilizados en este, el cuál ayudo a que este trabajo llegará a buen término.

Al MsC. Nicolás Santos por habernos brindado esta magnífica oportunidad.

Al Ingeniero Alejandro Ospino por su apoyo en el manejo de los programas auxiliares necesarios para correr la aplicación.

Al Ingeniero Edgar Eduardo Martinez por su apoyo en la metodología geoestadística.

Al MsC. Eduardo A. Jiménez por su colaboración en la metodología geoestadística.

Al geólogo Edwar Herrera Otero quien siempre enriqueció las discusiones con ideas y sugerencias muy oportunas.

Al ingeniero William Navarro por su orientación y por la información brindada.

A la Universidad Industrial de Santander UIS por el conocimiento brindado durante estos cinco años.

Al Grupo de Modelamiento de Procesos de Hidrocarburos GMPH, que nos brindo todas las herramientas necesarias para la realización de este trabajo.

A nuestras familias y amigos por el apoyo incondicional y por la paciencia durante todo este tiempo.

TABLA DE CONTENIDO

	Pág
INTRODUCCION	2
1. ESTADO DEL ARTE.....	4
1.1 RESEÑA HISTÓRICA.....	4
1.2 CONCEPTOS DE GEOESTADÍSTICA.....	6
1.3 VARIABLES REGIONALIZADAS.....	7
1.4 VARIOGRAMAS	11
1.5 MODELAMIENTO DE VARIOGRAMAS	14
1.6 ANISOTROPIA	20
1.7 TÉCNICA DE INTERPOLACIÓN GEOESTADÍSTICA DE KRIGING.....	22
1.8 GEOESTADÍSTICA MULTIVARIADA (CO- KRIGING).....	34
2. PROGRAMAS DE MODELAMIENTO GEOESTADÍSTICOS – LIBRERÍA GSLIB	38
2.1 PRINCIPIOS DE LA SIMULACIÓN ESTOCÁSTICA	38
2.2 ALGORITMO DE TIPO GAUSSIANO	50
2.3 ALGORITMOS DE TIPO INDICADOR.....	59
2.4 PROGRAMAS DE SIMULACIÓN GAUSSIANA.....	66
2.5 PROGRAMAS DE SIMULACIÓN DE INDICADOR SECUENCIAL.....	72
2.6 PROGRAMAS ÚTILES	77

3. CARACTERÍSTICAS GEOLOGICAS DE LOS AMBIENTES FLUVIALES	80
3.1. MOVIMIENTO DE PARTICULAS	80
3.2. ESTRUCTURAS SEDIMENTARIAS	84
3.3. MODELOS DE FACIES FLUVIALES	85
3.4. TIPOS DE DEPÓSITO	93
3.5. CONTROLES DE CANAL Y MORFOLOGÍA	100
3.6. ARQUITECTURA DE LOS ELEMENTOS DE CANALES	102
4. MODELAMIENTO DEL YACIMIENTO	103
4.1. DESCRIPCIÓN DEL CAMPO	103
4.1.1. Unidad Guadalupe Masivo	109
4.2. MARCO ESTRUCTURAL DEL CAMPO	112
4.3. MODELO PETROFÍSICO	113
4.4. MODELAMIENTO GEOESTADÍSTICO	123
4.5. ANALISIS GEOESTADISTICO	135
5. RESULTADOS	148
5.1 PERMEABILIDAD	148
5.2 POROSIDAD	157
6. CONCLUSIONES	166
7. RECOMENDACIONES	168
8. BIBLIOGRAFIA	169

LISTA DE TABLAS

	Pág
Tabla 4.1 Inventario de Registros de Pozo. Campo Castilla.....	114
Tabla 4.2. Análisis básicos y especiales Campo Castilla.....	115
Tabla 4.3. Intervalos corazonados para los pozos de Castilla.	124
Tabla 4.4. Topes de la unidades en los pozos trabajados para el campo Castilla.	125
Tabla 4.5. Valores de los registros eléctricos y los algoritmos generados.	126
Tabla 4.6. Interpretación Variogramas Permeabilidad	143
Tabla 4-7. Interpretación Variogramas Porosidad.....	144

LISTA DE FIGURAS

	Pág
Figura 1.1. Elementos del Variograma.....	13
Figura 1.2. Modelo Lineal	15
Figura 1.3. Modelo Esférico	16
Figura 1.4. Modelo Exponencial.....	17
Figura 1.5. Modelo de Potencia	18
Figura 1.6. Modelo Gaussiano	18
Figura 1.7. Modelo Efecto Hueco.....	19
Figura 1.9. Anisotropía Zonal.....	22
Figura 1.10. Grupo de muestras hipotético en un proceso de estimación	24
Figura 1.11. Estimación del valor de una variable en una celda de una malla	26
Figura 2.1 Ejemplo de archivo de parámetro para <i>sgsim</i>	67
Figura 2.2 Ejemplo de archivo de parámetro para <i>gtsim</i>	71
Figura 2.3 Ejemplo de archivo de parámetro para <i>sisim</i>	73
Figura 2.4 Ejemplo de archivo de parámetros para <i>pixelplt</i>	78
Figura 2.5 Ejemplo de archivo de parámetros para <i>vargplt</i>	79
Figura 3.1. Flujo de partículas en una corriente.....	81
Figura 3.2. Flujo laminar y turbulento.....	82
Figura 3.3. Flujo de corriente sobre un lecho de arena, limo y arcilla	83
Figura 3.4. Ripples de corriente.....	85
Figura 3.5. Canal trenzado	87
Figura 3.6. Tipo de depósito asociado a un canal meandriforme.....	88
Figura 3.7. Esquema de planta y cortes asociados a la trayectoria de un canal ...	89
Figura 3.8. Diferencia entre la línea central del canal y la línea del thalweg.....	90

Figura 3.9. Ejemplo de un canal meandriforme caracterizado por los depósitos de Point bar y Oxbow lake	91
Figura 3.10. Ejemplo de un meandro que corta verticalmente una sección estratigráfica.	91
Figura 3.11. Formación de un Oxbow lake en un canal meandriforme.....	92
Figura 3.12. Patrones de canales y perfiles característicos de acuerdo a la pendiente del río	93
Figura 3.13. Perfil característico de un canal mostrando el canal mismo, el natural levee y la superficie de inundación	94
Figura 3.14. Formación de un point bar. Vista de planta y perfil.....	95
Figura 3.15. Canal meandriforme mostrando la formación y migración de un point bar.....	96
Figura 3.16: (A). Vista de planta . (B) vista de perfil.....	96
Figura 3.17. Etapas en la formación de un Natural Levee	97
Figura 3.18. Ejemplo de un Natural Levee.....	98
Figura 3.19. Ejemplo de una llanura de inundación en un canal meandriforme	99
Figura 3.20. Formación de un Oxbow lake en un canal meandriforme.....	100
Figura 3.21. Tipos de depósitos aportados en un canal de acuerdo a la posición de la trayectoria	101
Figura 4.1. Mapa Localización del campo.....	103
Figura 4.2. Columna estratigráfica	106
Figura 4.3 Secuencia depositacional del Campo Castilla	108
Figura. 4.4 Ambiente de depositación.....	110
Figura 4.5 Sección Delgada Unidad Guadalupe Masivo.....	112
Figura 4.6 Mapa Estructural De La Unidad Guadalupe Masivo	113
Figura 4.7. Gráfico de RHOB vs PHICORE. Castilla 12. Unidad K2.....	117
Figura 4.8. Relación entre porosidad calculada y porosidad medida. Castilla-12 Unidad K2.	118
Figura 4.9 Gráfica para identificar tipos de roca para Castilla-12, Unidad K2.....	120

Figura 4.10. Comportamiento de la curva GR con la permeabilidad de corazones. Castilla-12-K2	121
Figura 4.11. Comportamiento de la gráfica de porosidad vs permeabilidad de corazones. Castilla-12-K2.....	123
Figura 4.12 Registros eléctricos y a la derecha la curva generada para las electrofacies en profundidad.	127
Figura 4.13 Sección horizontal unidad Guadalupe Masivo	128
Figura 4.14 Corte transversal unidad Guadalupe Masivo	129
Figura 4.15. Grid de Referencia.....	131
Figura 4.16. Definición del Grid Paralelo	133
Figura 4.17. Definición del Grid Proporcional	133
Figura 4.18. Convención de Orientación del Grid	134
Figura 4.19. Software GAMV	136
Figura 4.20. Variograma Horizontal De Permeabilidad En Dirección 0°	137
Figura 4.21. Variograma Horizontal De Permeabilidad En Dirección 45°	137
Figura 4.22. Variograma Horizontal De Permeabilidad En Dirección 90°	138
Figura 4.23. Variograma Horizontal De Permeabilidad En Dirección 135°	138
Figura 4.24. Variograma Vertical De Permeabilidad	139
Figura 4.25 Variogramas De Permeabilidad	139
Figura 4.26. Variograma Horizontal De Porosidad En Dirección 0°	140
Figura 4.27. Variograma Horizontal De Porosidad En Dirección 45°	140
Figura 4.28. Variograma Horizontal De Porosidad En Dirección 90°	141
Figura 4.29. Variograma Horizontal De Porosidad En Dirección 135°	141
Figura 4.30. Variograma Vertical De Porosidad.....	142
Figura 4.31 Variogramas De Porosidad	142
Figura 4.33 Elipse De Anisotropía	145
Figura 4.34. Software SGSIM	146
Figura 5.1 Modelo resultante de Permeabilidad.....	148
Figura 5.2 Vista lateral de Permeabilidad. Layer 1	149

Figura 5.3 Vista lateral de Permeabilidad. Layer 10	150
Figura 5.4 Vista lateral de Permeabilidad. Layer 30	150
Figura 5.5 Vista lateral de Permeabilidad. Layer 45	151
Figura 5.6 Vista lateral de Permeabilidad. Layer 60	151
Figura 5.7 Vista frontal de Permeabilidad. Layer 10	152
Figura 5.8 Vista frontal de Permeabilidad. Layer 30	152
Figura 5.9 Vista frontal de Permeabilidad. Layer 60	153
Figura 5.10 Vista frontal de Permeabilidad. Layer 100	153
Figura 5.11 Vista frontal de Permeabilidad. Layer 140	154
Figura 5.12 Vista en Superficie de Permeabilidad. Layer 1	154
Figura 5.13 Vista en Superficie de Permeabilidad. Layer 15	155
Figura 5.14 Vista en Superficie de Permeabilidad. Layer 50	155
Figura 5.15 Vista en Superficie de Permeabilidad. Layer 120	156
Figura 5.16 Vista en Superficie de Permeabilidad. Layer 161	156
Figura 5.17 Modelo Resultante de Porosidad	157
Figura 5.18 Vista lateral de Porosidad. Layer 1	158
Figura 5.19 Vista lateral de Porosidad. Layer 20	159
Figura 5.20 Vista lateral de Porosidad. Layer 40	159
Figura 5.22 Vista frontal de Porosidad. Layer 1	160
Figura 5.23 Vista frontal de Porosidad. Layer 30	161
Figura 5.24 Vista frontal de Porosidad. Layer 60	161
Figura 5.25 Vista frontal de Porosidad. Layer 90	162
Figura 5.26 Vista frontal de Porosidad. Layer 130	162
Figura 5.27 Vista frontal de Porosidad. Layer 160	163
Figura 5.28 Vista en Superficie de Porosidad. Layer 1	163
Figura 5.29 Vista en Superficie de Porosidad. Layer 30	164
Figura 5.30 Vista en Superficie de Porosidad. Layer 60	164

RESUMEN

TITULO: IMPLEMENTACION DE METODOS GEOESTADISTICOS PARA EL MODELAMIENTO DE PROPIEDADES PETROFISICAS EN YACIMIENTOS DE DEPOSITACION FLUVIAL EN COLOMBIA: CAMPO CASTILLA*

AUTORES: SILVIO JOSE DIAZGANADOS DIAZ**
FELIPE ANDRES NIETO FERNANDEZ**

PALABRAS CLAVE: Caracterización de Yacimientos, Modelamiento Geoestadística, Campo Castilla, Deposición fluvial, Unidad Guadalupe masivo, modelamiento Petrofísico.

DESCRIPCION

En este trabajo se presenta una metodología completa para la caracterización de yacimientos mediante la aplicación de herramientas geoestadísticas. El proceso de caracterización de yacimientos tiene como objetivo construir un modelo del yacimiento, lo mas realista posible, mediante la incorporación de toda la información disponible. El enfoque clásico consiste en elaborar un modelo que se fundamente en la información estática del yacimiento, teniendo como etapa final del proceso, la validación del modelo con la información disponible. Es importante aclarar que el término validación implica un proceso por naturaleza puntual, generalmente asociado a asegurar solamente la coherencia requerida entre zonas productoras y propiedades petrofísicas.

El objetivo de la metodología propuesta, es servir de base para el estudio de los diferentes campos colombianos y así mejorar la capacidad de predicción del modelo del yacimiento mediante la integración *a priori* de los parámetros intrínsecos.

La metodología propuesta parte de la descripción detallada del campo castilla, haciendo énfasis en la unidad Guadalupe masivo, la cual incluye geología, marco estructural y litología de dicha unidad, así como los diferentes modelos petrofísicos que dan origen a la construcción del modelo geoestadístico. El cual se obtiene utilizando el software "SGSIM" de la librería "GSLIB". En este trabajo se presenta la aplicación de la metodología propuesta en un yacimiento de depositación fluvial en Colombia, como es el campo castilla (formación Guadalupe masivo, unidad k2). Los resultados obtenidos muestran el potencial de la técnica propuesta en el caso de un yacimiento tan estratigráficamente complejo como el que se presenta en un ambiente de depositación fluvial.

* Trabajo de Investigación

** Facultad de Ingenierías Físico-Químicas. Escuela de Ingeniería de Petróleos. Nicolás Santos Santos.

ABSTRACT

TITLE: IMPLEMENTATION OF GEOSTATISTICAL METHODS FOR MODELING THE PETROPHYSICAL PROPERTIES IN FLUVIAL DEPOSITS RESERVOIRS IN COLOMBIA, CASTILLA FIELD.

AUTHORS: SILVIO JOSE DIAZGANADOS DIAZ**
FELIPE ANDRES NIETO FERNANDEZ**

KEYWORDS: Reservoir characterization, geostatistical modeling, Castilla Field, Fluvial Deposits, Massive Guadalupe Unit, Petrophysical Modeling.

DESCRIPTION:

This work presents a complete methodology for the reservoir characterization by the application Of geostatistical tool.

The process of reservoir characterization has as objective to build a model of the reservoir, as realistic a possible, by the incorporation of all the information available. The classical approach consists in making a model based in the static information of the reservoir, having as a final stage of the process, the validation of the model with the available data. It's important to clarify that the term validation, implies a process punctual by nature, generally associated to ensure just the coherence required between producer zones and petrophysical properties.

The objective of the proposed methodology is to serve as a base for the study in the many different Colombian fields, and so, to improve the capability of reservoir model prediction by the prior integration of intrinsic parameters.

This methodology begins by the detailed description of Castilla field, specifically the Massive Guadalupe Unit, which includes geology, lithology and structural parameters, besides the different petrophysical models that builds the geostatistical model. This is obtained using the SGSIM software of the GSLIB library.

This project presents the application of the proposed methodology in a fluvial deposits reservoir in Colombia, specifically Castilla Field (Massive Guadalupe Formation, K2 unit). The results obtained show the potential of this technique applied to a complex structure reservoir, as the one that appears in an environment of fluvial deposits.

** Facultad de Ingenierías Físico-Químicas. Escuela de Ingeniería de Petróleos. Nicolás Santos Santos.

INTRODUCCION

El proceso de caracterización de yacimientos tiene como objetivo construir un modelo lo más realista posible, mediante la incorporación de toda la información disponible. El enfoque clásico consiste en elaborar un modelo que se fundamenta en la información estática del yacimiento, teniendo como etapa final del proceso la validación del modelo con la información disponible. En este trabajo se presenta una metodología novedosa para la caracterización de yacimientos altamente heterogéneos mediante la integración de la información del yacimiento al modelo estático actualizado.

El objetivo de la metodología propuesta es mejorar la capacidad de predicción del modelo del yacimiento mediante la integración a priori de los parámetros intrínsecos.

La metodología utiliza la herramienta software “SGSIM” especialmente diseñado para encontrar la mejor distribución petrofísica (de porosidad y permeabilidad) que haga honor tanto a los datos de los pozos como a sus historias de producción, siempre y cuando se tengan los modelos geológico, estratigráfico y petrofísico previamente definidos. Lo más importante de la metodología consiste en obtener un modelo experimental o variograma que tenga la mejor tendencia y que sea lo más continuo posible con el fin de ajustarlo a un modelo teórico predefinido.

Para encontrar la mejor distribución fue necesario hacer sensibilidad a los diferentes parámetros de búsqueda, como son: ángulos, tolerancias, ancho de bandas, etc. Con esto se obtuvieron los variogramas con las mejores tendencias en las diferentes direcciones.

Los resultados obtenidos con esta metodología suministran una información muy valiosa a la hora de perforar nuevos pozos, ya que a través de la distribución petrofísica resultante se tiene conocimiento de la ubicación de las mejores zonas productoras.

1. ESTADO DEL ARTE

En el presente capítulo, se detalla como se inició la geoestadística, las principales teorías que componen esta ciencia, los personajes que desarrollaron estas teorías y cuales fueron los diferentes campos en los que empezó a ser aplicada.

Además se presenta la definición teórica de la geoestadística, se indican cuales son los pasos principales que se deben seguir para desarrollar un estudio geoestadístico, y las teorías principales en las cuales se ha desarrollado y se desarrolla esta ciencia.

1.1 RESEÑA HISTÓRICA

Son muchos los ejemplos en las Ciencias de la tierra que conducen al tratamiento de variables distribuidas espacialmente, por ejemplo la estimación de recursos minerales, la caracterización de parámetros geofísicos en yacimientos de petróleo, la evaluación de variables ambientales, etc. Para el estudio de estos datos, que tienen asociada siempre una posición en el espacio y conocidos como variables regionalizadas (Matheron, 1970), se usan actualmente métodos geomatemáticos, dentro de los cuales la geoestadística ha jugado en las tres últimas décadas un papel importante y ha probado su superioridad entre tantos métodos de estimación de reservas en numerosos tipos de minas (Armstrong y Carignan, 1997). Su extensión a otros campos de aplicación, aunque reciente, ha mostrado valiosos resultados, como se puede apreciar en los trabajos de Galli (1997) en la industria del petróleo y Myers (1988) en estudios ambientales, además de una amplia gama de trabajos de otros autores en diferentes campos de aplicación de las geociencias. La geoestadística se define de forma general como el estudio de las variables numéricas distribuidas en el espacio (Chauvet, 1994). Este método exige

para su aplicación el cumplimiento de ciertas hipótesis de estacionaridad (Journel y Huijbregts, 1978; Armstrong y Carignan, 1997), las cuales pueden ser verificadas a través del análisis exploratorio de los datos experimentales. Su estimador, el kriging, término creado por G. Matheron en 1962, tiene como objetivo encontrar el Mejor Estimador Lineal Insesgado a partir de la información disponible (Chica-Olmo, 1987). En efecto, el estimador obtenido (Z^*) consiste en una combinación lineal del producto de una ponderación de pesos asociados a cada localización muestreada por el correspondiente valor Z del fenómeno estudiado, observando dos condiciones fundamentales:

1) Que el estimador sea insesgado: $E [Z^* - Z] = 0$; y 2) que la varianza $\text{Var} [Z^* - Z]$ sea mínima.

La estimación obtenida de esta forma, al igual que otros interpoladores, proporciona una imagen suave o lisa de la realidad, pero como en muchas ocasiones no es suficiente conocer sólo los valores estimados, sería conveniente una representación capaz de reflejar las fluctuaciones siempre presentes en la práctica, lo que es posible utilizando la dispersión del fenómeno estudiado (Journel, 1974). Como las variables regionalizadas son consideradas en geoestadística como una función aleatoria, es posible obtener una realización particular de ésta (Journel y Huijbregts, 1978), es decir, una simulación que no sea una representación suave o lisa de la realidad, sino más bien una de las posibles realizaciones del fenómeno estudiado, de modo que presente la misma dispersión, o lo que es igual, las mismas características de variabilidad y correlación espacial encontrada en los datos experimentales, y por tanto que en la realidad desconocida.

En este capítulo se presentan los elementos fundamentales de la teoría de la simulación geoestadística, la cual desde su introducción por Matheron (1962) y su desarrollo por Guibal (1972) y Journel (1974) ha jugado un importante papel en las aplicaciones de la geoestadística (Lantuéjoul, 1997); en la simulación condicional de yacimientos mineros (Journel, 1974; Chica-Olmo, 1987), en la simulación de

litofacies en la industria del petróleo (Le Loc'h y Galli, 1997); en la representación de modelos de yacimientos petroleros (Allard, 1993); en la modelación de procesos hidrogeológicos (Dietrich, 1995); en la presentación de procesos geofísicos (Christakos y Panagopoulos, 1992), etc.

1.2 CONCEPTOS DE GEOESTADÍSTICA

La geoestadística es una rama de la estadística que hace posible, con relativamente poca información, caracterizar cuantitativamente la variabilidad y la distribución espacial de variables en una forma matemáticamente útil. Dicho en otras palabras, esta ciencia permite crear mapas de isolíneas de una variable específica, mediante los cuales se puede predecir, de una manera certera, el comportamiento de la variable en sitios donde se desconoce su valor, a partir de puntos con información conocida de los que se disponga en la zona de estudio.

Las técnicas geoestadísticas fueron originalmente desarrolladas por ingenieros de minas para caracterizar la estructura espacial de los depósitos minerales, pero luego fueron extendidas a prácticamente cualquier problema que involucre alguna variable en una, dos o tres dimensiones. La geoestadística puede ser utilizada para cualquier variable cuya medición sea hecha sobre un dominio, y donde se espera que su valor esté determinado por su posición y su relación con sus vecindades (Franco F. y Mulford A., 1991).

Esta clase de variables son la base conceptual de la geoestadística, y se conocen en la literatura como variables regionalizadas. Estas, tal como lo enunció Matheron en 1965, se refieren a variables que caracterizan fenómenos que se distribuyen en el espacio y/o tiempo, y que presentan una cierta estructura de auto correlación.

La geoestadística busca extraer de la información experimental, (mediciones de la variable regionalizada), los rasgos y características principales del fenómeno, de modo que puedan utilizarse en métodos de estimación y simulación de las variables involucradas. Para lograr esta caracterización, la geoestadística se fundamenta en el análisis de variabilidad espacial del fenómeno a través del variograma; el cual cuantifica probabilísticamente la dispersión espacial de la variable regionalizada (Chica M., 1987). Este análisis es el fundamento del método geoestadístico de Kriging, cuya técnica permite estimar valores de la variable en sitios en donde se desconoce su valor, y de esta forma, generar mapas de distribución espacial de la variable regionalizada.

Los pasos principales de un estudio geoestadístico son:

- Análisis de los datos exploratorios.
- Análisis estructural (cálculo y modelamiento de los variogramas).
- Estimaciones (Kriging o simulaciones).

1.3 VARIABLES REGIONALIZADAS

La variable regionalizada es la representación de cualquier fenómeno que extendiéndose en el tiempo, muestra cierta estructura o distribución espacial y en cuya variación se aprecie una continuidad aparente.

Journel y Huijbregts (1978) plantean la definición de variable regionalizada, como una variable puramente descriptiva y envuelve una interpretación probabilística, refiriéndose a que, desde el punto de vista matemático una variable regionalizada es simplemente una función $f(x)$ que toma valores en todos los puntos x de coordenadas (x_i, y_i, z_i) en el espacio tridimensional. Sin embargo, es muy frecuente que estas funciones varíen tan irregularmente en el espacio que impida

un estudio matemático directo, y se hace necesario realizar un análisis de variabilidad de la información disponible, sugiriendo un estudio profundo de la función variograma como veremos más adelante.

En términos teóricos es oportuno aclarar que una variable aleatoria (VA) es una variable que puede tomar ciertos valores de acuerdo a cierta distribución de probabilidades. Un valor medido en cada punto x_i es considerado como una realización $z(x_i)$ de una VA $Z(x_i)$ cuya media es $m(x_i)$. En los puntos x donde no existen valores medidos es desconocida la propiedad que se estudia, pero están bien definidos y pueden asimismo considerarse variables aleatorias $Z(x)$. Al conjunto de todas las mediciones $z(x)$ en el área de estudio de la variable regionalizada puede considerarse como una realización particular del conjunto de VAs ($Z(x), x \in \text{área de estudio}$). A este conjunto de VAs se llama Función Aleatoria y se escribe $Z(x)$ (Journel y Huijbregts, 1978; Armstrong y Carignan, 1997). De modo que al extender el concepto de función aleatoria al espacio de una o más dimensiones, aparece la noción aleatoria y estructural de una variable regionalizada: primero $Z(x)$ como VA y segundo que las VAs $Z(x)$ y $Z(x+h)$ no son en general independientes, sino que están relacionadas por la estructura espacial de la variable regionalizada original $Z(x)$.

En la estadística clásica se puede dar el caso de dos muestras con la misma media y varianza, pero con una continuidad y distribución espacial entre muestras completamente diferentes. La importancia de la variable regionalizada y de la geoestadística, reside en el hecho de que logran capturar esa continuidad, pues que tienen en cuenta la geometría de la muestra: su volumen, forma, tamaño, etc. (Clark I., 1979).

1.3.1 Hipótesis restrictivas. Para poder aplicar los resultados de la teoría de las funciones aleatorias al análisis de la dispersión espacial de los datos, es necesario conocer parte de su ley de distribución; este proceso se llama inferencia estadística. Para encontrar una solución a esta, se requiere producir dos hipótesis restrictivas (Chica M,. 1987). Estas hipótesis se hacen en los dos primeros momentos de la ley de probabilidad de la función aleatoria, en decir, la media y la varianza. Por esta razón antes de estudiarla, es importante recordar la definición de varianza.

- **Varianza:** es una desviación absoluta promedio calculada para un conjunto de datos N, para esto se utiliza la suma de los cuadrados de las diferencias (desviaciones) entre cada dato y la media, dividido entre el número total de datos.

Luego se puede calcular:

$$Var(x) = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N} \dots\dots\dots Eq.1.1$$

Var(x)= Varianza

Xi = Dato

μ = Media del conjunto de datos

N = Número de datos

La varianza generalmente se denota con σ^2 y su correspondiente desviación estándar σ , es simplemente la raíz cuadrada de la varianza.

- **Hipótesis de Estacionaridad:** se define que existe estacionaridad en un área dada cuando todas las muestras sujetas a estudios presentan la

misma distribución de probabilidad, es decir, estadísticamente el valor de propiedad en estudio no varía en el espacio.

El modelo de la función aleatoria asume que cada medición es un valor particular de una variable aleatoria, la suposición de variable regionalizada implica que las variables aleatorias $Z(x_i)$ estén relacionadas espacialmente por alguna función. Esta función, depende solamente de la distribución espacial de la variable regionalizada según sea determinada para un primer o un segundo orden de estacionaridad (Chica M., 1987).

Una variable regionalizada es de primer orden de estacionaridad si su ley de distribución de probabilidad es invariable por traslación, es decir:

$$f[Z(x_1), Z(x_2), \dots, Z(x_n)] = f[Z(x_1+h), Z(x_2+h), \dots, Z(x_n+h)] \dots \dots \text{Eq.1.2}$$

La traslación vendría marcada por el módulo del vector h , el cual representa la distancia y orientación relativa entre dos puntos.

Una variable regionalizada es de segundo orden de estacionaridad si su media representada por E y su varianza representada por Var , existen y son invariantes por traslación, es decir:

$$E[Z(x)] = E[Z(x+h)] \dots \dots \dots \text{Eq.1.3}$$

$$Var[Z(x_1)] = Var[Z(x_1+h)] \dots \dots \dots \text{Eq.1.4}$$

- **Hipótesis Intrínseca:** es una teoría que permite superar la dificultad creada por la dispersión que puede mostrar la variación en el espacio de la

variable regionalizada. Se basa en la estacionaridad de los incrementos de la pareja de variables aleatorias $Z(x+h)$, $Z(x)$ y no de la variable aleatoria $Z(x)$ en sí misma.

Considerando que h representa la distancia y orientación relativa entre dos puntos; la diferencia $Z(x+h) - Z(x)$ depende solo de h , al igual que su media y varianza. Es decir, llamando a la media de esta diferencia $m(h)$, a su varianza $2\gamma(h)$ y reemplazando $Y(x+h) = Z(x+h) - Z(x)$, se tiene:

$$m(h) = E[Y(x+h)] \dots \dots \dots Eq.1.5$$

$$2\gamma(h) = Var[Y(x+h)] = E[\{Y(x+h) - m(h)\}^2] \dots Eq.1.6$$

La hipótesis intrínseca define que el valor del incremento de la variable en estudio entre dos puntos separados por una distancia h , sea nulo, es decir, que no haya diferencia entre dos observaciones cuya separación sea tan grande como h . Esta hipótesis se basa en la hipótesis de estacionaridad de segundo orden, que hace que la media sea invariable por traslación. Con base en lo anterior, $m(h)$ es 0 y se tiene que:

$$2\gamma(h) = E[\{Z(x+h) - Z(x)\}^2] \dots \dots \dots Eq.1.7$$

Que corresponde a la definición de variograma (Franco F. y Mudford A., 1991).

1.4 VARIOGRAMAS

Los variogramas cuantifican la correlación espacial de una determinada variable. Cualitativamente, el variograma de un litotipo describe tanto el tamaño como la

frecuencia espacial de ocurrencia. El primer paso del modelamiento de los variogramas es calcular el variograma experimental a partir de la información de pozo y luego ajustarlo con un modelo de función Gaussiana truncada (modelo teórico). Como por ejemplo: detectar direcciones de anisotropía, zonas de influencia y su extensión, variabilidad con la distancia. Se denota $\gamma(h)$.

El variograma esta compuesto por los siguientes elementos:

- **Efecto Nugget (Co):** el variograma por definición es nulo en el origen, pero en la práctica las funciones obtenidas pueden presentar discontinuidad en el origen. Puede ser obtenido trazando una línea recta entre los primeros puntos del variograma empírico y extender ésta hasta que se intercepte con el eje Y. Si esta intersección ocurre por debajo de cero, el valor asumido por este efecto es cero, pues valores negativos de $\gamma(0)$ no tienen significado y no es común.
- **Meseta (Sill):** es el valor de $\gamma(h)$ para el cual con el aumento de h su valor permanece constante, se representa como $(C_T = C + C_o)$ y se denomina meseta. Puede obtenerse trazando una línea paralela a la abscisa y que se ajuste a los puntos de mayor valor del variograma y su valor se lee en la intersección de esta línea con la ordenada.
- **El Alcance (Range):** La distancia h para la cual las variables $Z(x)$ y $Z(x+h)$ son independientes, se denomina alcance y se representa por (a), es decir, las distancias para la cual los valores de la variable dejan de estar correlacionados, o lo que es lo mismo, la distancia para la cual el variograma alcanza su meseta.

El alcance siempre tiene valor positivo y puede ser obtenido a partir de la intersección de las líneas descritas en los puntos anteriores, ese punto leído en la abscisa es una fracción del propio alcance, fracción que se detallará posteriormente en la explicación de los modelos teóricos.

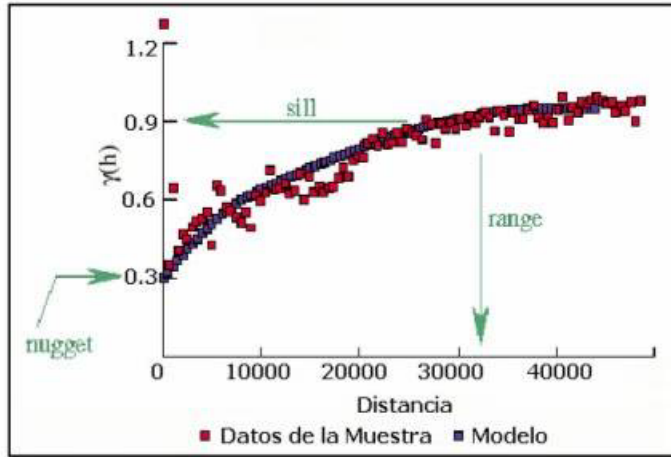


Figura 1.1. Elementos del Variograma

1.4.1 Cálculo del variograma (Experimental). Se define como la media aritmética de todos los cuadrados de las diferencias entre pares de valores experimentales separados una distancia h (Journel y Huijbregts, 1978), o lo que es lo mismo, la varianza de los incrementos de la variable regionalizada en las localizaciones separadas una distancia h .

La función $\gamma(h)$ se denomina variograma, la cual puede ser obtenida por la expresión.

$$\gamma(h) = \frac{1}{2Np(h)} \sum_{i=1}^{Np(h)} [Z(x_i) - Z(x_i + h)]^2 \dots\dots\dots Eq.1.8$$

- Donde:
- $Np(h)$ es el número de pares a la distancia h .
 - h es el incremento.
 - $Z(x_i)$ son los valores experimentales.
 - x_i localizaciones donde son medidos los valores $z(x_i)$.

Esta expresión de $\gamma(h)$ representa el útil más importante en todo estudio geoestadístico (Armstrong y Carignan, 1997; Weerts, y Bierkens, 1993; Chica, 1987). Su cálculo no consiste en una simple evaluación de su expresión, según se plantea en (Krajewski y Gibbs, 1993; Journel y Huijbregts, 1978; David, 1977; Xie y Myers, 1995a; Pannatier, 1993) esta operación está relacionada con los elementos siguientes:

- La dirección en la que será calculado el variograma, uno o dos ángulos que definen una dirección en el espacio α y/o β con tolerancias angulares $d\alpha$ y/o $d\beta$. El variograma calculado usando tolerancia angular de 90° se denomina “variograma medio”, “global” u “omnidireccional” como ya se indicó.
- El incremento o paso en el cálculo del variograma h y su tolerancia lineal dh , se recomienda que el valor de dh sea la mitad del incremento inicial.
- Una distancia, que representa la distancia máxima a que pueden estar alejados los segundos puntos del par con respecto a la línea que define la dirección de cálculo, conocido como ancho de banda.
- La distancia L_{\max} hasta la cual será calculado del variograma. Se recomienda que ésta sea la mitad de la distancia entre las muestras más alejadas (Armstrong y Carignan, 1997; Krajewski y Gibbs, 1993), aunque dependiendo de la geometría del fenómeno regionalizado en algunos casos puede ser calculado hasta una distancia superior.

1.5 MODELAMIENTO DE VARIOGRAMAS

Es el proceso mediante el cual, una vez que el variograma experimental ha sido graficado, se ajusta a uno de los modelos teóricos existentes, con el fin de caracterizar el comportamiento espacial de la variable. Cada uno de estos modelos está definido por una función matemática que permite determinar la

variabilidad o continuidad especial de la variable regionalizada (Franco F. y Mulford A., 1991).

Existen muchos modelos teóricos, algunos de los cuales se presentan a continuación, junto con sus principales características:

- **Modelo Lineal:** es el modelo más simple y utiliza una ecuación de la forma:

$$\gamma(h) = Ah \dots \dots \dots Eq.1.9$$

El término A, representa la pendiente de este modelo.

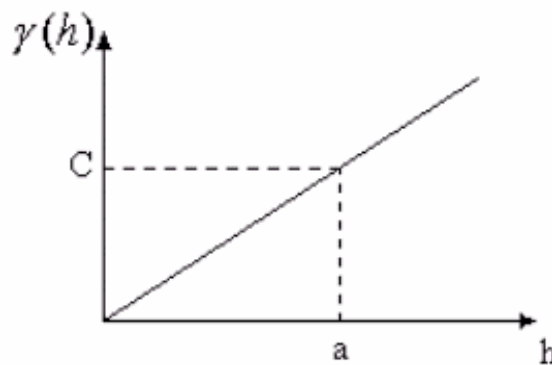


Figura 1.2. Modelo Lineal

Como se observa en la figura, este es uno de los modelos donde no se alcanza el umbral y por lo tanto no existe un rango de influencia. En este modelo de variograma los datos no se estabilizan, y por lo tanto no llegan a ser independientes a medida que se incrementa la separación entre ellos, sino que por el contrario el valor del variograma continúa cambiando con una tendencia marcada.

- **Modelo Esférico:** este modelo tiene la forma que se presenta en la siguiente figura:

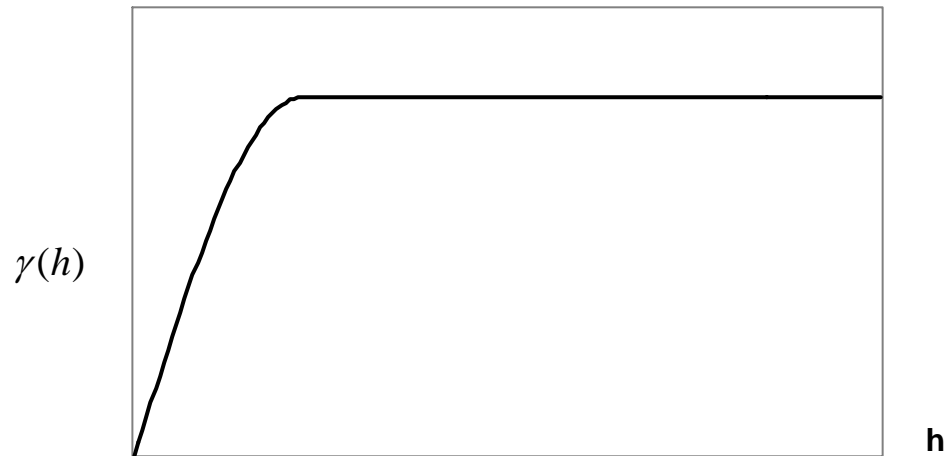


Figura 1.3. Modelo Esférico

Matemáticamente esta dada por la ecuación:

$$\gamma(h) = C \left[\frac{3h}{2a} - \frac{h^3}{2a^3} \right] \dots h \leq a$$

$$\gamma(h) = C \dots h > a \dots \dots \dots Eq.1.10$$

Este modelo a diferencia del lineal presenta un umbral (C) y un rango de influencia (a); lo que significa que para estos valores, los datos empiezan a ser independientes entre si; es decir, la variación entre los datos (valor de variograma) se estabiliza y deja de cambiar a medida que se incrementa la distancia de separación h.

- **Modelo Exponencial:** este modelo tiene la siguiente ecuación matemática:

$$\gamma(h) = C \left[1 - e^{-\frac{h}{a}} \right] \dots\dots\dots Eq.1.11$$

Este modelo se levanta a partir del origen más lentamente que el modelo esférico y nunca alcanza completamente su umbral, sin embargo, para conceptos prácticos se considera una línea recta tangente al origen intercepta el umbral del variograma a una distancia igual al rango de influencia.

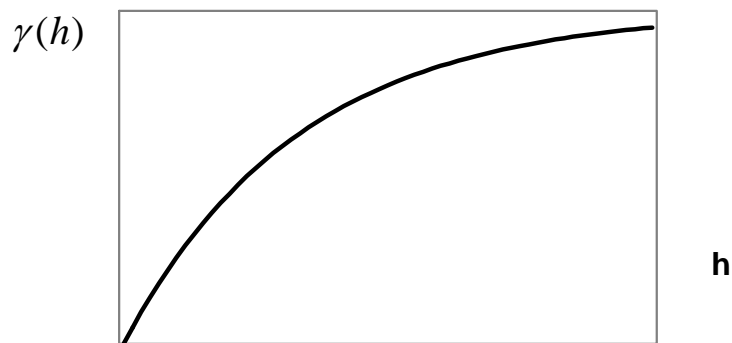


Figura 1.4. Modelo Exponencial

- **Modelo de Potencia:** este representado por la siguiente ecuación

$$\gamma(h) = C \left[h^n \right] \dots\dots\dots 0 < n < 2 \dots\dots\dots Eq.1.12$$

Cuando n es igual a 1, el modelo de potencia se convierte en un modelo lineal.

Este modelo que se ilustra en la siguiente figura, se ha utilizado en el estudio de la porosidad y conductividad hidráulica vertical en yacimientos de petróleos (Restrepo D., 2000)

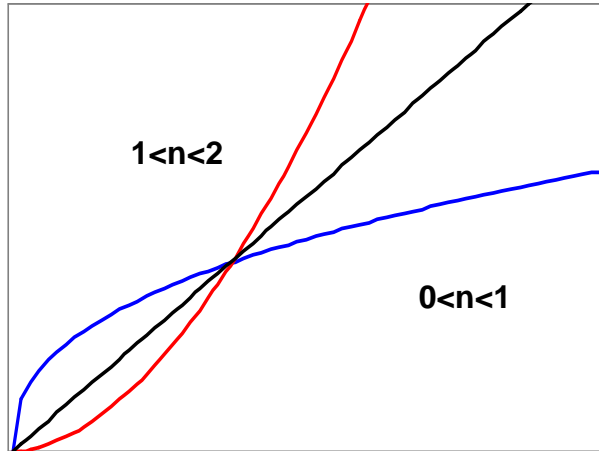


Figura 1.5. Modelo de Potencia

- **Modelo Gaussiano:** el esquema del modelo gaussiano se muestra en la siguiente figura y se expresa con la siguiente ecuación:

$$\gamma(h) = C \left[1 - e^{-(h/a)^2} \right] \dots \dots \dots Eq.1.13$$

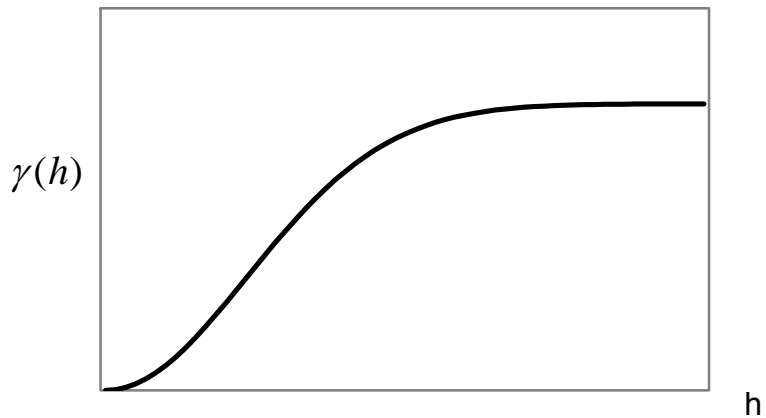


Figura 1.6. Modelo Gaussiano

Representa fenómenos continuos infinitamente diferenciables (sumamente continuos).

- **Modelo de Efecto Hueco:** se utiliza para representar fenómenos continuos con periodicidades y está definido por la ecuación:

$$\gamma(h) = C \left[1 - \frac{\text{Sen}(ah)}{ah} \right] \dots\dots\dots \text{Eq.1.14}$$

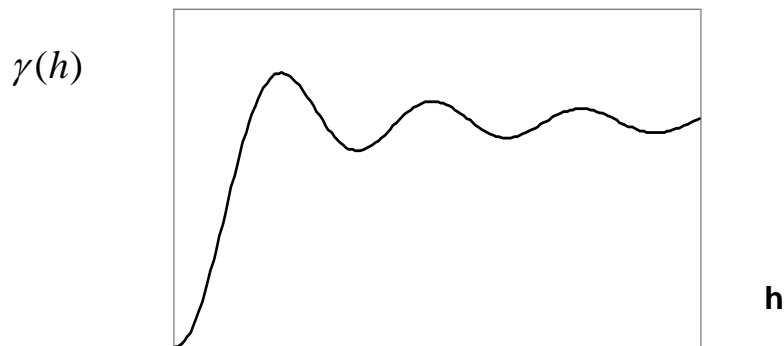


Figura 1.7. Modelo Efecto Hueco

- **Modelos Complejos:** es un modelo de variograma en cuya representación se combinan dos o más modelos teóricos de los que han sido definidos. Las combinaciones pueden incluir dos modelos de diferentes tipos, por ejemplo un esférico y un exponencial, o dos modelos del mismo tipo, por ejemplo, dos esféricos. Se usan en casos en que un solo modelo no satisface el modelamiento de un variograma experimental y se desea ajustar una parte usando un modelo, y la otra parte empleando otro.

1.6 ANISOTROPIA

Conviene aquí realizar un análisis sobre el comportamiento de la variabilidad del atributo en estudio. Se conoce que el variograma describe las características de continuidad espacial de la variable regionalizada en una dirección, pero este comportamiento pueden variar según la dirección que se analice, como se discute en Journel y Huijbregts (1978), David (1977), Zimmerman (1993), Krajewski y Gibbs (1993). Es por este motivo apropiado hacer un análisis del comportamiento de la continuidad en distintas direcciones, conocido como Anisotropía.

Se dice que una variable regionalizada tiene un comportamiento anisótropo cuando presenta direcciones particulares de variabilidad; tales direcciones privilegiadas corresponden generalmente a direcciones estructurales del fenómeno. El comportamiento anisótropo es fácilmente detectable calculando variogramas en diferentes direcciones del espacio (Chica M., 1987) y determinado a través de estos, la dirección en la que la variable tiene mayor continuidad. Como ya se mencionó, el parámetro h al que se refiere el variograma es un vector determinado por su magnitud y por su orientación; por tanto, es posible calcular el variograma para pares de puntos cuyos vectores \bar{h} , correspondan a una dirección específica.

La continuidad espacial de una variable puede presentar diferencias en los variogramas al cambiar la dirección de estos. De esta forma, cuando el variograma es una función tanto de la distancia como de la dirección, es necesario definir un modelo de variograma anisotrópico para la variable, especificando la dirección del eje en el que se presenta la anisotropía. Por el contrario, si el variograma es función solamente de la distancia, se puede considerar que el variograma corresponde a un modelo isotrópico.

Normalmente, son dos los tipos de anisotropía que se distinguen en el análisis de variogramas: la anisotropía geométrica y la anisotropía zonal (Franco F. y Mulford A., 1991).

- **Anisotropía Geométrica:** es la que más comúnmente ocurre y se presenta cuando al construir los variogramas de cualquier variable, se encuentran valores de rangos de influencia diferentes en direcciones diferentes.

Estos casos de anisotropía pueden ser tratados de una manera sencilla mediante una transformación simple que convierte el variograma en un modelo con un rango de influencia en uno.

Para variogramas anisotrópicos, se pueden volver a escribir las expresiones de los variogramas en las direcciones de los ejes mayor y menor del sistema coordinado de datos. Mediante esta transformación se crean variogramas isotrópicos que poseen el mismo umbral y rango de influencia.

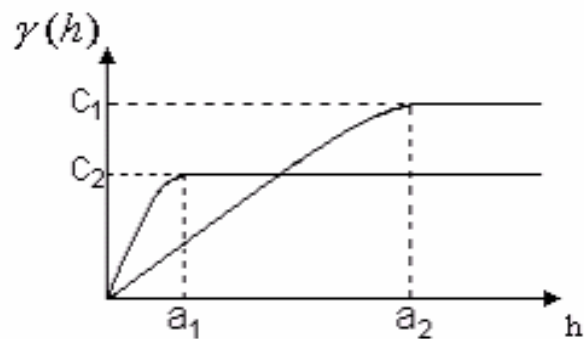


Figura 1.8. Anisotropía Geométrica

- **Anisotropía Zonal:** este tipo de anisotropía se presenta cuando el valor del umbral cambia con la dirección, mientras que el valor del rango de influencia permanece constante.

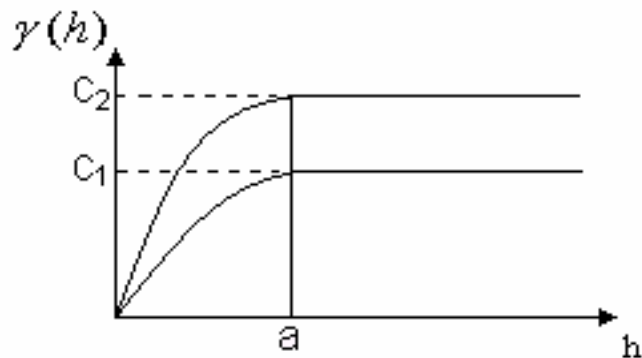


Figura 1.9. Anisotropía Zonal

1.7 TÉCNICA DE INTERPOLACIÓN GEOESTADÍSTICA DE KRIGING

Se ha visto cómo emplear los conceptos fundamentales de la geoestadística para construir un modelo que caracterice el comportamiento de una variable regionalizada cualquiera. Este modelo es la base para proceder a la estimación de la variable en los sitios del área de interés donde se desconoce su valor, a partir de las técnicas de interpolación de Kriging. El nombre de esta técnica es un homenaje al Dr. Diane Krige quien fue el primero en interesarse en las correlaciones espaciales existentes en los depósitos mineros (Franco F. y Mulford A., 1991).

Esta técnica permite estimar el valor de una variable en sitios de valor desconocido, mediante la asignación de ponderadores a cada uno de los puntos de medición en los que se conoce el valor de la variable. Estos ponderados no sólo están condicionados por la ubicación espacial de todos los puntos entre sí, y de estos con el punto o área a estimar; sino que se calculan para el mínimo error de estimación, de tal manera que el estimador que se obtiene es bastante confiable.

El Kriging es conocido como el método interpolador geoestadístico, es un estimador lineal insesgado, que presenta dos propiedades básicas que son: hacer que la suma de errores tienda a cero, y que el cuadrado de las desviaciones sea mínimo.

- **Proceso de estimación:** este proceso requiere definir el punto, área o volumen sobre el cual no se conoce el valor de la variable en estudio y se desea hacer el estimativo de él. Se denomina a este valor desconocido Z y al volumen, área o punto de interés, A . Normalmente un grupo n de muestras con valor $g_1, g_2, g_3, \dots, g_n$, que es denotado como S , puede servir como base para formar un estimador no lineal, es decir un ponderado promedio. Este estimador se designa como Z^* y es igual a:

$$Z^* = W_1g_1 + W_2g_2 + W_3g_3 + \dots + W_n g_n \dots\dots\dots Eq.1.15$$

Donde $W_1, W_2, W_3, \dots, W_n$, son los ponderados asignados a cada muestra.

El caso mas simple de todos se presenta cuando el valor de todos los ponderados es el mismo. En este caso, Z^* corresponde exactamente a la media aritmética de los valores de las muestras.

Considérese a modo de ejemplo, el grupo de cinco muestras conocidas y un punto A (en rojo), donde se desconoce el valor de la variable, el cual se presenta en la siguiente figura:

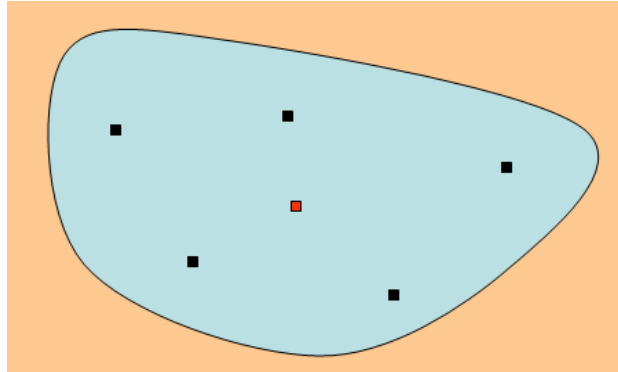


Figura 1.10. Grupo de muestras hipotético en un proceso de estimación

Las coordenadas de los seis puntos para una variable cualquiera dada se suponen conocidas. Una vez obtenido el variograma experimental, ajustado a uno de los modelos teóricos y leídos los valores del umbral, Efecto Nugget y rango de influencia, se procede a la estimación de la variable en el punto desconocido A.

El procedimiento más simple consiste en agregar el valor de la muestra más cercana (punto negro), a la muestra desconocida A. Este hecho convierte un error de estimación, ε , el cual es igual a la diferencia entre el valor, Z , del punto desconocido, y el valor estimado Z^* , asignado a él, que en este caso corresponde al valor de g_1 . Así:

$$Z^* = g_1 \dots \dots \dots Eq.1.16$$

$$\varepsilon = Z - Z^* \dots \dots \dots Eq.1.17$$

Teniendo en cuenta que se busca obtener un estimador imparcial, es decir, un estimador con el mínimo error de estimación; si se hicieran varias estimaciones, el error promedio deberá ser cero, $\bar{\varepsilon}$ (Clark I., 1979).

La confiabilidad del proceso de estimación se puede medir mediante el cálculo del error asociado a él. Si este error toma valores cercanos a cero, se puede decir que se cuenta con un buen estimador, pero si su valor es grande, no podrá existir confiabilidad en él. Estadísticamente la mejor estimación de la magnitud del error es la desviación estándar. De este modo la desviación estándar de un error de estimación, medirá la confiabilidad de un estimador. Sin embargo, hay que tener en cuenta que no importa cuantas mediciones se realicen, no se puede calcular la desviación estándar de los errores ya que no se conoce el error hecho. Por lo tanto se debe recurrir al concepto de varianza en el error de estimación. Es decir se tiene: $\varepsilon = Z - Z^*$

La varianza de los errores, σ_ε^2 , será igual a la desviación promedio al cuadrado del error medio:

$$\begin{aligned}\sigma_\varepsilon^2 &= \text{Promedio } (\varepsilon - \bar{\varepsilon})^2 \\ &= \text{Promedio de } \varepsilon^2 \text{ puesto que } \bar{\varepsilon} = 0 \\ &= \text{Promedio } (Z - Z^*)^2\end{aligned}$$

El mismo proceso de estimación debería ser repetido para todos los puntos muestreados. En la práctica este es difícil de realizar, de modo que hay que obtener una forma práctica de calcular la varianza. El proceso consiste en tomar el valor de la variable de estudio en el punto A, restarlo al valor de la variable en un punto conocido, elevar el resultado al cuadrado y repetir el proceso para todos los pares posibles de puntos, para finalmente, promediar los valores obtenidos. Esta definición corresponde a la del variograma, entre dos puntos A (desconocido) y 1 (conocido). Dada la distancia entre ellos, h, se puede evaluar la varianza en la estimación, leyendo el valor de la gráfica del modelo de variograma y multiplicándolo por dos. (Clark, I., 1979).

$$\sigma_{\varepsilon}^2 = 2\gamma(h) \dots \dots \dots Eq.1.18$$

El proceso es más complicado cuando en lugar de estimar un valor en un punto A, el interés se centra en un área o celda de una malla definida. La siguiente figura muestra una celda en cuyo centro está localizado el punto A. En este caso el valor de Z corresponderá al valor promedio de la variable en la celda, A será el área de la celda y las otras condiciones se mantiene:

$$Z^* = g1, \varepsilon = Z - Z^*$$

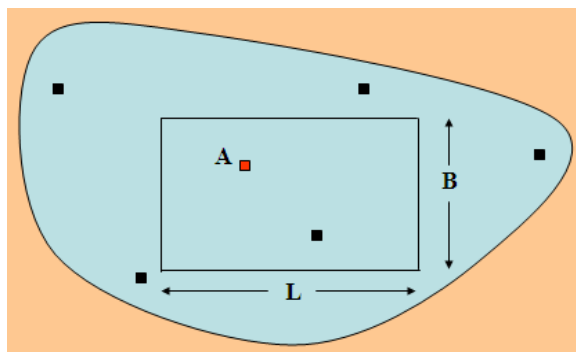


Figura 1.11. Estimación del valor de una variable en una celda de una malla

La varianza de la estimación de la variable sigue siendo el valor del variograma, pero ahora el variograma es calculado entre el valor de muestra g1, y el valor promedio de la variable en la celda A. El modelo de variograma hasta ahora utilizado calcula la diferencia de los valores de una variable entre dos puntos, pero no considera la probabilidad de tener tamaños diferentes para comparar, como en el caso entre un punto y una celda.

El valor del variograma entre un punto y una celda y cada punto dentro de la celda A y el promedio de esos valores pueden ser calculados normalmente. Este último

valor se define como $\bar{\gamma}(S, A)$, y corresponde a la media aritmética entre la muestra S y cada punto en la celda A, reemplazando el término $\gamma(h)$ en las ecuaciones donde este último aparece. $2\bar{\gamma}(S, A)$ Será entonces la varianza del error cometido al estimar cada punto dentro de la celda A y la muestra S. Sin embargo, lo que realmente se necesita es el valor del variograma entre el valor promedio de la celda A y la muestra S y no, entre todos los puntos individuales de la celda A y la muestra S.

Para corregir esta diferencia de valores, es necesario tomar en cuenta la variación de los valores en los puntos de la muestra dentro de la celda, la cual es evaluada mediante el uso de una función auxiliar F (L, B), que multiplicada por el umbral define la varianza de los valores entre todos los pares de puntos posibles dentro de la celda de dimensiones L y B. Los valores de esta función han sido tabulados para los diferentes modelos de variogramas estandarizados, es decir con rango de influencia y umbral igual a 1.

El valor que se obtiene de este proceso es equivalente al término $\bar{\gamma}(A, A)$, que es el valor del variograma promedio entre punto y punto de la celda. Al usar el punto de muestra para estimar el valor promedio de la celda en la figura anterior, se obtiene una varianza en la estimación igual a:

$$\sigma_{\varepsilon}^2 = 2\bar{\gamma}(S, A) - \bar{\gamma}(A, A) \dots \dots \dots Eq.1.19$$

Como generalmente se tiene más de una muestra disponible en el área de interés, es conveniente utilizarlas en su totalidad en el procedimiento de estimación. Para esta situación se puede suponer a Z^* como la media aritmética de los valores de las muestras, lo que suministra una forma muy simple del tipo de estimador ponderador promedio. Para este caso se tendrá lo siguiente:

$$Z^* = \frac{1}{n}(g_1 + g_2 + g_3 + \dots + g_n) = \frac{1}{n} \sum_{i=1}^n g_i \dots \dots \dots Eq.1.20$$

Considerando la nueva situación, el término $\bar{\gamma}(S, A)$ corresponde al valor del variograma entre cada punto en el grupo de muestras S y cada punto en la celda A, y el término $\bar{\gamma}(A, A)$ continúa siendo el variograma promedio entre punto y punto en la celda. Sin embargo, se presenta una fuente adicional de variación puesto que no se ha tenido en cuenta la variación entre puntos de muestra; indicada por el término $\bar{\gamma}(S, S)$, que corresponde al valor del variograma promedio entre punto y punto de muestras (Clark I., 1979).

La expresión final de la varianza en la estimación será para este caso:

$$\sigma_\varepsilon^2 = 2\bar{\gamma}(S, A) - \bar{\gamma}(S, S) - \bar{\gamma}(A, A) \dots \dots \dots Eq.1.21$$

Esta ecuación indica que la confiabilidad de un estimador depende de tres cantidades: la varianza entre las muestras y el área a ser estimada, $2\bar{\gamma}(S, A)$; la varianza entre las muestras, $\bar{\gamma}(S, S)$; y la variación entre los puntos dentro de la celda, $\bar{\gamma}(A, A)$.

- **Kriging Simple:** la hipótesis básica es la estacionaridad junto con el hecho de que se asume que la media de la función aleatoria es conocida y constante en todo el dominio de la estimación. el estimador de kriging simple es igual al valor conocido de la variable multiplicado por la correlación que existe entre la variable en el punto objetivo y la variable en el punto de observación.

Matemáticamente se representa de la siguiente manera:

$$\sigma_{SK}^2 = E[(Z - Z^*)^2] \dots \dots Eq.1.22$$

Donde: σ_{SK}^2 es la varianza.

E es la media.

Z^* es el valor a estimar.

Z es valor conocido.

- **Kriging Ordinario:** a diferencia del Kriging simple, este tipo de Kriging calcula la media a partir de datos que intervienen en la estimación, es decir, la media es desconocida.

Como se mencionó anteriormente en la estimación de la varianza, solo se consideró el promedio de todas las muestras como el valor del estimativo de una variable en el área de interés. La realidad es que existen casos en los cuales no parece razonable ponderar todas las muestras igualmente ya que algunas se encuentran a gran distancia del área de consideración A, mientras que otras están más cerca o dentro de ella. De este modo, es más razonable utilizar un promedio ponderado de los valores de las muestras, otorgando a las muestras más cercanas una mayor importancia. El nuevo estimador Z^* será de la siguiente forma:

$$Z^* = W_1g_1 + W_2g_2 + W_3g_3 + \dots + W_n g_n \dots \dots Eq.1.23$$

Donde $W_1, W_2, W_3, \dots, W_n$, son los ponderados, cuya suma debe ser igual a 1:

$$W_1 + W_2 + W_3 + \dots W_n = 1 \dots \dots Eq.1.24$$

Cuando esta condición se cumple y no existe presencia de tendencias en la continuidad de la variable, se puede afirmar que el estimador Z^* es un estimador imparcial, lo cual significa que al realizar un gran número de estimaciones, el error promedio será cero. El estimador Z^* , también es lineal porque es una combinación lineal de los valores de las muestras. La media aritmética es un caso especial de estimador lineal en donde todos los ponderados son iguales (Clark I., 1979).

A partir del concepto de ponderador lineal imparcial, se puede mostrar que en general la varianza en la estimación es definida como:

$$\sigma_{\varepsilon}^2 = 2 \sum_{i=1}^n W_i \bar{\gamma}(S_i, A) - \sum_{i=1}^n \sum_{j=1}^n W_i W_j \bar{\gamma}(S_i, S_j) - \bar{\gamma}(A, A) \dots Eq.1.25$$

Aunque cada uno de los términos de la ecuación ya ha sido definido, ahora es importante observar que donde previamente se había calculado $\bar{\gamma}(S, A)$, que es el variograma promedio entre cada muestra y el área desconocida, se forma ahora un ponderador promedio para cada muestra individual con el área A, $\bar{\gamma}(S_i, A)$. Por ejemplo, si se tienen n muestras tomadas alrededor del área A, el primer término de la ecuación anterior será:

$$\sum_{i=1}^n W_i \bar{\gamma}(S_i, A) = W_1 \bar{\gamma}(Punto1, A) + W_2 \bar{\gamma}(Punto2, A) + \dots + W_n \bar{\gamma}(Punto n, A)$$

El cálculo de este término significa un problema puesto que el variograma permite comparar puntos de igual tamaño, y no involucra casos como el de un punto y una celda. Para este procedimiento se emplea también una función auxiliar llamada H (B, L), que multiplicada por el umbral define la varianza de los valores entre cada muestra y la celda A, de dimensiones L y B. Al igual que la función F (L, B), los

valores de esta función han sido tabulados para modelos de variogramas estandarizados.

El último término de la ecuación de estimación de varianza no cambia de forma ya que en esta solamente se ha cambiado la forma del estimador y no el área que está siendo estimada. El término $\bar{\gamma}(S, S)$, el cual anteriormente evaluaba la variación en los valores entre las muestras, debe ahora tener en cuenta los diferentes ponderadores asociados a cada una de ellas. Si se tienen por ejemplo, 4 muestras, se tendrán de igual forma 4 ponderadores. De este modo, si se toma la muestra 4 con la muestra 2, se deben incluir los dos ponderadores correspondientes W_4 y W_2 , de forma que el término $W_i W_j \bar{\gamma}(S_i, S_j)$ será:

$$W_2 W_4 \bar{\gamma}(S_2, S_4)$$

Entonces, para formar el término equivalente $\bar{\gamma}(S, S)$, cada término $\bar{\gamma}(S_i, S_j)$ debe ser multiplicado por los correspondientes $W_i W_j$, antes de ser adicionados a la sumatoria del segundo término en la ecuación de estimación de la varianza.

Como se menciona anteriormente, el método de Kriging utiliza un estimador de tipo lineal con un ponderador promedio de los valores de las muestras para evaluar la propiedad en estudio. Este estimador se dice que es imparcial si la suma de los ponderados asignados a cada muestra es igual a 1. Sin embargo, existe un gran número de estimadores lineales imparciales, de modo que para definir cual es el mejor, se aplica como criterio el que tenga la menor varianza en el proceso de estimación. La expresión para la varianza en la estimación depende de tres cosas: la geometría básica de las muestras y el área desconocida, la forma del variograma y el ponderador asignado a cada muestra. De este modo para cualquier grupo muestral dado, la varianza puede ser cambiada solamente

mediante la alteración en los valores de los ponderadores. Como el objetivo es minimizar la varianza en la estimación de la variable de interés con respecto a los ponderadores y la varianza es función de ellos, el procedimiento consiste en diferenciar la varianza con respecto a los ponderadores e igualar esta diferencial a cero, como se muestra a continuación:

$$\frac{\partial \sigma_{\varepsilon}^2}{\partial W_i} = 0 \dots i = 1, 2, \dots, n \dots \dots \dots Eq.1.27$$

Este proceso de diferenciación arroja como resultado n ecuaciones con n incógnitas ($W_1, W_2, W_3, \dots, W_n$). Estos ponderadores suministrarán un estimador con el valor mínimo de la varianza en la estimación. Sin embargo, la suma de los ponderadores encontrados no será necesariamente igual a uno, ya que no existe nada en el sistema de ecuaciones que restrinja a los ponderadores de esa forma. Por lo tanto, es necesario satisfacer la siguiente ecuación:

$$\sum_{i=1}^n W_i = 1 \dots \dots \dots Eq.1.28$$

De esta forma, obtener el mejor estimador lineal imparcial requiere resolver (n + 1) ecuaciones y hasta ahora solo se tienen n incógnitas, lo que dificulta la resolución del sistema. Para superar este problema, es necesario introducir una nueva incógnita en la forma de un multiplicador de Lagrange, λ , que permita balancear el sistema. Por consiguiente, en lugar de minimizar la varianza de estimación σ_{ε}^2 , se minimiza ahora:

$$\sigma_{\varepsilon}^2 - \lambda \left(\sum_{i=1}^n W_i - 1 \right)$$

Con respecto a $W_1, W_2, W_3, \dots, W_N$ y λ (Clark I., 1979)

$$\partial \left[\frac{\sigma_{\varepsilon}^2 - \lambda \left(\sum_{i=1}^n W_i - 1 \right)}{\partial W_i} \right] = 0 \dots i = 1, 2, \dots, n \dots \text{Eq.1.29}$$

Una vez que la diferenciación ha sido realizada y las ecuaciones ordenadas, se obtiene el siguiente sistema:

$$\begin{aligned} W_1 \bar{\gamma}(S_2, S_1) + W_2 \bar{\gamma}(S_2, S_2) + \dots + W_n \bar{\gamma}(S_2, S_n) + \lambda &= \bar{\gamma}(S_2, A) \\ W_1 \bar{\gamma}(S_3, S_1) + W_2 \bar{\gamma}(S_3, S_2) + \dots + W_n \bar{\gamma}(S_3, S_n) + \lambda &= \bar{\gamma}(S_3, A) \\ \dots & \dots \dots \dots \\ W_1 \bar{\gamma}(S_n, S_1) + W_2 \bar{\gamma}(S_n, S_2) + \dots + W_n \bar{\gamma}(S_n, S_n) + \lambda &= \bar{\gamma}(S_n, A) \\ W_1 + W_2 + \dots + W_n &= 1 \end{aligned}$$

La mayoría de los términos en este sistema ya son conocidos. Si se toma por ejemplo la primera ecuación, el término del lado derecho equivale al valor promedio del variograma entre la muestra 1 y el área desconocida A. En el lado izquierdo de la ecuación existen (n + 1) incógnitas, W_i y λ , y el valor promedio del variograma, entre la muestra 1 y cada una de las otras muestras sucesivamente.

Es decir, cada uno de los términos $\bar{\gamma}$ equivale a la varianza en la estimación de la variable entre el punto de muestra 1 y cada uno de los puntos muestrales, y así sucesivamente con todos los puntos de muestra.

La segunda ecuación es idéntica a la primera, excepto porque en esta la muestra S_2 es la que está presente a lo largo de toda la ecuación. En la tercera

ecuación sucede de igual forma pero con la muestra S_3 , y así sucesivamente hasta la enésima ecuación con S_n , presente en todos los términos de la ecuación. La última ecuación del sistema corresponde al cumplimiento de la condición que la suma de los ponderadores sea igual a uno.

La resolución del sistema de ecuaciones presentado, producirá un grupo de ponderadores, que al remplazarlos, en la expresión final de la varianza nos dará el llamado mejor estimador lineal imparcial Z^* , es decir el valor de la variable que se desconoce en el área de interés. Este proceso es el que ha sido llamado kriging, nombre que se le aplica tanto al sistema de ecuaciones como al estimador producido (Clark I., 1979).

La varianza del estimador de kriging puede ser encontrada, reemplazando los valores de los ponderadores calculados en la siguiente ecuación:

$$\sigma_k^2 = \sum_{i=1}^n W_i \gamma(S, A) + \lambda - \gamma(A, A) \dots \dots \dots Eq.1.30$$

La varianza de kriging σ_k^2 , representa el error en términos de varianza que se ha cometido en el cálculo del estimador Z^* , y su raíz cuadrada representa el error estándar en la estimación.

1.8 GEOESTADÍSTICA MULTIVARIADA (CO- KRIGING)

Los conceptos presentados hasta aquí, extendidos a más de una variable, se denominan Geoestadística Multivariada (Wackernagel, 1995). Es posible encontrar casos de variables de interés que están insuficientemente muestreadas, pero que se conoce su correlación con otras variables en la zona de interés. Utilizando esta correlación es posible estimar una variable de interés a partir de la información de la propia variable además de las correlacionadas con ellas (Journel y Huijbregts,

1978; David, 1977; Myers, 1991a; Wackernagel, 1995; Myers, 1991d; ASCE Task, 1990; Christakos y Bogaert, 1996; Almeida y Journel, 1994; Carr y Mao, 1993). Esto es, el Co-Krigeaje, una extensión o generalización del krigeaje cuando más de una de las variables disponibles guardan relación entre sí. En este caso, se requiere conocimiento no sólo del modelo de semivariograma de cada una de las variables, sino además, del semivariograma cruzado entre las variables (Zhang et al., 1992; Myers, 1991a; D'Agostino y Zelenka, 1992; Pawlowsky et al., 1994; Myers, 1992; ASCE Task, 1990; Myers, 1991a; Carr y Myers, 1990; Wackernagel, 1994). Existen variantes de Co-Krigeaje más generales para la integración de datos (Almeida y Journel, 1994)

En este proceso, se pueden distinguir las siguientes situaciones (Wackernagel, 1995 y 1998):

- **Isotopía:** Se produce cuando todas las variables poseen valores medidos en todas las localizaciones. En este caso no es de interés aplicar el procedimiento multivariado, porque el Co-Krigeaje en este caso puede resultar equivalente al krigeaje, se dice variables autokrigeables.
- **Heterotopía total:** Cuando las variables poseen valores medidos en localizaciones diferentes. En este caso no es de interés tampoco aplicar procedimiento multivariado, además, de que no es posible obtener el semivariograma cruzado experimental.
- **Heterotopía parcial:** Esta situación se produce cuando algunas (la mayor parte) de las localizaciones muestreadas poseen valores medidos de todas las variables, un caso importante es cuando las muestras de la variable de interés están incluidas como un subconjunto de las demás variables. En este caso

pueden ser calculados los semivariogramas cruzados y resulta ventajoso utilizar el procedimiento Co-Krigeaje.

Semivariogramas cruzados

El semivariograma cruzado se obtiene por la ecuación:

$$\gamma_{AB}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z_A(x_i) - Z_A(x_i + h)][Z_B(x_i) - Z_B(x_i + h)] \dots \dots \dots Eq.1.31$$

Donde Z_A y Z_B son variables correlacionadas, Z_A la variable de interés y Z_B la variable auxiliar o secundaria.

Los criterios para el cálculo del semivariograma cruzado son análogos al caso univariado, mientras el semivariograma directo toma sólo valores positivos, el cruzado puede tomar valores negativos, lo que indica correlación inversa entre las variables. Un aspecto importante en el modelado de los semivariogramas cruzados es que deben satisfacer la desigualdad de Cauchy – Schwarz (Wackernagel, 1995):

$$|\gamma_{AB}(h)| \leq \sqrt{\gamma_A(h) \gamma_B(h)} \dots \dots \dots Eq.1.32$$

Una forma de modelar los semivariogramas cruzados consiste en ajustar independientemente los semivariogramas de las variables Z_A , Z_B y el de la suma $Z_A + Z_B$, los cuales están relacionados por la siguiente expresión.

$$\gamma_{AB}(h) = \frac{1}{2} [\gamma_{A+B}(h) - \gamma_A(h) - \gamma_B(h)] \dots \dots \dots Eq.1.33$$

En Issaks y Srivastava (1989) se presentan elementos para el cálculo y ajuste de los semivariogramas en el caso multivariado, además del Modelo Lineal de Corregionalización como procedimiento para modelar semivariogramas directos y cruzados.

- **Co-kriging Simple:** Es el caso más simple de cokriging y la hipótesis básica es la estacionaridad de todas las variables junto con el hecho de que se asume que las medias de todas las variables son conocidas.

Al igual que antes, las condiciones para la optimización son:

- 1) Estimador insesgado: $E(Z^*_{cok}) = E(Z)$
- 2) $Var[Z - Z^*_{cok}]$ mínima

- **Co-kriging ordinario:** Al igual que en el caso de kriging ordinario, se asume que las medias de las variables son desconocidas y se imponen condiciones para filtrarlas.

El estimador propuesto es:

$$Z^*_{cok}(x) = \sum_{\Phi=1}^N \lambda_{\Phi}(x) Z(x_{\Phi}) + \sum_{J=1}^K \sum_{\Phi=1}^N \beta_{\Phi}(x) S_j(x_{\Phi}) \dots \dots \dots Eq.1.34$$

Donde: Z y S son las variables correlacionadas

λ y β son los multiplicadores de Lagrange

En esta tesis se aplicara la técnica de kriging ordinario, por ser esta la que mejor se ajusta a los requerimientos del proyecto y a la interpolación de las variables involucradas.

2. PROGRAMAS DE MODELAMIENTO GEOESTADÍSTICOS – LIBRERÍA GSLIB

2.1 PRINCIPIOS DE LA SIMULACIÓN ESTOCÁSTICA

Se considera la distribución sobre un campo A de uno ó más atributos $Z(u)$, donde $u \in A$. La simulación estocástica es el proceso de construcción de modelos alternativos, probabilísticamente iguales y de alta resolución, de una distribución espacial de $Z(u)$; cada realización es denotada mediante el superíndice $l = \{Z^l(u), u \in A\}$. La simulación es llamada “condicional” si las realizaciones resultantes hacen honor a valores de datos pertenecientes a sus locaciones $Z^l(u_\alpha) = Z(u_\alpha), \forall l$.

La variable $Z(u)$ puede ser categórica, por ejemplo, indicando la presencia o ausencia de un tipo de roca en particular, o puede ser continua tal como lo es la porosidad de un yacimiento.

La simulación difiere del Kriging o de cualquier algoritmo de interpolación en dos aspectos importantes:

- En la mayoría de algoritmos de interpolación, incluyendo el Kriging, la meta es proveer la mejor de todas las respuestas, por lo tanto única, del estimado local de una variable o cualquiera de los componentes de su tendencia sin considerar específicamente las estadísticas espaciales resultantes de los estimados, tomadas juntas. En la simulación, la reproducción de rasgos globales (textura) y estadísticas (histogramas, covarianza), toman precedencia sobre la exactitud local. El Kriging proporciona un conjunto de representaciones locales, tales como $Z^*(u)$, $u \in A$, donde la exactitud local prevalece.

La simulación proporciona representaciones alternativas globales, $Z^l(u), u \in A$, donde la reproducción de patrones de continuidad espacial prevalece.

- A excepción de asumir un modelo Gaussiano para errores, el Kriging provee únicamente una medida incompleta de la exactitud local y no se aprecia exactitud conjunta al considerar varias locaciones juntas. Las simulaciones son consideradas para proveer dichas medidas de exactitud, de ambos tipos, locales o involucrando varias locaciones. Estas medidas son dadas entre valores de simulación alternativas L en cualquier locación (exactitud local) o campos simulados alternativos (globales o exactitud conjunta).

Los diferentes algoritmos de simulación imparten diferencias estadísticas globales y rasgos espaciales en cada realización. Por ejemplo, valores categóricas simulados pueden ser hechos para hacer honor a un modelo de covarianza, así como en simulaciones de relación Gaussiana un acercamiento híbrido podría ser considerado para generar modelos numéricos que reflejen ampliamente los diferentes tipos de rasgos. Por ejemplo, se puede comenzar con un proceso basado en objetos o una simulación de indicador categórica para generar la arquitectura geométrica de varias litofacies, siguiendo con un algoritmo Gaussiano para simular la distribución de propiedades petrofísicas continuas en cada litología por separado, luego un proceso de perturbación continua de una imagen original hasta que se concuerde con algunas características preestablecidas escritas dentro de una función objetiva, podría ser usada para modificar localmente las propiedades petrofísicas para que se ajusten, por decir algo, a los datos de pruebas de pozos.

Ningún algoritmo de simulación por si solo es lo suficientemente flexible para permitir la reproducción de una amplia variedad de rasgos y estadística encontrada en la práctica. Es la responsabilidad de cada usuario, seleccionar un conjunto de algoritmos apropiados y ajustarlos si es necesario.

2.1.1. Representación de heterogeneidades mayores. Las heterogeneidades mayores a gran escala son usualmente las características más consecuentes de los modelos numéricos generados. Por ejemplo, los límites de las facies son con frecuencia una parte crítica para el desempeño del flujo en un yacimiento. Esas heterogeneidades deben ser priorizadas en el proceso de la simulación. Un modelo de función aleatoria $Z(u)$ no puede reproducir heterogeneidades o discontinuidades severas, tales como las encontradas al cruzar un límite físico, tal como la puede ser un litotipo.

Si el fenómeno que está en estudio es una mezcla de diferentes poblaciones físicas y/o estadísticas, la geometría de la mezcla debería ser modelada y simulada primero; luego, los atributos que hacen parte de la población homogénea pueden ser simulados (existen casos en los que ese tipo de secuencia debe ser reversada). Por ejemplo, la arquitectura espacial de las litofacies más representativas en un yacimiento de petróleo debe ser representada primero, a través de una previa selección de la facie/tipo de roca. Luego, puede hacerse la simulación de la distribución de propiedades petrofísicas. Una aproximación de dos pasos a la simulación no es solo más consistente con el fenómeno físico fundamental sino que también evita la extensión de los requerimientos de estacionaridad/homogeneidad fundamentales para la mayoría de los modelos de función aleatoria continua.

Del mismo modo, la reproducción de patrones de continuidad espacial de valores extremos debería ser dada con prioridad, especialmente en situaciones donde

patrones conectados de valores extremos condicionen los resultados obtenidos a partir de los modelos numéricos de simulación. Por ejemplo, la representación geométrica de fracturas intersecando múltiples facies, puede ser más importante que la representación de la geometría de las facies en sí.

2.1.2 Estrategias de búsqueda. La mayoría de los Kriging y los algoritmos de simulación consideran un número limitado de datos condicionales cercanos. La primera razón para esto, es limitar los requerimientos de CPU y memoria. El tiempo de cómputo requerido para solucionar un sistema Kriging se incrementa mientras que el número de datos va al cubo. Los requerimientos de almacenamiento para la matriz del Kriging principal se incrementan mientras el número de datos va al cuadrado.

Por otra parte, si se adopta una vecindad de búsqueda global requerirá conocer la covarianza para la distancia de separación más larga entre los datos. La covarianza es típicamente pobremente conocida para distancias más allá de $\frac{1}{2}$ o $\frac{1}{3}$ del tamaño del campo. Una búsqueda de vecindad local no hace referencia a valores de covarianza fuera de la búsqueda elipsoidal.

Una tercera razón para una vecindad de búsqueda limitada es el permitir el re-escalado local de la media cuando se usa KO. Todos los datos deben haber sido interpretados juntos para establecer un variograma confiable; sin embargo, al momento de la estimación es comúnmente mejor relajar la decisión de la localidad estacionaria y usar solo datos cercanos.

Un número de restricciones son usadas para establecer que datos cercanos deben ser considerados:

- Solo aquellos datos que se encuentran centrados en la búsqueda elipsoidal en la locación estimada son considerados.
- Los datos permisibles pueden ser posteriormente restringidos por un máximo especificado nd_{max} . El datos más cercano, hasta nd_{max} , es retenido. En todos los programas de Kriging y en programas de simulación donde los datos originales son buscados independientemente a partir de nodos simulados, la parte más cercana es medida mediante una distancia Euclidiana. En el algoritmo de simulación secuencial los nodos grid son buscados por distancia de variograma.
- Una búsqueda de octante es una opción a tomar en cuenta para asegurar que los datos son tomados en todos los lados del punto que está siendo estimado. Esto es particularmente importante cuando se trabaja en 3D con datos usualmente alineados junto a las perforaciones. Se especifica este tipo de búsqueda al escoger el número de datos $noct$ a usar en cada octante.

Si muy pocos datos (menos que nd_{min}) son encontrados, entonces la locación del nodo siendo considerado es dejada sin información. Esto restringe la estimación o simulación a áreas donde hay suficientes datos.

Se han implementado tres diferentes algoritmos de búsqueda en diferentes programas:

- Búsqueda exhaustiva: el acercamiento más simple es chequear sistemáticamente todos los datos nd y retener el más cercano nd_{max} para encontrar las tres restricciones encontradas anteriormente. Esta estrategia es ineficiente cuando existen muchos datos y ha sido adoptado en programas de Kriging 2D.

- **Búsqueda de súper bloque:** esta estrategia de búsqueda es un algoritmo eficiente para ser usado en casos cuando se van a estimar muchos puntos, usando vecindades de datos locales, con el mismo conjunto de datos originales. El algoritmo se basa en una clasificación inicial y ordenamiento de los datos de acuerdo a una red regular de bloques paralelepípedos. Esta red es independiente a la de puntos/bloques siendo estimada o simulada. Típicamente, el tamaño de la red de búsqueda es mucho más grande que la estimación o simulación final del espaciamiento del nodo grid.

Cuando se estima un punto cualquiera, solo aquellos datos dentro del superbloque cercano tienen que ser chequeados. Un gran número de datos son entonces eliminados porque han sido clasificados en superbloques más allá de los límites de búsqueda.

La primera tarea es construir una plantilla de súper bloques, centrada en el que contenga el nodo que está siendo estimado. Un segundo truco computacional es ordenar todos los datos por índice numérico del súper bloque. Un arreglo (de tamaño igual al número total de bloques en la red del súper bloque) es construido para almacenar el número acumulativo de datos para cada súper bloque y todos los súper bloques con menores índices de bloques.
- **Búsqueda espiral:** la estrategia espiral es un algoritmo eficiente para casos cuando los datos están (o han sido re-localizados) en un grid regular. Esta idea es para considerar valores en nodos grid más lejanos del punto que está siendo estimado. El ordenamiento de nodos cercanos para chequear es establecido en base a la distancia del variograma. Una plantilla de búsqueda espiral es construida así:

 1. Considerar un nodo grid central en $i_x = 0$, $i_y = 0$, $i_z = 0$.

2. El programa que se base en una búsqueda espiral requiere el tamaño máximo de búsqueda para ser introducido como parámetros de dimensionamiento máximo MAXCTX, MAXCTY, MAXCTZ, los cuales son enteros impares. La máxima distancia de búsqueda, en términos de nodos grid, es entonces $n_x = \text{MAXCTX}/2$, $n_y = \text{MAXCTY}/2$, $n_z = \text{MAXCTZ}/2$.
3. Computar el semivariograma entre el nodo corregido i_x, i_y, i_z y todos los nodos j_x, j_y, j_z , además de una pequeña contribución de la correspondiente distancia Euclidiana

$$\text{dis}(j_x, j_y, j_z) = \gamma(j_x, j_y, j_z) + \varepsilon \cdot h_{j_x, j_y, j_z}$$
4. Ordenar el arreglo $\text{dis}(j_x, j_y, j_z)$ en orden ascendente.
5. Ahora, dado un nodo K_x, K_y, K_z , los nodos más cercanos son encontrados mediante espirales de acuerdo a los órdenes de j_x, j_y, j_z .

Cuando se construye una plantilla de búsqueda espiral, los valores de covarianza para todos los valores j_x, j_y, j_z dentro de la búsqueda elipsoidal son computados y almacenados para ser mas tarde recuperados para construir las matrices del Kriging. Este procedimiento permite hacer mucho más rápido el cómputo de la covarianza cada vez que se necesite.

La herramienta software “SGSIM”, con la cual vamos a trabajar, utiliza este tipo de búsqueda.

2.1.3 Programas de variograma de GSLIB. Las dos subrutinas de variogramas provistas por GSLIB difieren en su capacidad de manejar esquemas de datos regulares e irregulares. Ambos pueden manejar datos en 1D, 2D, 3D. La primera subrutina de variograma *gam* es para datos enmallados. La otra subrutina *gamv* es para datos espaciados irregularmente.

Las subrutinas pueden manejar muchas direcciones diferentes, variables y tipos de variograma en una corrida simple. Por ejemplo, es posible considerar cuatro diferentes variables, ocho direcciones, todos los cuatro variogramas directos, dos variogramas cruzados y el correlograma, en una sola subrutina.

Esta flexibilidad ha sido alcanzada si se especifican por parte del usuario las variables *tail*, *head*, el tipo de variograma, para cada variograma que se va a computar.

- **Datos enmallados:** cuando los datos están en un grid regular, las direcciones son especificadas dando el número de nodos grid que deben ser cambiados para mover de un nodo grid al nodo más cercano que yace a lo largo del vector direccional.

Con la subrutina *gam* las direcciones son especificadas en todos los ejes, el azimut y el ángulo de inclinación. No se permiten tolerancias de dirección o de lag. En algunos casos esta definición de dirección es muy restrictiva y puede resultar en que no haya suficientes pares siendo combinados. La solución es promediar direcciones múltiples de *gam* con sus pesos mediante el número de pares, o almacenar el grid como datos irregularmente espaciado y usar *gamv* para completar los variogramas (a costa de mayor tiempo computacional).

- **Datos espaciados irregularmente:** los ángulos y las distancias lag son introducidas explícitamente para datos espaciados irregularmente. El ángulo azimut es medido en sentido horario de norte a sur y el ángulo de inclinación medido en grados negativos debajo de la horizontal: Las tolerancias angulares son requeridas para ambos casos, estas pueden sobreponerse, causando que los

pares se reporten a más de una dirección. La unidad de distancia lag y la tolerancia lag son también requeridos.

Por las razones expuestas anteriormente, en este trabajo utilizaremos el “GAMV”. En el último capítulo se mostrará en detalle.

2.1.4 Mapas de variograma VARMAP. Los variogramas son representados tradicionalmente en curvas de 1 D $\gamma(h)$ como una función de una distancia h a lo largo de una dirección particular. Es en ocasiones útil tener una vista global de los valores del variograma en todas las direcciones. El mapa de variograma es una gráfica en 2D de $\gamma(h_1, h_2)$ de un semivariograma de muestra para todos los vectores de separación experimentalmente disponibles $h = (h_1, h_2)$. El valor $\gamma(0) = 0$ grafica en el centro de las figuras. El valor $\gamma(h_1, h_2)$ grafica en escala de grises o colores los valores (h_1, h_2) a partir del centro. LE volumen del variograma es la representación en 3D de $\gamma(h_1, h_2, h_3)$.

Las direcciones de anisotropía son usualmente evidentes a partir de un mapa de variograma. La visualización de un volumen de variograma es hecha mejor usando una rutina gráfica que permita un seccionamiento continuo del volumen a lo largo de cualquiera de los 3 ejes h_1, h_2, h_3 .

2.1.5 La aproximación de la simulación secuencial. La aproximación permite el trazado del valor de una variable $Z(u)$ a partir de su distribución condicional dado el valor de la covarianza más ajustada a la misma locación u . El principio de simulación secuencial es una generalización de esa idea: el condicionamiento es extendido para que incluya todos los datos disponibles dentro de la vecindad de u , incluyendo los datos originales y todos los valores previamente simulados.

Se puede considerar la distribución conjunta de N variables aleatorias Z_i , con N posiblemente muy grande. Las N variables aleatorias Z_i , pueden representar los mismos atributos en los N nodos de un grid denso discretizando el campo A , o pueden representar N atributos diferentes medidas en la misma locación, o representarían una combinación de K atributos diferentes definidos en los N' nodos del grid con $N=K*N'$.

Luego se considera el condicionamiento de estas N variables aleatorias por un grupo de n datos de cualquier tipo simbolizados mediante $I(n)$. El correspondiente N variador de la función de distribución acumulativa condicional (ccdf) es:

$$F_{(N)}(z_1, \dots, z_N | (n)) = \text{Prob} \{ Z_i \leq z_i, i = 1, \dots, N | (n) \} \dots \dots \dots \text{Eq.2.1}$$

Esta expresión es completamente general sin ninguna limitación; algunos o todas las variables Z_i pueden ser categóricas. La aplicación sucesiva de la relación de probabilidad condicional da a conocer que delinear una muestra variada N a partir de *ccdf*, puede ser hecho en N pasos sucesivo, cada uno de ellos involucrando un *ccdf* univariado con nivel incremental de condicionamiento:

- Delinear un valor $Z_1^{(l)}$ a partir de *ccdf* univariado de Z_1 , dado el dato original (n) . El valor $Z_1^{(l)}$ es considerado ahora como un datum condicionador para todos los delineamientos subsecuentes; así, los conjuntos de información (n) es actualizado a $(n+1) = (n) \cup \{ Z_1 = z_1^{(l)} \}$.
- Delinear un valor $Z_2^{(l)}$ a partir de un *ccdf* univariado Z_2 , dado el grupo de datos actualizados $(n+1)$, luego se actualiza el conjunto de información a $(n+2) = (n+1) \cup \{ Z_2 = z_2^{(l)} \}$.
- Secuencialmente, se considera a todos las N variables aleatorias como Z_i .

El conjunto $\{z_i^{(l)}, i = 1, \dots, N\}$ representa una realización conjunta simulada de las N variables aleatorias Z_i . Si otra realización es necesaria, $\{z_i^{(l)}, i = 1, \dots, N\}$, se repite todo el proceso de delineamiento secuencial. Este procedimiento requiere la determinación de N *ccdf*'s univariados.

$$\text{Pr ob} \{Z_1 \leq z_1 | (n)\} \dots \dots \dots \text{Eq.2.2}$$

$$\text{Pr ob} \{Z_2 \leq z_2 | (n+1)\}$$

$$\text{Pr ob} \{Z_3 \leq z_3 | (n+2)\}$$

Este principio de simulación secuencial es independiente del algoritmo usado para establecer la secuencia en la Eq.2.2 de *ccdf* univariados. En el programa *sgsim*, todos los *ccdf* son asumidos de modo Gaussiano y sus medias y varianzas son dadas por unas series de N sistemas de Kriging simple.

En la mayoría de aplicaciones de las N variables aleatorias cuyos *ccdf* N -variadas van a ser reproducidas, corresponden a N locaciones u_i de una función aleatoria $Z(u)$ relacionada a un atributo simple z . El algoritmo de simulación secuencial requiere el conocimiento de las N *ccdf* univariadas. Excepto, si $Z(u)$ es modelado como una función aleatoria Gaussiana, estos *ccdf* univariados no son Gaussianos ni completamente determinados por sus únicas medias y varianzas. Afortunadamente, si solo se busca la representación de la covarianza- z (o variograma) por medio del conjunto de datos simulados $\{z^{(l)}(u_i), i = 1, \dots, N\}, \forall l$, basta con que cada uno de los N *ccdf* univariados en la secuencia de la Eq. 2.2 identifiquen la media y varianza del Kriging simple de la correspondiente variable aleatoria $Z(u_i)$ dado el dato considerado ($n+i-1$).

El tipo de *ccdf* de cualquiera de estos N *ccdf* univariados, pueden ser cualquiera, por ejemplo, puede ser una distribución de 3 parámetros, donde el tercero está siendo usado para impartir a los valores simulados $z^{(i)}(u_i)$, las propiedades más allá del modelo covarianza-z.

2.1.6 Consideraciones para la implementación. Las siguientes son las diferentes consideraciones necesarias para la implementación de este tipo de algoritmos:

- La aplicación estricta del principio de simulación secuencial es necesaria para la determinación de *ccdf* más y más complejos, ya que el tamaño del conjunto de datos condicionales se incrementa desde (n) a $(n+N-1)$. En la práctica, el argumento que el dato más cercano afecta la influencia de los datos más remotos; por lo tanto, los datos más cercanos son retenidos para condicionar cualquiera de los N *ccdf* existentes. Puesto que el número de valores simulados previamente puede llegar a ser abrumador ya que i aumenta progresivamente desde 1 hasta $N \gg n$, aún cuando estos sean remotos.
- La limitación de la vecindad de los datos condicionales ocasiona que las propiedades estadísticas del conjunto $(N+n)$ de variables aleatoria serán reproducidas solo hasta la máxima distancia de muestra encontrada en la vecindad. Por ejemplo, la búsqueda debe extenderse por lo menos tan lejos como distancia al cual el variograma es generado; esto requiere condicionamiento extensivo mientras que la secuencia avanza de 1 a N .
- La teoría no especifica la secuencia en la cual los N nodos deberían ser simulados. En la práctica, sin embargo, es mejor hacer una consideración de una secuencia aleatoria. Además, sin los N nodos son recorridos, por ejemplo, a través

de todas las filas, cualquier inicio a partir de teoría rigurosa puede ocasionar un esparcimiento correspondiente a artificios junto a las filas.

2.2 ALGORITMO DE TIPO GAUSSIANO

El modelo de función aleatoria Gaussiana es único en estadística debido a su extrema simplicidad analítica y por el hecho de ser el distribuidor límite de muchos teoremas analíticos globalmente conocido como “Teorema de Límite Central”

En pocas palabras, si el fenómeno espacial continuo $\{z(u), u \in A\}$ es generado por la suma de un número (no muy grande) de fuentes independientes $\{y_k(u), u \in A\}, k = 1, \dots, K$, con distribuciones espaciales similares, entonces su distribución espacial puede ser modelada por un modelo de función aleatoria

Gaussiana Multivariada $Z(u) = \sum_{k=1}^K Y_k(u) \approx \text{Gaussiana}$.

La restricción limitante no es el número K , o el hecho de que los componentes $Y_k(u)$ están igualmente distribuidos, pero si la hipótesis de independencia de los $Y_k(u)$.

Si los errores humanos y de mediciones pueden algunas veces ser considerados como eventos o procesos independientes, en las ciencias de la tierra los diversos procesos geológicos/biológicos que han generado el fenómeno observado, son raramente independientes uno del otro, ni son aditivos. No obstante, los modelos Gaussianos multivariados son extremadamente afines, bien entendidos y tienen un record establecido de aplicaciones exitosas. Estas consideraciones heurísticas son suficientes para convertir al modelo Gaussiano en la elección privilegiada para el modelamiento de variables continuas, a menos que se pruebe lo contrario.

La definición tradicional de una distribución Gaussiana (o normal) multivariada como es expresada a través de su expresión matricial de su multivariada puede

ser encontrada en muchos textos, pero ese término es de muy poco uso práctico. Es mejor definir este modelo a través de sus propiedades características.

La función aleatoria $\{Y(u) = Y(u), u \in A\}$ es normal multivariada si y solo si:

- Todos los subgrupos de la función aleatoria, por ejemplo $\{Y(u), u \in B \in A\}$, son también normales multivariados.
- Todas las combinaciones lineales de las variables aleatorias componentes de $Y(u)$ son distribuidas normalmente

$$X = \sum_{\alpha=1}^n \omega_{\alpha} Y(u_{\alpha}) \dots \dots \dots Eq.2.3$$

$\forall n, \forall$ los pesos ω_{α} , mientras que $u_{\alpha} \in A$.

- Un valor de covarianza cero, ocasiona independencia total

$$\text{Si } Cov\{Y(u), Y'(u)\} = 0 \dots \dots \dots Eq.2.4$$

Las dos variables aleatorias $Y(u)$ y $Y'(u)$, no están simplemente sin correlacionar, sino que también son independientes.

- Todas las distribuciones condicionales de cualquier subgrupo de la función aleatoria $Y(u)$, dadas las realizaciones de cualquier otro subgrupo, son normales (multivariados). Por ejemplo, la distribución condicional de las K variables aleatorias $\{Y_k(u'_k), k = 1, \dots, K, u'_k \in A\}$, dadas las realizaciones $y(u_{\alpha}) = y_{\alpha}, \alpha = 1, \dots, n$, es normal variada, $\forall K, \forall u'_k, \forall n, \forall u_{\alpha}, \forall y_{\alpha}$.

El caso de que $K=1, U_i=U_0$, donde los modelos de variable aleatoria $Y(U_0)$, la incertidumbre acerca de un valor sin muestrear específico $Y(U_0)$ es de interés particular: el "ccdf" de $Y(U_0)$, dados los n datos Y_{α} , es normal y completamente caracterizado por:

- Su media o expectativa condicional identificado para el SK estimado de $Y(U_0)$

$$\begin{aligned}
 E\{Y(u_0) | y(u_\alpha) = y_\alpha, \alpha = 1, \dots, n\} &\equiv [y(u_0)]_{SK}^* \\
 &= m(u_0) + \sum_{\alpha=1}^n \lambda_\alpha [y_\alpha - m(u_\alpha)] \quad \dots\dots Eq.2.5
 \end{aligned}$$

Donde $m(u) = E\{Y(u)\}$ es el valor esperado de la variable aleatoria $Y(u)$, no necesariamente estacionaria, los n pesos λ_α son dados por el sistema SK.

$$\sum_{\beta=1}^n \lambda_\beta C(u_\beta, u_\alpha) = C(u_0, u_\alpha), \alpha = 1, \dots, n \quad \dots\dots Eq. 2.6$$

Donde $C(u, u') = Cov\{Y(u), Y(u')\}$ es la covarianza no necesariamente estacionaria, de la función aleatoria $Y(u)$.

- Su varianza o varianza condicional es la varianza del SK:

$$\begin{aligned}
 Var\{Y(u_0) | y(u_\alpha) = y_\alpha, \alpha = 1, \dots, n\} &= \\
 C(u_0, u_0) - \sum_{\alpha=1}^n \lambda_\alpha C(u_0, u_\alpha) &\quad \dots\dots Eq. 2.7
 \end{aligned}$$

Para la simulación estocástica, la normalidad de todos los “ccdf” es un verdadero punto a favor; realmente, la determinación de la secuencia (Eq. 2.2) de “ccdf” sucesivos reduce a resolver la secuencia correspondiente a sistemas de Kriging Simple.

2.2.1 Transformación normal score. La primera condición necesaria para que la función aleatoria estacionaria $Y(u)$ sea Multivariada normal es que su “ccdf” univariado sea normal

$$\text{Pr ob} \{Y(u) \leq y\} = G(y), \forall y \dots \dots \dots \text{Eq.2.8}$$

Donde $G(\cdot)$ es la “ccdf” Gaussiana estándar; $Y(u)$ es asumida estandarizada, es decir, con media cero y varianza unitaria.

Desafortunadamente, la mayoría de los datos de las ciencias de la tierra no presentan histogramas Gaussianos simétricos. Esto no es un problema mayor ya que una transformación no lineal puede transformar cualquier “cdf” continuo en cualquier otro “cdf”.

Se toman dos variables aleatorias Z y Y con “cdf” $F_Z(z)$ y $F_Y(y)$ respectivamente. La transformación $Y = \varphi(Z)$ identifica las probabilidades cumulativas correspondientes a los valores p en los intervalos de probabilidad Z y Y :

$$F_Y(y_p) = F_Z(z_p) = p, \forall p \in [0,1]$$

$$y = F_Y^{-1}(F_Z(z)) \dots \dots \dots \text{Eq.2.9}$$

Con $F_Y^{-1}(\cdot)$ siendo el “cdf” inverso o función quantile de la variable aleatoria Y :

$$y_p = F_Y^{-1}(p), \forall p \in [0,1].$$

Si Y es una normal estándar con “cdf” $F_Y(y) = G(y)$, la transformada $G^{-1}(F_Z(\cdot))$ es la transformada Normal Score. En la práctica, los n datos de muestras z son ordenados en valores ascendentes $z^{(1)} \leq z^{(2)} \leq \dots \leq z^{(n)}$.

La frecuencia cumulativa correspondiente al valor K -ésimo más grande del datum

$$z \text{ es } F_Z(z^{(k)}) = k/n \text{ o } F_Z(z^{(k)}) = \sum_{j=1}^k \omega_j \in [0,1] \text{ si un conjunto de pesos corregidos } \omega_j$$

han sido aplicados a los n datos. Entonces la transformada Normal Score de $z^{(k)}$ es el quantile k/n de la normal estándar cdf.

$$y^{(k)} = G^{-1}\left(\frac{k}{n}\right) \dots \dots \dots Eq.2.10$$

Si la muestra marginal de cdf $F_z(z)$ es muy discontinua debido, por ejemplo, a que existen muy pocos datos, se podría considerar su suavización anterior a la transformación Normal Score, lo cual se lograría con el programa *histsmth*, lo que también podría usarse para la transformación en reversa.

Si los pesos son usados para corregir la muestra del cdf $F_z(z)$, es el correspondiente histograma con los pesos usados de los datos de la transformación Normal Score que son normalizados, y no los histogramas del Normal Score sin los datos de pesos.

2.2.2 Simulación gaussiana secuencial. El algoritmo más apropiado para generar realizaciones de campos Gaussianos multivariados es provisto por el principio de simulación secuencial tratado anteriormente en el numeral 2.1.5. Cada variable es simulada secuencialmente de acuerdo a su cdf normalizada completamente caracterizada a través de un sistema de Kriging Simple de tipo (Eq. 2.6). Los datos de condicionamiento consisten en todos los datos originales y todos los valores simulados previamente encontrados dentro de una vecindad de una locación que está siendo simulada.

La simulación condicional de una variable continua $z(u)$ modelada por una función aleatoria estacionaria de tipo Gaussiana, $Z(u)$ procede de la siguiente manera:

- Se determina la univarianza de cdf, $F_z(z)$, representativa en toda el área de estudio y no solo donde los datos de muestra z estén disponibles. La corrección puede ser necesaria si los datos z están localizados preferencialmente; también puede necesitarse una suavización mediante una extrapolación.

- Usando el cdf $F_z(z)$, se efectúa una transformación Normal Score de los datos z a los datos y , con una normalización estándar cdf.
- Se verifica la normalidad de la bivariable de los datos y y del Normal Score. Si no puede ser retenido el modelo Gaussiano multivariado, entonces se consideran modelos alternativos tales como una mezcla de poblaciones Gaussianas o un algoritmo basado en indicador para la simulación estocástica.
- Si un modelo de función aleatoria Gaussiana multivariable puede ser adoptado para la variable y , se procede con el programa *sgsim* y la simulación secuencial:
 1. Se defina un patrón aleatorio que visite cada nodo de la malla (no necesariamente regular) una vez. En cada nodo u , se retiene un número específico de datos condicionadores de vecindad incluyendo los datos originales y los valores y de nodo de malla simuladas previamente.
 2. Se usa el Kriging Simple con el modelo del variograma Normal Score para determinar los parámetros (media y varianza) del ccdf de la función aleatoria Y (i) en un lugar u .
 3. Se determina un valor simulado $y^{(l)}(u)$ a partir de ese ccdf.
 4. Se adiciona el valor simulado $y^{(l)}(u)$ al conjunto de datos.
 5. Se procede al siguiente nodo y se hace un lazo hasta que todos los nodos simulados son simulados.
- Se regresa la transformación de los valores normales simulados $\{y^{(l)}(u), u \in A\}$ a valores de la variable original $\{z^{(l)}(u) = \varphi^{-1}(y^{(l)}(u)), u \in A\}$, donde son usadas interpolaciones y extrapolaciones.

Realizaciones múltiples

Si se desean realizaciones múltiples $\{z^{(l)}(u), u \in A\}, l = 1, \dots, L$, se repite el algoritmo anterior L veces del siguiente modo: el mismo patrón aleatorio que recorre los nodos. En este caso, la configuración de datos, además del sistema de Kriging Simple, son los mismos de una realización a la otra, por lo que solo necesitan ser resueltas solo una vez y reduce el tiempo de simulación.

Kriging Simple o Kriging Ordinario

La decisión previa a la estacionaridad requiere que el Kriging Simple con media igual a cero, sea usado en los pasos anteriormente mencionados del algoritmo de simulación Gaussiana. Si los datos son lo suficientemente abundantes para considerar inferencia de un modelo de función aleatoria no estacionaria, se puede proceder así:

- Considerar un arreglo de medias previas de un Kriging Simple local (diferentes de cero) para ser usado en la expresión de los estimados de KS. Estas, cuentan para hacer el paso al proceso de Kriging Ordinario mediante el aporte de medias no estacionarias locales obtenidas a partir de otra fuente de información.
- Dividir el área en subzonas distintas y considerar para cada una de ellas un diferente modelo de función aleatoria estacionario, lo que implica inferencia de un histograma y un variograma de Normal Score diferentes para cada subzona.
- Considerar un variograma Normal Score estacionario, inferido a partir de todo el conjunto de datos, y una media no estacionaria para $Y(u)$. La media no estacionaria, $E\{Y(u)\}$, en cada locación u , es, implícitamente re-estimada a partir de los datos de vecindad a través de Kriging Ordinario. El resultado de tal re-escalamiento local de la media modelo, es, usualmente, una reproducción más pobre del histograma estacionario Y y del modelo del variograma.

Similarmente, la simulación Gaussiana secuencial con una tendencia externa puede ser implementada mediante el reemplazo del estimado y el sistema del Kriging Simple por un estimado del Kriging con tendencia, donde la Normal Score local es linealmente calibrada al valor colocado de una variación suavizada de una variable secundaria (diferente de z). Ya sea Kriging Ordinario o con tendencia el usado para reemplazar el estimado del Kriging Simple, con todo el rigor, es la varianza del Kriging Simple la que debe ser usada para la varianza del ccdf Gaussiano. Una aproximación consistiría en cada varianza del KO o KT derivada localmente por un factor constante obtenido a partir de una calibración previa.

2.2.3 Truncaciones múltiples de un campo gaussiano. Si solo existen dos categorías ($k=2$), una siendo implementada por la otra, se pueden obtener realizaciones no condicionales de la única variable indicadora $i(u,)$ mediante el truncamiento de realizaciones continuas $\{y^{(l)}(u), u \in A\}$ de un campo aleatorio Gaussiano estándar $Y(u)$:

$$i^{(l)}(u) = 1, \dots \text{si} \dots y^{(l)}(u) \leq y_p$$

$$i^{(l)}(u) = 0, \dots \text{sin o} \dots \dots \dots \text{Eq.2.11}$$

Con $y_p = G^{-1}(p)$, siendo p la proporción deseada de los indicadores iguales a uno:
 $E\{I(u)\} = p$.

Ya que el modelo de función aleatoria Gaussiana está completamente determinada por su covarianza $C_y(h)$, existe una relación uno a uno entre esa covarianza y la covarianza de indicador luego de la truncación al valor de p. Mediante la inversión de la relación

$$\text{Pr ob}\{Y(u) \leq y_p, Y(u+h) \leq y_p\} = p^2 + \frac{1}{2\pi} \int_0^{\arcsin C_y(h)} \exp\left(-\frac{y_p}{1+\sin \theta}\right) d\theta, \text{ se puede terminar}$$

la covarianza $C_y(h)$ del campo aleatorio Gaussiano lo cual resultaría con el variograma indicador requerido.

Como la covarianza $C_y(h)$ es el único parámetro del modelo $Y(u)$, no puede ser usado para reproducir más de una covarianza indicadora. Truncaciones múltiples de las mismas realizaciones $\{y^{(l)}(u), u \in A\}$, a valores de entrada diferentes producirá múltiples indicadores categóricos con las proporciones marginales correctas, pero con covarianzas de indicador y correlaciones cruzadas arbitrariamente controladas por el modelo Gaussiano. Por esto, unas covarianzas de indicador sin control, implican realizaciones con variabilidad espacial sin control.

Las series de valores de salidas que definen las distintas categorías pueden ser hechas variables en el espacio. Esto permite la variación de la proporción de cada categoría, por ejemplo, dentro de una secuencia estratigráfica.

Otra consecuencia de las Truncaciones múltiples de las mismas realizaciones continuas, es que la secuencia espacial de las categorías simuladas es arreglada. Por ejemplo, a lo largo de cualquier dirección se encontrarán secuencias de categorías 1, 2, 3, 4 ó 4, 3, 2, 1. El algoritmo no permite la generación de secuencias discontinuas tales como 1, 4, 3, 2.

Esta restricción es útil en algunos casos, limitante en otros. El origen común de la distribución continua de todas las categorías resulta en una agradable distribución espacial continua.

Por último, ya que la función aleatoria Gaussiana $Y(u)$ que está siendo simulada es continua, el condicionamiento de los datos categóricos requiere aproximación adicional.

2.3 ALGORITMOS DE TIPO INDICADOR

Modelos de función aleatoria indicadores, siendo binarios, son idealmente apropiados para la simulación de variables categóricas controladas por estadística de dos puntos tales como las probabilidades del siguiente tipo:

$$\text{Pr ob}\{I(u) = 1, I(u + h) = 1\} = E\{I(u) \cdot I(u + h)\} \dots\dots\dots \text{Eq.2.12}$$

$$\text{Pr ob}\{I(u) = 1, I(u + h) = 1\} = \text{Pr ob}\{I(u) = 1 | I(u + h) = 1\} \cdot p$$

$I(u)$ es el modelamiento de la función aleatoria, la variable de indicador binaria $i(u)$ se ajusta a 1, si cierta categoría o evento prevalece en la locación u , 0 si no, y $E\{I(u)\} = p$ es la proporción estacionaria. Se nota que la probabilidad de transición anterior es la covarianza de indicador a través de una simulación de indicador estocástica que permite la identificación de una serie de probabilidades de transición de tipo (Eq. 2.12) para diferentes vectores de separación h . La pregunta es si la estadística de dos puntos del tipo (Eq. 2.12) es suficiente para caracterizar la geometría de la categoría $\{u \text{ tal que } i(u)=1\}$.

Si la respuesta a la pregunta anterior es negativa, entonces se necesita condicionar el modelo de función aleatoria indicador $I(u)$ a estadística de orden mayor tal como la de tres puntos:

$$\text{Pr ob}\{I(u) = 1, I(u + h) = 1, I(u + h') = 1\} = E\{I(u) \cdot I(u + h) \cdot I(u + h')\} \dots\dots \text{Eq.2.13}$$

O estadística Multivariada implícita tal como unidades s de la forma y tamaño tal que $i(u) = 1, \forall u \in s$.

El condicionamiento a estadística de variable explícita 3,4,..., K del tipo (Eq.2.13) puede ser obtenida a través de procedimientos “annealing” (muestras de posibles estados de un sistema por medio de selección aleatoria de nuevos parámetros. Al final de la simulación, la colección de datos escogidos aleatoriamente en el espacio de búsqueda entrega la información acerca del mismo) o Kriging de indicador generalizado y ecuaciones normales. El condicionamiento a estadística multipunto implícita, tal como distribuciones de parámetros de forma y tamaño, sea mejor con procesos basados en objetos y otros procesos de punto marcado.

La contribución más importante del formalismo del indicador, es la evaluación directa de probabilidades condicionales como las requeridas por el principio de simulación secuencial. De este modo:

- Si la variable a ser simulada es ya un indicador binario $i(u)$, se ajusta a 1 si la locación u pertenece a la categoría $s \in A$, o cero si es lo contrario; entonces

$$Pr ob \{ I(u) = 1 | (n) \} = E \{ I(u) | (n) \} \dots \dots \dots Eq.2.14$$

- Si la variable $z(u)$ a ser simulada es continua, su *ccdf* puede también ser escrita como una expectativa condicional del indicador:

$$Pr ob \{ Z(u) \leq z | (n) \} = E \{ I(u; z) | (n) \} \dots \dots \dots Eq.2.15$$

En ambos casos, el problema de la evaluación de la probabilidad condicional es planeado dentro de la evaluación de la expectativa condicional de un indicador específico de la variable aleatoria. Esta evaluación se basa en teoría de regresión bien establecida, por ejemplo, Kriging. Este aplicado al indicador apropiado de la variable aleatoria provee el mejor estimado de la expectativa condicional del indicador, por esto, se dan las probabilidades condicionales (Eq. 2.14 y Eq. 2.15).

2.3.1 Simulación de variables categóricas. Se puede considerar la distribución espacial de K categorías exclusivas mutuamente $s_k, k=1, \dots, K$. Esta lista es además exhaustiva, por ejemplo, cualquier locación u pertenece a una y solo una de estas categorías.

Se refiere a $i(u; s_k)$ como el indicador de clase s_k , ajustado a 1 si $u \in s_k$, o a cero si ocurre lo contrario.

La exclusión mutua y la exhaustividad ocasionan las siguientes relaciones:

$$i(u; s_k) \cdot i(u; s_{k'}) = 0, \forall k \neq k'$$

$$\sum_{k=1}^K i(u; s_k) = 1 \dots \dots \dots Eq.2.16$$

El Kriging directo de la variable indicador $i(u; s_k)$ proporciona un estimado (de hecho, un modelo) para la probabilidad que s_k prevalece en la locación u . Por ejemplo, usando un Kriging de indicador simple:

$$Pr ob^* \{I(u; s_k)\} = 1 | (n) = p_k + \sum_{k=1}^K \lambda_\alpha [I(u; s_k) - p_k] \dots \dots Eq.2.17$$

Donde $p_k = E \{I(u; s_k)\} \in [0, 1]$ es la frecuencia marginal de la categoría inferida s_k . Los pesos λ_α son dados por un sistema de Kriging Simple tipo estacionario tradicional usando la covarianza de indicador de categoría s_k .

En casos donde las proporciones promedio p_k varían localmente, se puede proporcionar explícitamente sistemas de Kriging de indicador simple con proporciones locales de variaciones suavizadas $p_k(u)$ y la relación (Eq. 2.12), o re-estimar estas proporciones a partir de datos indicadores disponibles en la vecindad de la locación u . Tales re-estimaciones locales de p_k , hacen mas aporte para el uso del Kriging Ordinario.

Si la categoría, por decir s_{k_0} , es predominante, su probabilidad de ocurrencia puede ser reducida como el complemento hasta 1 de las otras (K-1) probabilidades

$$\text{Pr ob}^*\{I(u; s_{k_0})\} = 1 - \sum_{k \neq k_0} \lambda_\alpha [I(u; s_k) - p_k] \dots \text{Eq.2.18}$$

- **Simulación Secuencial**

En cada nodo u a lo largo del patrón aleatorio, el Kriging indicador seguido por una corrección de relación de orden, proporciona K probabilidades estimadas $p_k^*(u|\cdot), k = 1, \dots, K$. La información condicionante (\cdot) , consiste en datos originales y valor de indicador simulados previamente para la categoría s_k .

Seguido se define cualquier ordenamiento de las K categorías, por decir 1, 2, ..., K. este ordenamiento define un escalamiento tipo *cdf* del intervalo de probabilidad [0,1]:

$$\left[0, p_1^*(\cdot)\right], (p_1^*(\cdot), p_2^*(\cdot) + p_1^*(\cdot)), \dots, \left(1 - \sum_{k=1}^{K-1} p_k^*(\cdot)\right) \dots \text{Eq.2.19}$$

Se representa un número aleatorio p uniformemente distribuido en [0,1]. El intervalo en el que p cae, determina la categoría simulada en la locación u . Se actualizan todos los grupos de datos de indicadores K con esta nueva información de simulación, y se procede a la siguiente locación u' a lo largo del patrón aleatorio. El ordenamiento arbitrario de las K probabilidades $p_k^*(\cdot)$ no afecta que categoría es representada ni la distribución espacial de las categorías debido a la distribución uniforme de p .

2.3.2 Simulación de variables continuas. La distribución espacial de una variable continua $z(u)$ discretizada en K clases exclusivas mutuamente

$s_k : (z_{k-1}, z_k), k = 1, \dots, K$, puede ser interpretada y simulada como la distribución espacial K indicadores de clase. La pérdida de resolución dentro de la clase puede ser recuperada en parte mediante el uso a priori de distribución dentro de ella, tal como un modelo de potencia.

Una ventaja de considerar la variable continua $z(u)$ como un mosaico de K clases, es la flexibilidad para modelar la distribución espacial de cada clase mediante un variograma de indicador diferente.

Otra ventaja sobre la mayoría de otras aproximaciones a estimación y simulación es la posibilidad del acceso y manejo de información apropiada.

Un modelo simple de función aleatoria Gaussiana no tiene la flexibilidad, por decir en suficientes parámetros libres, para manejar una mezcla de poblaciones o cuentas para información apropiada. Las mezclas de modelos Gaussianos con diferentes covarianzas, por ejemplo, una por litotipo, requiere previos datos categóricos para separar los tipos de facies: tales datos categóricos pueden no estar disponibles a la escala requerida. El uso de programación cuadrática para incluir información apropiada, tal como restricciones inequitativas, es engorroso, de largo uso de CPU y proporciona únicamente una solución parcial. Por ejemplo, este tipo de programación usado para reforzar la restricción activa $z(u_\alpha) \in (a_\alpha, b_\alpha)$, limita la solución a los límites del intervalo, a_α o b_α , en vez de permitir una solución dentro del intervalo mismo.

- **Indicadores de Clase Cumulativo**

Como oposición a las poblaciones verdaderamente categóricas, cuyas necesidades no tienen orden alguno, las clases $(z_{k-1}, z_k), k = 1, \dots, K$, de una variable

continua $z(u)$ son ordenadas secuencialmente. Por esta razón, es mejor caracterizar las clases con indicadores de clase acumulativo

$$i(u; z_k) = 1, \text{ si } z(u) \leq z_k$$

$$i(u; z_k) = 0 \dots \dots \dots \text{Eq.2.20}$$

El tipo (z_{k-1}, z_k) es definido por el producto $i(u; z_k)[1 - i(u; z_{k-1})] = 1$. A excepción de la primera y última clase, la inferencia de los variogramas de indicadores acumulativos es más fácil que la inferencia de variogramas indicadores de clase, particularmente, si estos tienen pequeñas probabilidades marginales. Además, al considerar indicadores acumulativos (tipo *cdf*) en vez de indicadores de clase (tipo *pdf*), permite el uso de datos indicadores que transportan información a través de “cutoffs”; así existe menos pérdida de información. Por último, los indicadores acumulativos están directamente relacionados al *cdf* de variables continuas bajo estudio.

- **Representación de Variogramas**

La simulación a partir de *ccdf* derivado de un indicador, garantiza la representación de los variogramas indicadores para los z_k de entrada considerados, hasta fluctuaciones ergódicas. Esto no garantiza la reproducción del variograma original z , a menos que un completo Co-Kriging de indicador con una muy buena discretización sea implementada (gran número de K de entrada).

Mejor dicho, es el madograma $\gamma_h(h) = E\{|Z(u) - Z(u+h)|\}$, el reproducido ya que esta función estructural es la integral de todos los variogramas de indicador. No existe una razón a priori para dar el privilegio de la representación del variograma sobre el madograma.

Si por alguna razón el variograma tradicional z debe ser representado, se puede usar una opción con Kriging indicador o revertir a un algoritmo de tipo Gaussiano. Debe notarse que con esta última opción lo que se representa es el variograma de la transformación Normal Score de z . Otra opción es pos-procesar la simulación de z mediante “annealing” para imponer el variograma requerido. Se debe tener en cuenta que la discretización de clase finita induce un ruido adicional (efecto nugget).

- **Simulación Secuencial**

En cada nodo u a ser simulado a lo largo del patrón aleatorio, el Kriging indicador (KS o KO) proporciona un modelo ccdf a través de K estimados de probabilidad

$$F^*(u; z_k | (n)) = \Pr ob^* \{ Z(u) \leq z_k | (n) \}, k = 1, \dots, K.$$

La interpolación dentro de las clases proporciona una continuidad para todos los valores de entrada $z \in [Z_{\min}, Z_{\max}]$.

La simulación Monte Carlo de una realización $z^{(l)}(u)$ es obtenida mediante el trazo de un número aleatorio uniforme $p^{(l)} \in [0,1]$ y retomando el quantile $p^{(l)}$ del ccdf:

$$z^{(l)}(u) = F^{*-1}(u; p^{(l)} | (n)) \dots \dots \dots Eq.2.22$$

$$F^*(u; z^{(l)}(u) | (n)) = p^{(l)}$$

El conjunto de datos de indicadores (para todos los z_k) son actualizados a partir del valor simulado $z^{(l)}(u)$, y se produce a la siguiente locación u' a lo largo del patrón aleatorio.

Una vez que todas las locaciones u han sido simulados, se obtiene una imagen estocástica $\{z^{(l)}(u), u \in A\}$. El proceso de simulación secuencial completo con un

nuevo patrón aleatorio puede ser repetido para obtener otra realización independiente $\{z^{(l')}(u), u \in A\}, l' \neq l$.

El simulador de indicador secuencial tanto para variables categóricas como para continuas es implementado en el programa *sisim*, donde se permite una combinación de Kriging Simple y Ordinario. Información de baja relevancia codificada como datos de indicador extraviados o datos previos de probabilidad puede ser considerada, mientras que el mismo indicador de covarianza $C_I(h; z)$ sea usado para los datos claves como para los que no lo son tanto.

2.4 PROGRAMAS DE SIMULACIÓN GAUSSIANA

2.4.1 Simulación gaussiana secuencial *sgsim*. El algoritmo Gaussiano secuencial se explicó anteriormente, donde los parámetros de dimensionamiento *sgsim.inc* deben ser adoptados para cada estudio de simulación. El tamaño del grid deber ser ajustado explícitamente y debe dársele suficiente espacio para la tabla de covarianza de donde se toman los datos. La representación de estructuras de covarianza de gran rango puede ser deficiente si la tabla de toma de datos de la covarianza es muy pequeña.

La librería GSLIB junto con el programa *sgsim* deben ser compilados siguiendo las instrucciones apropiadas y los parámetros requeridos son mostrados y explicados a continuación:

```

1 Parameters for SGSIM
2 *****
3
4 START OF PARAMETERS:
5 ../data/cluster.dat          \file with data
6 1 2 0 3 5 0                 \ columns for X,Y,Z,vr,wt,sec.var.
7 -1.0      1.0e21            \ trimming limits
8 1                             \transform the data (0=no, 1=yes)
9 sgsim.trn                    \ file for output trans table
10 0                             \ consider ref. dist (0=no, 1=yes)
11 histsmth.out                \ file with ref. dist distribution
12 1 2                          \ columns for vr and wt
13 0.0      15.0               \ zmin,zmax(tail extrapolation)
14 1          0.0              \ lower tail option, parameter
15 1          15.0             \ upper tail option, parameter
16 1                             \debugging level: 0,1,2,3
17 sgsim.dbg                    \file for debugging output
18 sgsim.out                    \file for simulation output
19 1                             \number of realizations to generate
20 50      0.5      1.0        \nx,xmn,xsiz
21 50      0.5      1.0        \ny,ymn,ysiz
22 1       0.5      1.0        \nz,zmn,zsiz
23 69069                          \random number seed
24 0      8                       \min and max original data for sim
25 12                             \number of simulated nodes to use
26 1                             \assign data to nodes (0=no, 1=yes)
27 1      3                       \multiple grid search (0=no, 1=yes),num
28 0                             \maximum data per octant (0=not used)
29 10.0   10.0   10.0           \maximum search radii (hmax,hmin,vert)
30 0.0    0.0    0.0           \angles for search ellipsoid
31 4      0.60   1.0           \ktype: 0=SK,1=OK,2=LVM,3=EXDR,4=COLC
32 ../data/ydata.dat           \ file with LVM, EXDR, or COLC variable
33 4                             \ column for secondary variable
34 1      0.1                     \nst, nugget effect
35 1      0.9  0.0  0.0  0.0    \it,cc,ang1,ang2,ang3
36          10.0 10.0 10.0      \a_hmax, a_hmin, a_vert

```

Figura 2.1 Ejemplo de archivo de parámetro para *sgsim*

- Datafl: datos de entrada en un archivo formateado simplificado. Si el dato no existe, se generará una simulación incondicional.
- Icolx, icoly, icolvr e icolsec: son los números de columnas para las coordenadas X, Y y Z, la variable a ser simulada, el peso de los datos y el posible uso de una variable secundaria. Se pueden ajustar el número de columnas de las coordenadas a cero, lo que indica que la simulación es en 2D o 1D. Para pesos de datos iguales, se ajusta *icolwt* a cero.

- Tmin y tmax: son todos los valores que se introducen en el simulador, los que estén por debajo de *tmin* o por encima de *tmax* se ignoran.
- Itrans: si se ajusta a cero, no se realiza ninguna transformación, la variable se asume de una vez normal estándar (los resultados de la simulación permanecerán también invariables). Si es 1, se realiza la transformación.
- Ismooth: ajustada a cero, el histograma de datos, posiblemente con pesos esparcidos, es usado para la transformación; si es 1, ocurre que los datos son transformados de acuerdo a los valores en otro archivo.
- Smthfl: archivo con los valores a usar para la transformación a Normal Score.
- Icolvr e icolwt: son las columnas en *smthfl* para la variable y los pesos (se ajusta a 1 y 2, si *smthfl* es la salida a partir de *histsmth*).
- Zmin y Zmax: mínimos y máximos valores de datos permitidos. Son usados en el procedimiento de transformación inversa.
- Ltail y Lpar: especifica la implementación de la transformación inversa en el punto más bajo de la distribución; Ltail = 1, implementa una interpolación lineal al límite menor Zmin, y Ltail = 2, implementa una interpolación del modelo de potencia con $\omega = Lpar$ al límite inferior Zmin.
- Utail y Upar: especifican la implementación de la transformación inversa en el extremo superior de la distribución; Utail = 1, implementa una interpolación lineal al límite superior Zmax, y Utail = 2, implementa una interpolación de modelo de potencia con $\omega = utpar$ al límite superior Zmax; y utail = 4, implementa una extrapolación de modelo hiperbólico con $\omega = utpar$.
- Idbg: es el nivel de “debugging” entero entre 0 y 3. Entre más alta sea, más información saldrá escrita.
- Dbgfl: es el archivo para la salida del “debugging”.

- Outfl: es el archivo de salida donde se almacena el grid de salida. Este archivo contendrá los resultados, ciclo más rápido en X, luego en Y, luego en Z y por último, simulación por simulación.
- Nsim: es el número de simulaciones a generar.
- Nx, Xmn, Xsiz: es la definición del sistema del enmallado (eje x).
- Ny, Ymn, Ysiz: es la definición del sistema del enmallado (eje y).
- Nz, Zmn, Zsiz: es la definición del sistema del enmallado (eje z).
- Ndmin y Ndmax: número mínimo y máximo de datos originales que deben ser usados para simular un nodo grid. Si existen menos datos que Ndmin, el nodo no es simulado.
- Mcnode: el número máximo de nodos previamente simulados para usar en la simulación de otro nodo.
- Sstart: si se ajusta a cero, los datos y los nodos grid previamente simulados son buscados por separado, los datos mediante una búsqueda de súper bloque y los nodos en una búsqueda en espiral. Si es ajustado a 1, los datos se asignan a nodos grid y se usa una búsqueda en espiral, donde no se consideran los valores *ndmin* y *ndmax*.
- Multigrid: se realiza una simulación de grid múltiple si se ajusta en 1, si es lo contrario, se hace una búsqueda espiral estándar para nodos previamente simulados.
- Nmuilt: número de refinamientos de grid múltiple a considerar, solo si *multigrid* se ha ajustado a 1.
- Noct: número de datos originales a usar por octante. Si es ajustado a 0, entonces no se usa; por el contrario, este sobrepasa al parámetro *ndmax* y los datos son divididos en octantes; y el dato *oct* más cercano en cada octante es utilizado para la simulación del nodo grid.
- Radius_{hmax}, Radius_{hmin}, Radius_{vert}: búsqueda en la máxima dirección horizontal, mínima dirección horizontal y dirección vertical.

- Sang1, Sang2, Sang3: los parámetros de ángulo que describen la orientación de la búsqueda elipsoidal.
- Ktype: tipo de Kriging (0 = KS, 1 = KO, 2 = KS con media variable localmente, 3 = K con tendencia externa, 4 = CoK) usado mediante loops a través de todos los nodos. Por teoría se requiere el KS, solo en casos donde el número de datos originales encontrados es lo suficientemente grande puede usarse KO sin correr el riesgo del esparcimiento de los valores de los datos más allá de su rango de influencia.
- Rho y Varred: coeficiente de correlación y factor de reducción de varianza para el uso de CoK (solo usado si Ktype = 4). El factor de reducción de varianza modifica la varianza del Kriging luego del CoK. Por defecto, deber ser 1, sin embargo, dependiendo en la continuidad de la realización de la variable secundaria, la varianza del modelo resultante puede ser muy alta. Un factor de reducción de alrededor de 0,6 podría ser requerido. Actualmente se está estudiando un proceso mejor entendible y más automático de este parámetro.
- Secfl: el archivo para la media variante localmente, la variable de tendencia externa o la variable para el CoK.
- Nst o C0: número de estructuras de semivariogramas y la constante de nugget anisotrópica.
- Para cada estructura **nst**, se debe definir **it**, el tipo de estructura; **cc**, el parámetro **c**; **ang1**, **ang2**, **ang3**, los ángulos que definen la anisotropía geométrica; **aa_{max}**, rango horizontal máximo; **aa_{hmin}**, rango horizontal mínimo; y **aa_{vert}**, rango vertical. El modelo del semivariograma es aquel proveniente del Normal Score. La varianza del Kriging es directamente interpretada como la varianza de la distribución condicional; consecuentemente, la constante nugget **c0** y los parámetros **c** (sill), deben ser sumados a 1. Se debe recordar que el modelo de potencia no es un modelo legítimo para un fenómeno multi-Gaussiano.

2.4.2 Truncaciones múltiples de un campo gaussiano gtsim. Como cualquier programa de simulación, deber ser compilado siguiendo ciertas características, a continuación se presentan los parámetros requeridos para el programa principal.

```

1 Parameters for GTSIM
2 *****
3
4 START OF PARAMETERS:
5 sgsim.out          \file with input Gaussian realizations
6 gtsim.out         \file for output categorical realizations
7 1                 \number of realizations
8 50  50  1         \nx,ny,nz
9 3                 \number of categories
10 1  0.25          \ cat(1) global proportion(1)
11 2  0.25          \ cat(2) global proportion(2)
12 3  0.50          \ cat(3) global proportion(3)
13 0                \proportion curves (0=no, 1=yes)
14 propc01.dat      \ file with local proportion (1)
15 1                \ column number for proportion
16 propc02.dat      \ file with local proportion (2)
17 1                \ column number for proportion
18 propc03.dat      \ file with local proportion (3)
19 1                \ column number for proportion
20

```

Figura 2.2 Ejemplo de archivo de parámetro para *gtsim*

- Gausfl: archivo de entrada de datos con las realizaciones Gaussianas que serán truncadas para crear las realizaciones de variable categóricas.
- Outfl: archivo de salida para las realizaciones de variable categórica.
- Nsim: número de realizaciones.
- Nx, Ny y Nz: tamaño del grid de simulación.
- Ncat: número de categorías.
- Cat () y Proportion (): los códigos de enteros para cada categoría y su correspondiente proporción global.
- Pcurve: si se ajusta a cero, no hay curvas de proporción y las proporciones globales son sumadas para establecer las entradas para el truncamiento. Si es puesto en 1, los archivos de curva de proporción son considerados para la entrada de la variación local. Se necesitan archivos de proporción $n_{cat} - 1$.

- Propfl: archivo que contiene un grid regular de 3D (del mismo tamaño del de entrada en las realizaciones Gaussianas) que a su vez incluye las proporciones locales de las categorías *ncat* – 1.
- Icolprop: columna con la proporción de variación local.

2.5 PROGRAMAS DE SIMULACIÓN DE INDICADOR SECUENCIAL

Este programa es para la simulación tanto de variables categóricas de codificación entera, como para variables continuas con datos a indicar a partir de *cdf*. A continuación se explican los parámetros requeridos para el programa principal.

```

1 Parameters for SISIM
2 *****
3
4 START OF PARAMETERS:
5 1 \1=continuous(cdf), 0=categorical(pdf)
6 5 \number thresholds/categories
7 0.5 1.0 2.5 5.0 10.0 \ thresholds / categories
8 0.12 0.29 0.50 0.74 0.88 \ global cdf / pdf
9 ../data/cluster.dat \file with data
10 1 2 0 3 \ columns for X,Y,Z, and variable
11 direct.ik \file with soft indicator input
12 1 2 0 3 4 5 6 7 \ columns for X,Y,Z, and indicators
13 0 \ Markov-Bayes simulation (0=no,1=yes)
14 0.61 0.54 0.56 0.53 0.29 \ calibration B(z) values
15 -1.0e21 1.0e21 \trimming limits
16 0.0 30.0 \minimum and maximum data value
17 1 0.0 \ lower tail option and parameter
18 1 1.0 \ middle option and parameter
19 1 30.0 \ upper tail option and parameter
20 cluster.dat \ file with tabulated values
21 3 0 \ columns for variable, weight
22 0 \debugging level: 0,1,2,3
23 sisim.dbg \file for debugging output
24 sisim.out \file for simulation output
25 1 \number of realizations
26 50 0.5 1.0 \nx,xmn,xsiz
27 50 0.5 1.0 \ny,ymn,ysiz
28 1 1.0 10.0 \nz,zmn,zsiz
29 69069 \random number seed
30 12 \maximum original data for each kriging
31 12 \maximum previous nodes for each kriging
32 1 \maximum soft indicator nodes for kriging
33 1 \assign data to nodes? (0=no,1=yes)
34 0 3 \multiple grid search? (0=no,1=yes), num
35 0 \maximum per octant (0=not used)
36 20.0 20.0 20.0 \maximum search radii

```

```

37 0.0 0.0 0.0 \angles for search ellipsoid
38 0 2.5 \0=full IK, 1=median approx. (cutoff)
39 0 \0=SK, 1=OK
40 1 0.15 \One nst, nugget effect
41 1 0.85 0.0 0.0 0.0 \ it,cc,ang1,ang2,ang3
42 10.0 10.0 10.0 \ a_hmax, a_hmin, a_vert
43 1 0.10 \Two nst, nugget effect
44 1 0.90 0.0 0.0 0.0 \ it,cc,ang1,ang2,ang3
45 10.0 10.0 10.0 \ a_hmax, a_hmin, a_vert
46 1 0.10 \Three nst, nugget effect
47 1 0.90 0.0 0.0 0.0 \ it,cc,ang1,ang2,ang3
48 10.0 10.0 10.0 \ a_hmax, a_hmin, a_vert
49 1 0.10 \Four nst, nugget effect
50 1 0.90 0.0 0.0 0.0 \ it,cc,ang1,ang2,ang3
51 10.0 10.0 10.0 \ a_hmax, a_hmin, a_vert
52 1 0.15 \Five nst, nugget effect
53 1 0.85 0.0 0.0 0.0 \ it,cc,ang1,ang2,ang3
54 10.0 10.0 10.0 \ a_hmax, a_hmin, a_vert
55

```

Figura 2.3 Ejemplo de archivo de parámetro para *sisim*

- Vartype: tipo de variable (1 = continua, 0 = categórica).
- Ncat: numero de entradas o categorías.
- Cat: valores de entrada o códigos de categoría (debe haber *ncat* valores en esta línea de salida).
- Pdf: valores globales del *cdf* o *pdf* (debe haber *ncat* valores en esta línea a de salida).
- Datafl: datos de entrada en un archivo GEO-EAS simplificado.
- Icolx, Icoly, Icolz e Icolvr: número de columnas para las coordenadas X, Y, Z y la variable a ser simulada. Se pueden ajustar una o dos columnas de coordenadas a 0, lo que indica que la simulación es en 2D o 1D.
- Directik: archivo a partir del cual se leen los valores de indicador ya transformados. Los valores extraviados son identificados tan bajos como *tmin*, lo que correspondería al intervalo de restricción. Por el contrario, los datos *cdf* deben crecer constantemente de 0 a 1 y las probabilidades categóricas tienen que estar entre 0 y 1, y además sumar 1.
- Icolsx, Icolsy, Icolsz e Icolli: columnas para las coordenadas X, Y, Z y las variables de indicador.

- $Imbsim$: se ajusta a 1 si se usa el CoK, de lo contrario es 0.
- $B(z)$: si el $Imbsim$ es 1, se necesitan valores de calibración $B(z)$.
- $Tmint$ y $Tmax$: límites de valores por los cuales por debajo o por encima, los demás serán ignorados.
- $Zmin$ y $Zmax$: valor de atributo mínimo y máximo cuando se considera una variable continua.
- $Ltail$ y $Ltpar$: especifica la extrapolación en el extremo inferior: $Ltail = 1$, implementa una interpolación lineal al límite inferior $zmin$; $Ltail = 2$, implementa una interpolación de modelo de potencia con $\omega = Ltpar$ al límite inferior $zmin$; y $Ltail = 3$, implementa una interpolación lineal entre quantiles tabulados (solo para variables continuas)
- $Middle$ y $Midpar$: especifica la interpolación dentro del medio de la distribución: $middle = 1$, implementa una interpolación lineal; $middle = 2$, implementa una interpolación de modelo de potencia con $\omega = midpar$; y $middle = 3$, permite una interpolación lineal entre valores de quantile tabulados (solo para variables continuas).
- $Utail$ y $Utpar$: especifica la extrapolación en el extremo superior de la distribución: $utail = 1$, implementa una interpolación lineal al límite superior $zmax$; $utail = 2$, implementa una interpolación de modelo de potencia, con $\omega = utpar$ al límite superior $zmax$; $utail = 3$ implementa una interpolación lineal entre los quantiles tabulados; y $utail = 4$, implementa una interpolación de modelo hiperbólico con $\omega = utpar$. La extrapolación del extremo hiperbólico está limitado por $zmax$ (solo para variables continuas).
- $Tabfl$: si la opción escogida para cualquiera de las tres regiones, es la interpolación lineal entre valores tabulados, entonces es abierto el archivo de formato Geo-EAS simplificado para cargar los valores. Una opción legítima es exactamente el mismo archivo como los datos condicionales.

- Icolvrt e Icolwtt: los números de columnas para los valores y los pesos de los mismos en *tabfl*. Se puede anotar que los pesos pueden ser usados pero no son requeridos, simplemente se ajusta el número de la columna menor o igual a cero.
- ldbg: nivel de “debugging” entre 0 y 3. Entre más alto sea, mayor cantidad de información saldrá.
- Dbgfl: archivo para la salida del debugging.
- Outfl: archivo donde se escribe el grid de salida, contendrá resultados, rapidez del ciclo en X, luego en Y, luego en Z y por último simulación por simulación.
- Nsim: número de simulaciones a generar.
- Nx, Xmn, Xsiz: es la definición del sistema del enmallado (eje x).
- Ny, Ymn, Ysiz: es la definición del sistema del enmallado (eje y).
- Nz, Zmn, Zsiz: es la definición del sistema del enmallado (eje z).
- Seed: semilla del número aleatoria (número entero, impar y grande).
- Ndmax: número máximo de datos originales que serán usados para simular un nodo grid.
- Ncnode: número máximo de nodos previamente simulados para usar en la simulación de otro nodo.
- Maxsec: número máximo de “soft data” (en locaciones de nodo) que serán usados para la simulación de un nodo. Este es particularmente útil para restringir el número de “soft data” cuando una variable secundaria exhaustiva informa todos los nodos del grid.
- Sstrat: si se ajusta a cero, los datos y los nodos del grid simulados previamente son buscados por separado, los datos mediante una búsqueda de súper bloque y los nodos en una búsqueda en espiral. Si es ajustado a 1, los datos se asignan a nodos grid y se usa una búsqueda en espiral, donde no se consideran los valores *ndmin* y *ndmax*.

- Multigrid: se realizará una simulación de grid múltiple si se ajusta este valor en 1 (por el contrario, se considerará una búsqueda espiral estándar).
- Nmult: número de refinamientos de grid múltiple a considerar, solo si *multigrid* se ha ajustado a 1.
- Noct: número de datos originales a usar por octante. Si es ajustado a 0, entonces no se usa; por el contrario, el dato *oct* más cercano en cada octante es utilizado para la simulación del nodo grid.
- Radius_{hmax}, Radius_{hmin}, Radius_{vert}: búsqueda en la máxima dirección horizontal, mínima dirección horizontal y dirección vertical.
- Sang1, Sang2, Sang3: los parámetros de ángulo que describen la orientación de la búsqueda elipsoidal.
- Mik, Mikcat: si se ajusta *mik* a 0, entonces se realiza un Kriging indicador completo en cada locación del nodo del grid para establecer una distribución condicional. Si se ajusta *mik* a 1, entonces se usa la aproximación media, es decir, un variograma simple para todas las categorías. De otro modo, solo se necesita que se resuelva un sistema Kriging y el tiempo de cómputo se reduce significativamente. El variograma correspondiente a la categoría *mikcat* será usado.
- Ktype: el tipo de Kriging (0 = KS, 2 = KO) usado a través del loop por todos los nodos. Por teoría se requiere el KS, solo en casos donde el número de datos originales encontrados es lo suficientemente grande puede usarse KO si correr el riesgo del esparcimiento de los valores de los datos más allá de su rango de influencia. Los valores globales *pdf* (especificados con cada categoría) son usado para KS.
- Nst o C0: número de estructuras de semivariogramas y la constante de nugget anisotrópica.
- Para cada estructura **nst**, se debe definir **it**, el tipo de estructura; **cc**, el parámetro **c**; **ang1**, **ang2**, **ang3**, los ángulos que definen la anisotropía

geométrica; aa_{max} , rango horizontal máximo; aa_{hmin} , rango horizontal mínimo; y aa_{vert} , rango vertical.

2.6 PROGRAMAS ÚTILES

Se presentan programas de utilidad que parten de los algoritmos geoestadísticos, sin embargo son útiles en la producción de gráficas y figuras.

Los generados por lenguaje de descripción en una página de visualización, son gráficos elementales de aproximadamente 3”*4”, donde pueden ser pre-visualizados o inclusive impresos, opciones que no se tenían en versiones anteriores del programa.

Los comandos de visualización para los encabezados de la gráfica están contenidos en archivos de plantilla nombrados indiferentemente, de donde son leídos mediante un editor.

La motivación detrás de estas muestras del visualizador es dar al usuario una forma de obtener rápidos sumarios gráficos y estadísticos de sus datos. Claro que se debe aprender un poco acerca del lenguaje de descripción del visualizador si se desea adaptar el programa de salida.

- **Mapas de escalas de grises y colores *pixelplt***. para hacer gráficos de cortes en 2D a partir de conjuntos de datos en 3D, ya sea generado por un Kriging o algoritmos de simulación, se pueden usar mapas de grises o escalas de colores, variando simplemente datos del archivo y de entrada. A continuación se muestra el código del programa

```

Parameters for PIXELPLT
*****

START OF PARAMET
simulacionnewPOROK2.out          \file with gridded data
1                                \ column number for variable
-1.0e21 1.0e21                    \ data trimming limits
EP_POROK2newCORREGIDA.ps        \file with PostScript output
1                                \realization number
63 0.5 10.0                       \nx,xmn,xsiz
161 0.5 10.0                      \ny,ymn,ysiz
215 0.5 10.0                      \nz,zmn,zsiz
1                                \slice orientation: 1=XY, 2=XZ, 3=YZ
100                               \slice number
POROSIDAD                        \Title
Este                             \X label
Norte                            \Y label
0                                \0=arithmetic, 1=log scaling
1                                \0=gray scale, 1=color scale
0                                \0=continuous, 1=categorical
0.0 0.5 0.05                     \continuous: min, max, increm.
4                                \categorical: number of categories
1 3 Code_One                      \category(), code(), name()
2 1 Code_Two
3 6 Code_Three
4 10 Code_Four

Color Codes for Categorical Variable Plotting:
1=red, 2=orange, 3=yellow, 4=light green, 5=green, 6=light blue,
7=dark blue, 8=violet, 9=white, 10=black, 11=purple, 12=brown,
13=pink, 14=intermediate green, 15=gray

```

Figura 2.4 Ejemplo de archivo de parámetros para pixelplt

- Graficando variogramas *vargplt*:** este programa toma el formato de salida especial usado por los programas del variograma y crea muestras gráficas para equipos de visualización. Este es un programa de muestreo directo y no provee ninguna facilidad para el cálculo del variograma o ajuste de un modelo. Se encuentran ajustes de variogramas interactivos de acceso público. Este es el único programa del GSLIB que trabaja interactivamente. Se requerirá un nombre de archivo que contenga los resultados del variograma (a partir de gam o gamv). El escalamiento de la distancia y los ejes del variograma pueden ser ajustados manual o automáticamente. Se pueden mostrar múltiples variogramas en una sola

gráfica o se puede iniciar un visualizador nuevo en cualquier momento. La siguiente gráfica muestra el código usado para las visualizaciones:

```
Parameters for VARGPLT
*****

START OF PARAMETERS:
vargamplt.ps          -file for PostScript output
2                    -number of variograms to plot
0.0  20.0            -distance limits (from data if max<min)
0.0   1.2            -variogram limits (from data if max<min)
Normal Scores Semivariogram -Title for variogram
gam2.out             -1 file with variogram data
2   1   1   1   1   - variogram #, dash #, pts?, line?, color

Color Codes for Variogram Lines/Points:
1=red, 2=orange, 3=yellow, 4=light green, 5=green, 6=light blue,
7=dark blue, 8=violet, 9=white, 10=black, 11=purple, 12=brown,
13=pink, 14=intermediate green, 15=gray
```

Figura 2.5 Ejemplo de archivo de parámetros para vargplt

- **Valores del variograma a partir de un modelo *vmodel*:** este programa escribe un archivo de salida con el mismo formato del programa *gam*, para que puedan ser graficadas con el *vargplt*. Los usos primarios del *vmodel* son sobreponer un modelo en puntos experimentales y también proveer una utilidad para chequear la definición del modelo del semivariograma.

3. CARACTERÍSTICAS GEOLOGICAS DE LOS AMBIENTES FLUVIALES

Todas las masas de roca tienen algunas características o aspectos que constituyen su estructura, estas estructuras geológicas están relacionadas con todos los accidentes tectónicos de una determinada región y son formados por movimientos epirogénicos y orogénicos.

La estructura es la forma en que han sido depositadas las rocas y de acuerdo a esta forma de depositación podemos inferir los mejores depósitos con características de roca, tales como porosidad y permeabilidad, más favorables para el entrapamiento de hidrocarburos.

Un objetivo en la exploración y desarrollo de yacimientos fluviales es determinar el espesor y el ancho de los cuerpos conglomerados de areniscas (principalmente depósitos de cinturones de canal). La mayoría de las técnicas existentes para determinar las dimensiones de yacimientos fluviales tienen muchas desventajas. Uno de los principales objetivos para estimar una mejor aproximación al problema es hacer uso del conocimiento teórico y estudios de campo. Es importante distinguir la arquitectura del sistema fluvial caracterizando cada uno de los canales, y los cinturones de canal en conjunto. La correlación entre pozos y una apropiada analogía entre los datos de espesor y ancho de canal nos dan un valor potencial para caracterizar el yacimiento fluvial.

3.1. MOVIMIENTO DE PARTICULAS

La capacidad de una corriente es la medida del tamaño de partículas que esta puede cargar. La competencia de una corriente para una carga suspendida es un balance entre las fuerzas de levantamiento que ejercen turbulencia y la fuerza de la gravedad. La competencia es una medida de la capacidad de la corriente para poder transportar partículas de diferentes tamaños. Ver Figura 3.1.

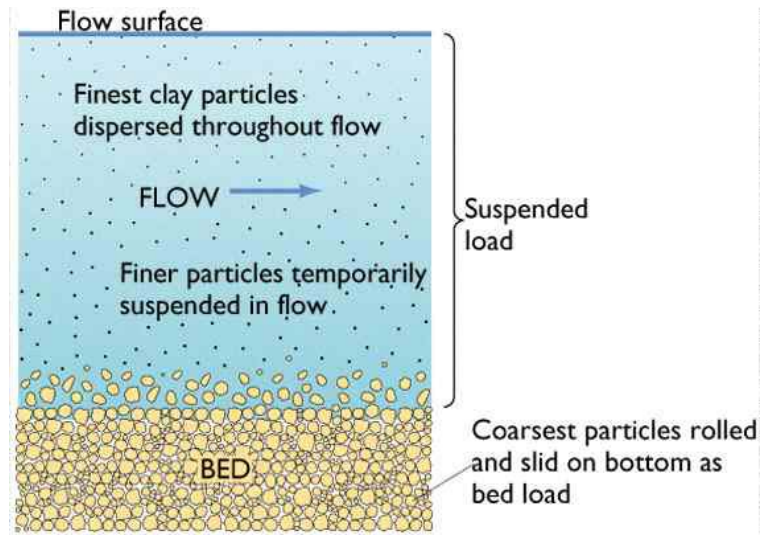


Figura 3.1. Flujo de partículas en una corriente

La energía con que la corriente puede transportar estas partículas se refleja de acuerdo a cantidad de partículas transportadas por el medio. Esto describe un patrón de flujo establecido por la fuerza y velocidad de la corriente. Se puede visualizar imaginariamente el trayecto recorrido por una partícula en una determinada corriente. Existen tres patrones principales de flujo :

- Flujo Laminar: Las partículas se mueven dentro del fluido en capas paralelas
- Flujo Turbulento: Las líneas de flujo se cruzan en un patrón complejo.
- Flujo de pendientes fuertes y suaves: El flujo es determinado por el ángulo de la pendiente. Ver Figura 3.2.

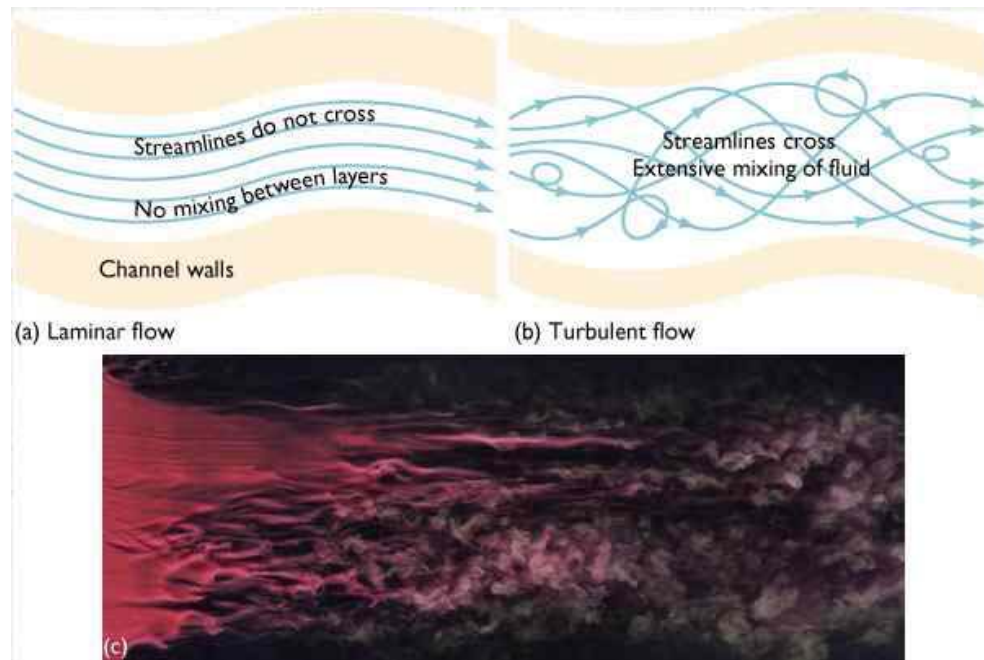


Figura 3.2. Flujo laminar y turbulento

Las tres formas de movimiento de flujo: Laminar, turbulento y de pendiente, difieren en la velocidad y la habilidad de transportar granos de arena y otros sedimentos erodados de las tierras circundantes y del piso del río. Los flujos laminares son suaves y solo pueden transportar partículas más pequeñas. Los flujos turbulentos dependiendo de su velocidad pueden mover partículas de tamaños de *pebbles* a *cobbles*. Las corrientes más rápidas (las de mayor energía) pueden mover partículas más gruesas y a mayores velocidades.

Los tipos de depósitos generados en sistemas fluviales dependen en gran medida de la relación existente entre carga y descarga de partículas. Cuando una corriente de turbulencia decrece, su fuerza de transporte cae y deposita parte de su carga. Cuando una corriente eleva durante un mayor flujo, el agua que sobrepasa el canal inunda el sistema, generando depósitos de granos más finos, conocidos como las llanuras de inundación (floodplainshale).

Estos tipos de depósitos se pueden diferenciar de acuerdo a la forma de depositación y a la capacidad de descarga. Los movimientos más importantes son (Ver Figura 3.3):

- Carga Suspendida (*Suspended Load*).
 - Partículas dispersas
 - Partículas Finas Suspendidas
- Transporte en lecho de roca (*Bed Load*).
 - Partículas Rodadas (*Rolled*)
 - Partículas en Deslizamiento (*Slid*)
 - Saltación (*Saltation*)

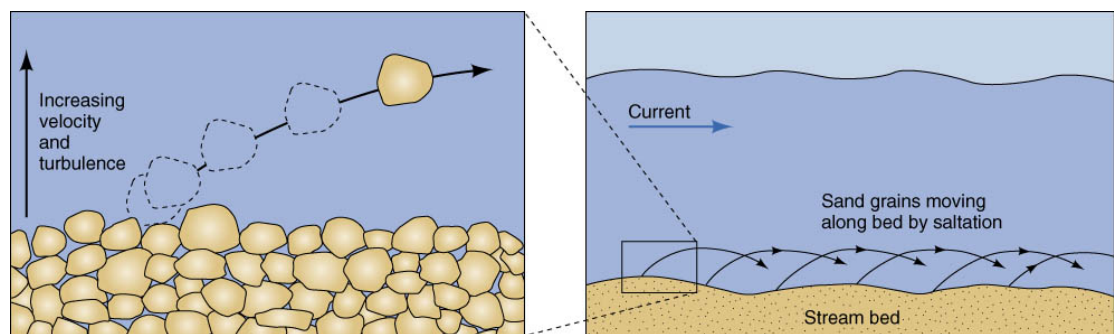


Figura 3.3. Flujo de corriente sobre un lecho de arena, limo y arcilla

El flujo de corriente sobre un lecho de arena, limo y arcilla puede transportar partículas en dos formas: (1) Por saltación, deslizamiento y rodamientos a lo largo del fondo. (2) por carga suspendida

Los aluviones generan una clase de depósito caracterizado por el flujo de pendientes fuertes y suaves. El cambio de pendiente permite una turbidez y por consiguiente una mezcla de partículas en la corriente. Los aluviones están

relacionados con los escombros de erosión que los ríos depositan como sedimento y cubren los pisos de los valles, generalmente en la base de las montañas.

3.2 ESTRUCTURAS SEDIMENTARIAS

Las litofacies en los sistemas de arenas en sistemas fluviales resultan de los procesos de transporte de partículas tales como tracción, transporte en suspensión y saltación. La morfología de los cuerpos de arena formados por estos tipos de depósito depende primariamente de tres parámetros, tamaño del grano de arena, profundidad del flujo y la velocidad del flujo. Otros parámetros tales como viscosidad y temperatura, ejercen un menor control sobre la morfología del lecho de transporte.

El movimiento de los granos de arena crea ondas y dunas en el lecho de la corriente. Las Ondas (*Ripples*) en las arenas son formadas por crestas estrechas separadas por amplios valles, formando superficies onduladas. Los *ripples* migran lateralmente generando estructuras de estratificación cruzada. Esta es bastante útil en la determinación de la dirección de la corriente.

El movimiento en los lechos de la corriente puede ser afectado por una mayor energía generando un movimiento de partículas con mayor velocidad y por lo tanto produciendo depósitos con estructuras de formas de Duna (*Dune*); de mayor altura y pendiente que las ondículas (*ripples*), generadas a pendientes más suaves. Las dunas tienen las mismas estructuras de estratificación cruzada que las ondas.

A más altas velocidades el movimiento de la duna genera pendientes más escarpadas y las partículas en suspensión generan turbulencia afectando el transporte de las partículas del fondo del lecho. Esto genera ondas y dunas dentro de las mismas dunas y por ultimo el movimiento de retroceso forma las estructuras de Antidunas (*Antidune*).



Figura 3.4. Ripples de corriente

3.3. MODELOS DE FACIES FLUVIALES

Una facie es una unidad de sedimento o roca sedimentaria que conserva las mismas características, como textura, estructuras sedimentarias u otras características físicas o químicas (litofacies). Si se encuentra ensamblado un fósil, se puede definir como una biofacie.

Los modelos de facies son usados para interpretar ambientes de depositación. En las facies fluviales los modelos están relacionados con la definición de estructuras sedimentarias y tipos de depósitos asociados a un patrón de flujo.

Examinar y conocer las características de los depósitos y como ellos cambian lateral y verticalmente; comparar observaciones con modelos de facies nos ayudan a comprender sus características dentro del yacimiento.

3.3.1. Patrones de rios de canales. Para el estudio de los patrones de flujo es importante caracterizar las facies asociadas y los diferentes tipos de depósito. Los patrones de canales se refiere a las características de un canal vista en planta. Los depósitos de los ríos varían dependiendo del patrón del canal, entonces se desarrollan diferentes tipos de modelos de facies con diferentes patrones de canal.

- Meandriforme: Un canal, curvado; tiende a formarse en pendientes de bajas a moderadas, en un material de grano fino.
- Trenzado: El canal se divide alrededor de islas, tiende a formarse en pendientes más altas y con material más grueso.
- Lineal: Raros; se forman en muy bajas pendientes.
- Anastomosados.

En la naturaleza se pueden encontrar mezclas asociadas de los diferentes tipos de patrones.

Los depósitos de los ríos consisten de depósitos de canal (overbank) y de llanuras de inundación (floodplain). Los de canal se pueden caracterizar con los siguientes elementos: canales (point bars, channels), Natural levee, Crevasse Splays. El lodo y limo son los depósitos característicos de las llanuras de inundación. Las características de los depósitos de los ríos dependen del patrón del canal.

En la caracterización de yacimientos fluviales, desde el punto de vista geoestadístico, estos tipos de depósitos son considerados de gran importancia, para la definición de modelos de objetos. Estos modelos se caracterizan de acuerdo a su arquitectura fluvial, con el fin de permitir dar identidad a cada objeto el cual puede ser modelado con parámetros estadísticos y funciones matemáticas. Todo esto permite construir estratigráficamente el ambiente de depositación de un determinado yacimiento.

Ríos Trenzados. (*braided*): Cuando el nivel de un río cae rápidamente, la competencia y la capacidad de carga del río se ve disminuida, entonces la carga es depositada. Estos depósitos obstruyen el flujo del río y causan la división de la corriente formando pequeños canales.

Estos tipos de ríos son caracterizados por altas relaciones ancho / espesor, normalmente mayores que 40 y comúnmente excediendo 300. Leopold and Wolmand (1.975) encontraron que ellos tienen pendientes más escarpadas que otros tipos de ríos.

La geometría de una corriente trenzada genera estructuras aisladas como islas. Esto permite crear corrientes separadas como trenzas o pequeños flujos de agua independientes. Si una corriente de esta tiene la capacidad para mover cargas grandes de sedimentos, estas tienden a depositarlos como barras que localmente dividen el flujo. Un patrón trenzado presenta abundancia de material grueso y tiende a formar corrientes que tienen gran capacidad de descarga y a formar bancos fácilmente erodables. Ver Figura 3.5.

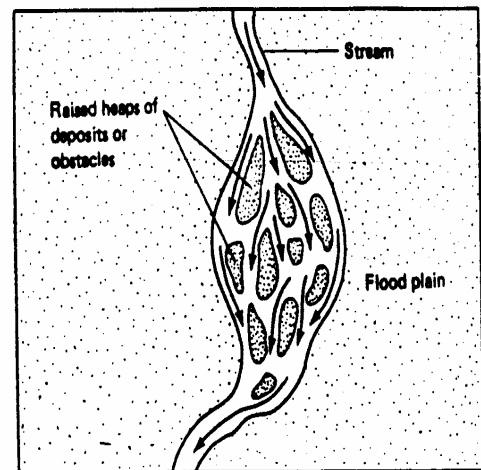


Figura 3.5. Canal trenzado

Ríos Meandriiformes (*meandering*): En contraste con los ríos trenzados, los ríos meandriiformes típicamente contiene un canal que se envuelve a través de su recorrido rompiendo y depositando material encontrado a su paso. Estos son característicos de pendientes más suaves. De acuerdo a la geometría establecida por este tipo de canal, erodará en la parte externa del meandro y depositará en la parte interna del mismo, generando depósitos de tipo point bar (barras puntuales). Ver Figuras 3.6 y 3.7.

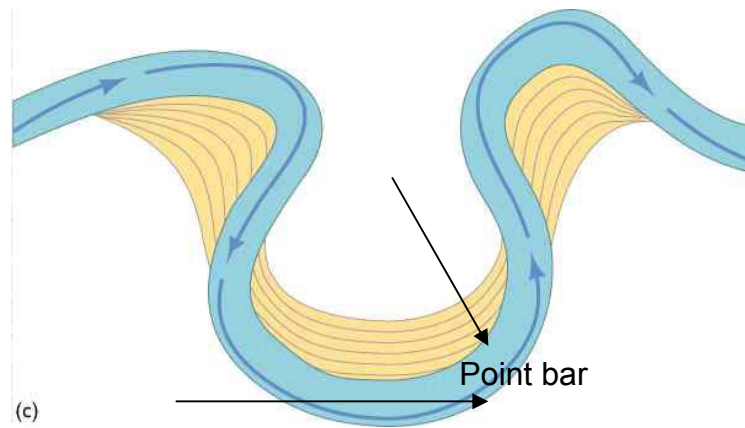


Figura 3.6. Tipo de depósito asociado a un canal meandriiforme

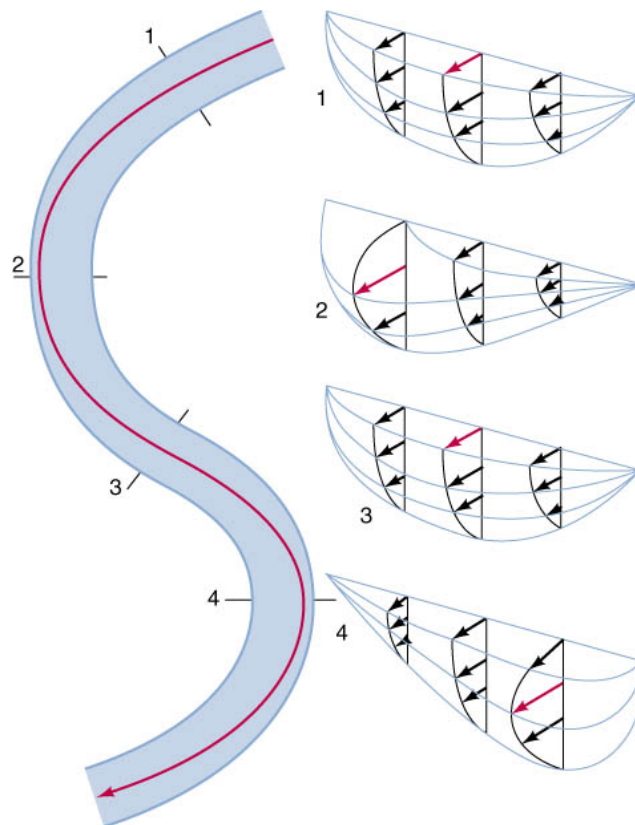


Figura 3.7. Esquema de planta y cortes asociados a la trayectoria de un canal

En el esquema de planta y cortes asociados a la trayectoria de un canal meandriforme. La línea roja muestra la dirección del thalweg

Algunos elementos geométricos importantes a describir en un modelo de facie de tipo meandriforme incluyen principalmente (Ver Figura 3.8):

- Ancho de canal
- Profundidad
- Relación ancho / profundidad.
- Ubicación del centro del canal
- Ubicación de la línea del thalweg.

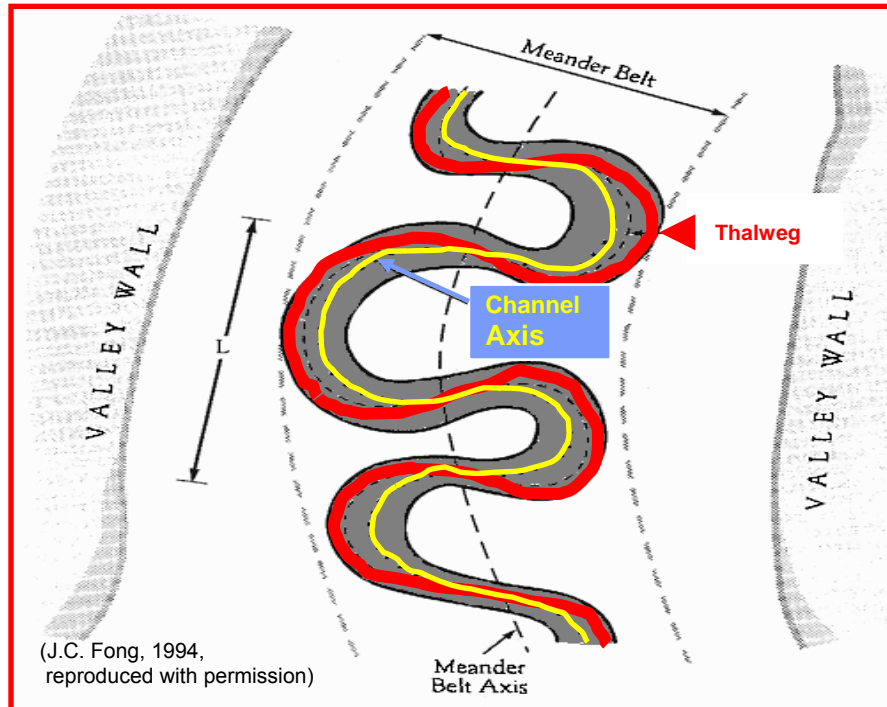


Figura 3.8. Diferencia entre la línea central del canal y la línea del thalweg

La línea central no corresponde con la línea del thalweg en todas las trayectorias del río. La línea del thalweg se define como la unión de los puntos donde el canal presenta mayor profundidad. Está relacionada con la mayor energía de la corriente. En la figura anterior es importante observar la geometría de cada una de las secciones de la trayectoria del río. Para la trayectoria más recta el perfil característico está asociado con geometrías más parabólicas y la línea del thalweg está en el centro del canal. En las trayectorias curvas los cortes muestran un desplazamiento del thalweg hacia la parte externa del meandro.

Generalmente el canal meandriforme tiene relaciones ancho /espesor menores de 40, y para ríos anastomosados la relación puede ser menor que 10. Son caracterizados por una predominancia de material suspendido, bancos cohesivos y pendientes ligeramente suaves.

En la primera etapa de formación de un meandro, la corriente lineal se ve afectada por los obstáculos encontrados creando variaciones en la fuerza de la trayectoria. Esto aumenta la sinuosidad de los meandros incipientes hasta alcanzar superficies cada vez más cóncavas, para luego cortar el meandro en dos puntos de avulsión. El depósito aislado se denomina *oxbow lake* (arco de buey), el cual consta de material fino con vegetación típica característica de zonas de inundación. Ver Figuras 3.9, 3.10 y 3.11.



Figura 3.9. Ejemplo de un canal meandriforme caracterizado por los depósitos de Point bar y Oxbow lake



Figura 3.10. Ejemplo de un meandro que corta verticalmente una sección estratigráfica.

Aunque los ríos meandriformes ocurren en pendientes suaves estos pueden llegar a tener un gran carácter erosivo en pendientes fuertes y creando perfiles característicos como el observado en la figura

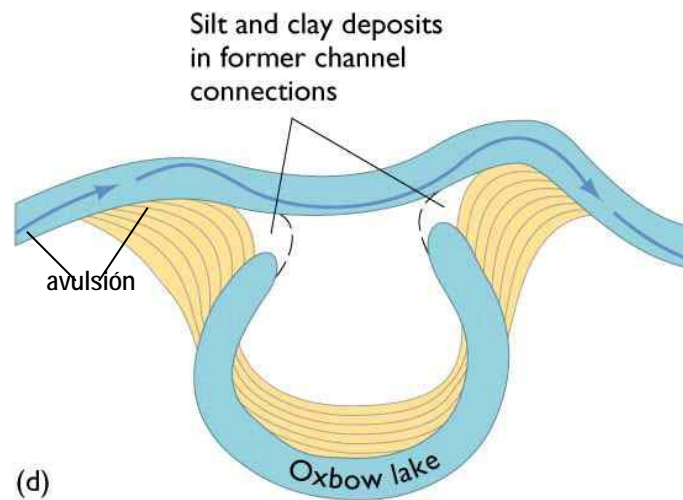


Figura 3.11. Formación de un Oxbow lake en un canal meandriforme

Ríos Anastomosados (*anastomosing*): Los ríos anastomosados son caracterizados por dos o más canales estables que exhiben de baja a alta sinuosidad. Ellos parecen ocurrir en áreas donde hay rápida subsidencia, pero donde las pendientes permanecen relativamente bajas (0,5 %). En este tipo de ríos los canales son más profundos que anchos, generando perfiles principalmente verticales.

3.3.2. Clasificación de canales con base en la sinuosidad. La Figura 3-12 muestra los patrones de canales y perfiles característicos de acuerdo a la pendiente del río.

Sinuosidad	single-channel	multiple-channel
low (<1,5)	Straight	braided
high (>1,5)	meandering	anastomosing

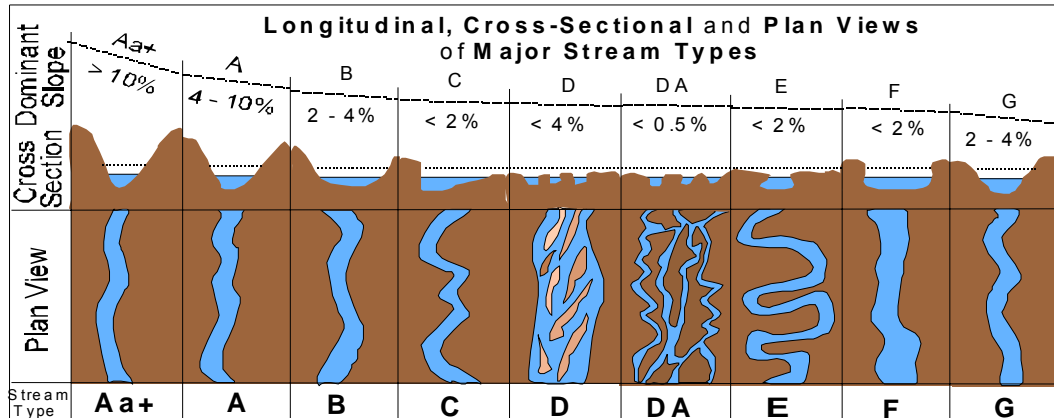


Figura 3.12. Patrones de canales y perfiles característicos de acuerdo a la pendiente del río

3.4. TIPOS DE DEPÓSITO

Cada vez que una corriente de agua puede llevar cierta cantidad de carga, en cualquiera de sus formas, se produce una depositación debido al aumento de dicha carga o por disminución de velocidad de transporte. Esto genera diferentes tipos de depósitos caracterizados por su forma y por el tamaño de grano que lo conforman.

Algunos de los principales tipos de depósitos asociados a canales son (Ver Figura 3.13):

- Depósitos de canal.
- Point bar.
- Natural levee.
- Crevasse Splays.

- FloodPlain Shale.

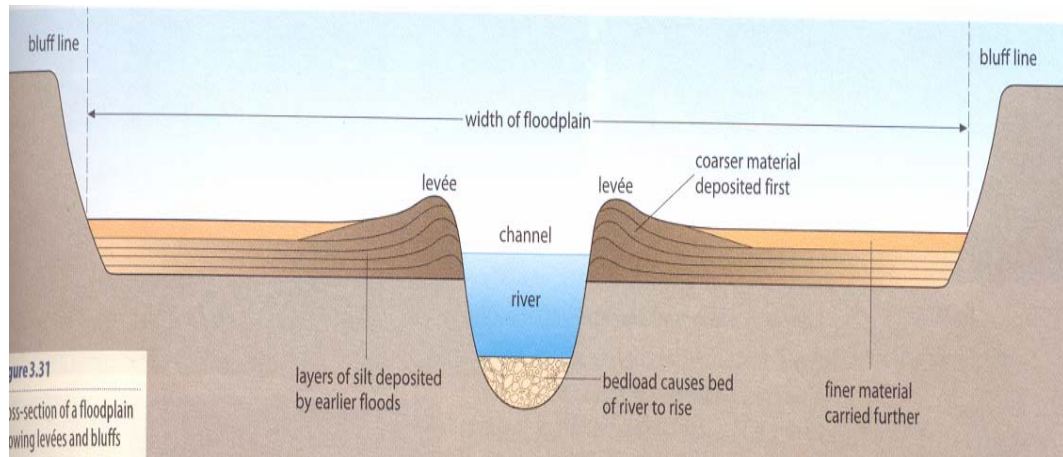


Figura 3.13. Perfil característico de un canal mostrando el canal mismo, el natural levee y la superficie de inundación

3.4.1. Depósitos de canal. Este tipo de facies, llenan canales sinuosos abandonados. Es vista como la de mejor calidad del yacimiento debido a la alta energía de depositación y consecuentemente al tamaño de grano más grande (compuestos la mayoría de veces por arena y grava). Los canales de relleno de canal son geoméricamente definidos por un ancho, espesor máximo y la posición relativa de la línea central del canal.

3.4.2 Formacion de un point bar. Los depósitos de point bar son cuerpos deposicionales formados dentro del perfil del meandro. Las fuertes corrientes disminuyen a lo largo de los point bar, descargando material en la parte interna del meandro, acrecentando el depósito en la dirección del canal. La migración del point bar se manifiesta hacia la parte interna de la curvatura del meandro. Ver Figuras 3.14 y 3.15.

Dicho con más precisión, esos depósitos se describirán mejor como acumulaciones de arena y grava en forma de cuarto creciente.

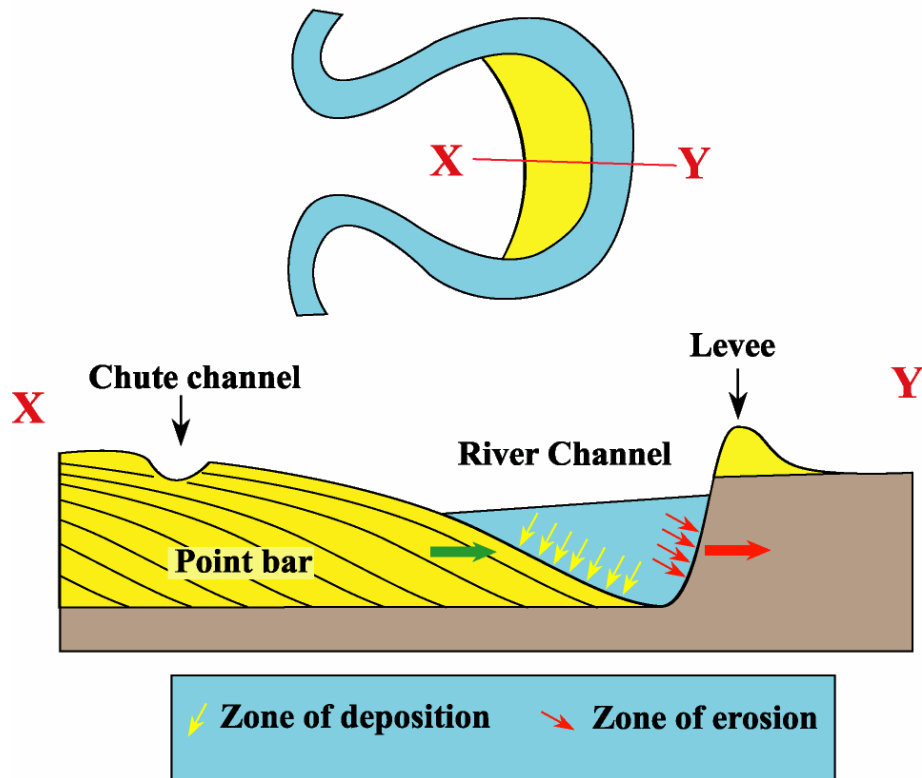


Figura 3.14. Formación de un point bar. Vista de planta y perfil



Figura 3.15. Canal meandriforme mostrando la formación y migración de un point bar

3.4.3. Formación del natural levee. Este tipo de facie es el depósito de borde de canal. Estos depósitos se desarrollan a lo largo y de forma paralela al canal, disminuyendo el tamaño de grano a medida que se separa del canal. Estas arenas son consideradas de menor calidad que las arenas de canal. Los parámetros definidos para el *Natural Levee* son, el ancho w , la altura total s , y la profundidad por debajo del tope del canal k .

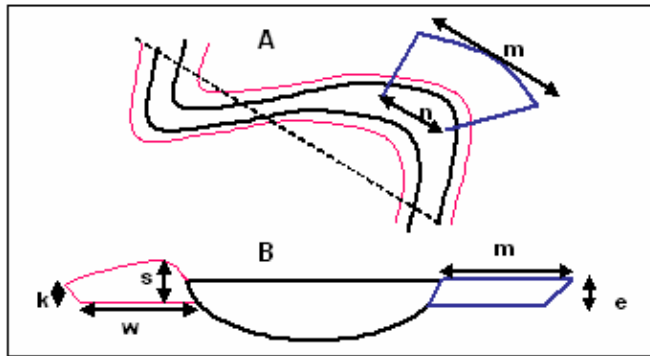


Figura 3.16: (A). Vista de planta . (B) vista de perfil

En la figura anterior se muestran las. Vistas de planta y de perfil del canal, el natural levee y un crevasse splay. de estos tipos de depósitos con sus respectivos elementos geométricos.

The formation of levees

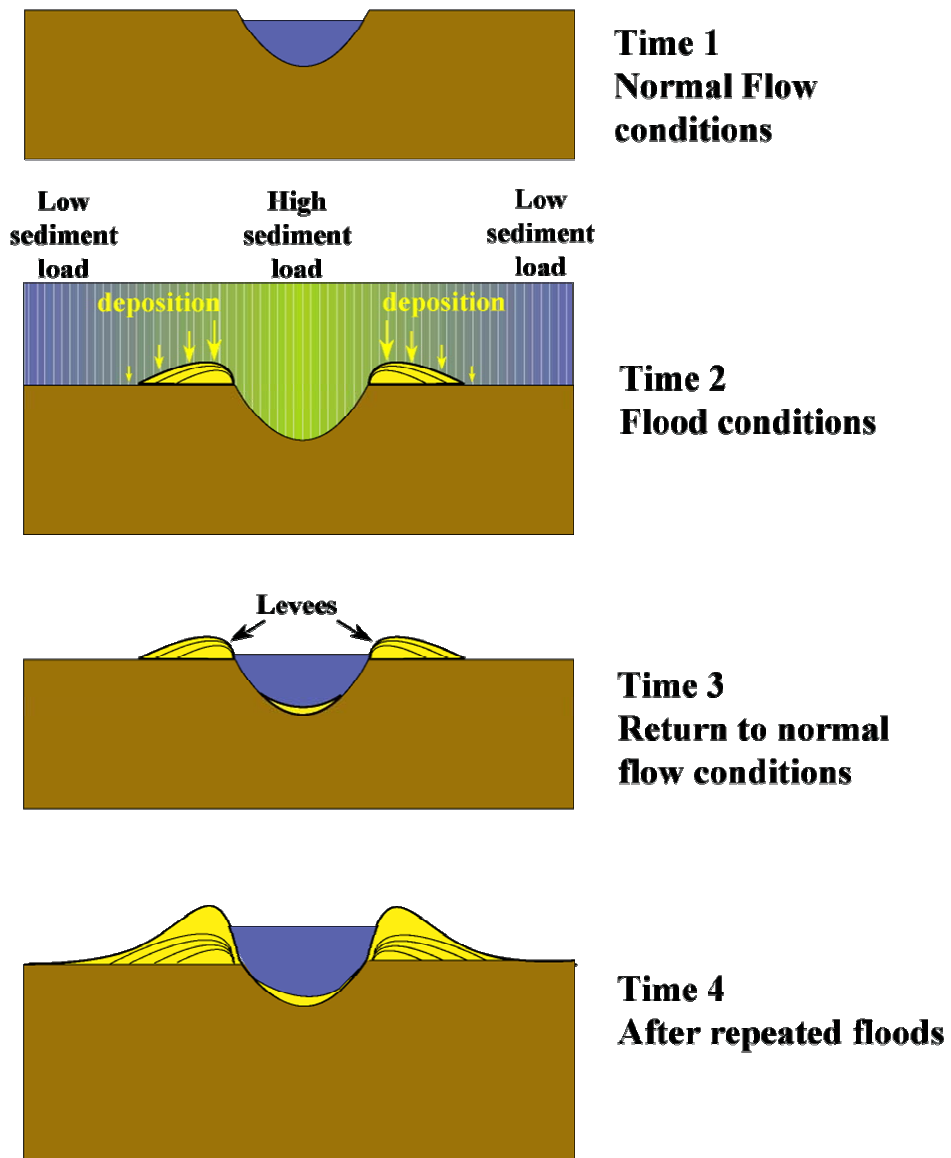


Figura 3.17. Etapas en la formación de un Natural Levee

La formación de los levees parten de un flujo normal que delimita los sedimentos restringidos al canal. Luego el agua y los sedimentos sobre pasan los limites de borde del canal, en condiciones de inundación. El flujo retorna a las condiciones normales y se determina el depósito del levee. El proceso repetitivo de estos ciclos de inundación – depositación genera el incremento en altura del depósito del levee. Ver Figuras 3.17 y 3.18.

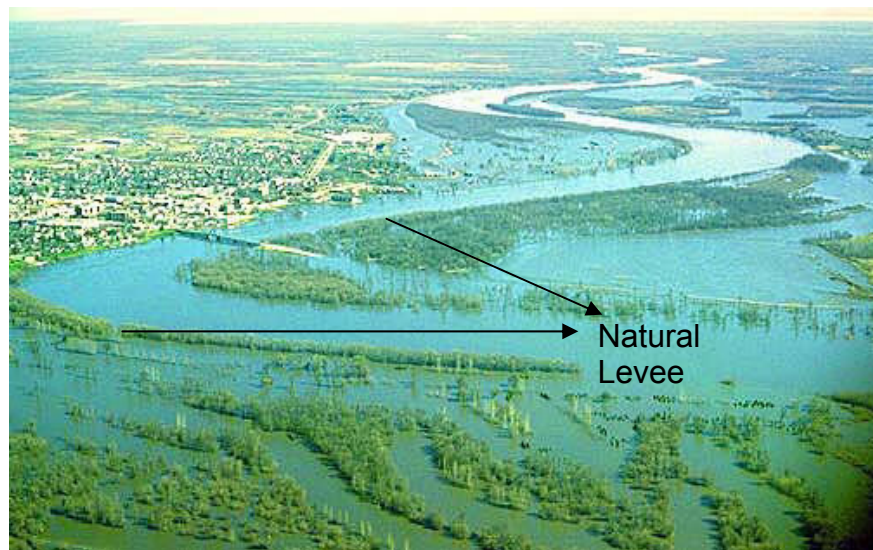


Figura 3.18. Ejemplo de un Natural Levee

3.4.4. Formación de los crevasse splays. Los crevasse splays son como depósitos de deltas que se forman a los márgenes de canales principales. Son interdigitados y típicamente consisten de areniscas de grano fino a medio con abundantes estructuras sedimentarias, raíces de plantas y bioturbación. Este tipo de arenas es de menor calidad de las que se encuentran en las dos facies anteriores.

3.4.5. Llanura de inundación. Este tipo de facies, la llanura de inundación (*floodplain Shale*), es considerada como el *background* de los principales tipos de depósitos de canal, cuando son encontrados en una sucesión estratigráfica. Consiste de sedimentos finos depositados por suspensión a partir de las aguas que inundan la planicie adyacente al canal. El *floodplain Shale* puede ser visto como una matriz en la cual son embebidos o incrustados los objetos que representan las arenas de buena calidad del yacimiento. Uno puede considerar la llanura de inundación como un objeto en el cual se presenta una matriz de arena con un porcentaje menor al 10% en relación con el *shale* de los depósitos.

Un floodplain es una planicie de pendiente suave construida por sucesivas capas de depósitos de aluvión ricos en material fino y dados por diferentes periodos de inundación. Durante fuertes periodos de lluvia, cantidades de agua incrementan los ríos cubriendo las tierras adyacentes al canal. Ver Figuras 3.19 y 3.20.



Figura 3.19. Ejemplo de una llanura de inundación en un canal meandriforme

Los sedimentos presentes en las llanuras de inundación son ricos en materia orgánica, son excelentes terrenos para muchos cultivos y material vegetal.

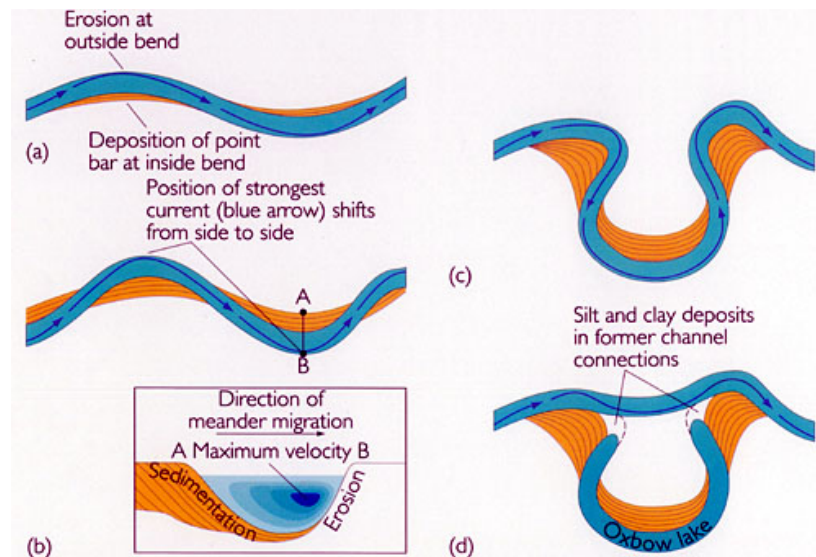


Figura 3.20. Formación de un Oxbow lake en un canal meandriforme

Los oxbow lake son depósitos característicos de sectores de inundación formados por el aislamiento de un sector del canal debido al rompimiento del mismo. El depósito del oxbow lake crea un ambiente característico de vegetación dada por materiales finos. Este tipo de depósito puede caracterizar lentes aislados de inundación que se encuentran dentro de los depósitos de canales de meandros.

En esta tesis, se hará mas énfasis en canales, natural levee y en los flood plain shales, como se mostrará en el siguiente capítulo.

3.5. CONTROLES DE CANAL Y MORFOLOGÍA

En la definición de la morfología del canal varios elementos son importantes para controlar el comportamiento del canal:

- Descarga
- Carga de Sedimento
- Ancho
- Profundidad
- Velocidad
- Pendiente
- Rugosidad de la capa
- Densidad de Vegetación

En el tipo de depósito y la arquitectura del canal los principales elementos que definen la geometría del modelo del canal son el ancho, la profundidad y la sinuosidad. Ver Figura 3.21.

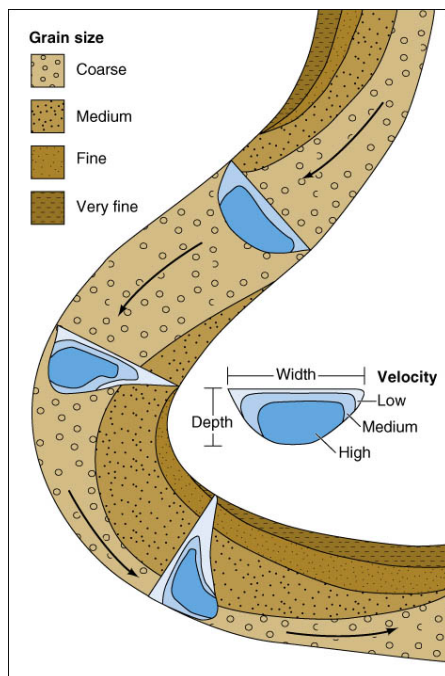


Figura 3.21. Tipos de depósitos aportados en un canal de acuerdo a la posición de la trayectoria

3.6. ARQUITECTURA DE LOS ELEMENTOS DE CANALES

Los tipos de depósitos generados por los ríos se pueden clasificar de acuerdo a diferentes componentes descriptivos y genéticos, que conforman la arquitectura del canal. Estos son:

1. De acuerdo a la naturaleza de las superficies límite superior e inferior del depósito. Superficies erosivas o gradacionales, planares, irregulares o curvadas (Convexas o cóncavas).
2. De acuerdo a su geometría: Tabulares, lentes, cuñas, en forma de cuchara, rellenos de canal en forma de U.
3. De acuerdo a su escala: espesor, extensión paralela lateral y perpendicular a la dirección de flujo.
4. De acuerdo a su litología: Ensamblaje de litofacies y secuencia vertical
5. Geometría interna: Naturaleza y disposición interna de superficies límite (paralelas, truncadas, onlap, download).
6. Patrones de paleo corrientes: orientación de indicadores de flujo relativo.

La integración de cada uno de estos elementos permite definir la arquitectura del canal. Con el fin de establecer un modelo de objetos el cual pueda ser modelado por propiedades petrofísicas y litológicas, es necesario interactuar con los diferentes parámetros para construir un canal que describa de forma más correcta de acuerdo a las características que describen un patrón de canal.

4. MODELAMIENTO DEL YACIMIENTO

4.1. DESCRIPCIÓN DEL CAMPO

El Campo Castilla se localiza en la Cuenca de los Llanos, 35 Km. al Sur de la ciudad de Villavicencio, en el Departamento del Meta. (Fig. 4.1)

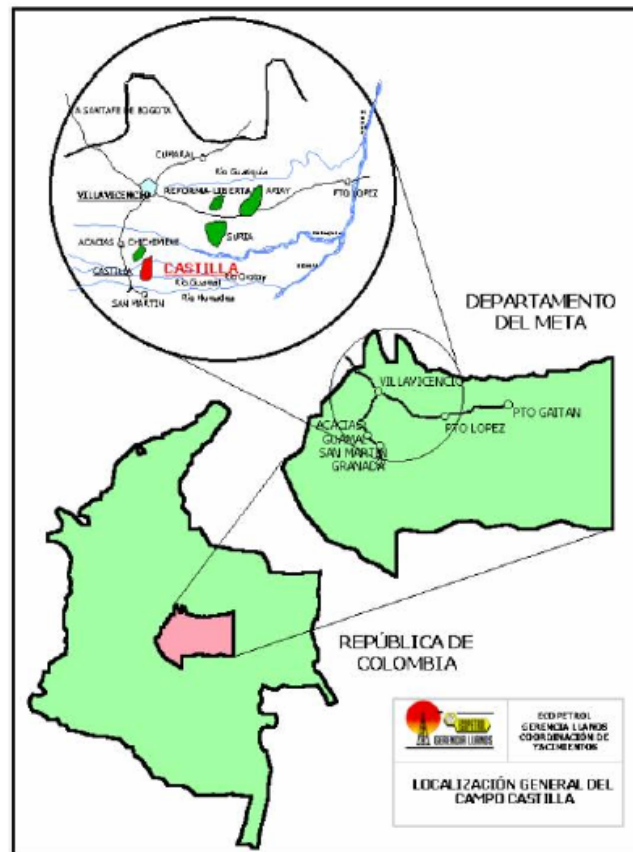


Figura 4.1. Mapa Localización del campo

El Campo Castilla fue descubierto por Chevron en el año de 1,969 con la perforación del pozo Castilla 1, pero inició producción en el año de 1,977, con el contrato de asociación Cubarral suscrito entre ECOPETROL y la compañía Chevron.

El 30 de enero de 2000 terminó el contrato de asociación Cubarral en el que Chevron administraba y operaba en los Campos Castilla y Chichimene. Con un contrato de operación, Chevron manejó el campo entre el 31 de enero y el 30 de julio de 2000, tras el cual la Gerencia Llanos asumió la operación del campo a partir del 31 de julio de 2000.

El campo Castilla cuenta actualmente con 42 pozos perforados, de los cuales 40 pozos activos producen crudo pesado de 13.7°API, con un potencial de producción de 52 Kbpd aproximadamente.

De los 29 pozos perforados hasta el año 2000, 3 fueron perforados por Ecopetrol en los bloques denominados Castilla Este y Castilla Norte. A pesar de estas denominaciones, estos tres pozos pertenecen a la misma estructura del Campo Castilla. Adicionalmente en el año 2001 Ecopetrol perforó los pozos Castilla 25H y Castilla 26, en el 2002 los pozos Castilla 27 a 29 y en el 2003 Castilla 30 a 31.

Desde el punto de vista de yacimientos, el campo posee un Estudio Integrado de Yacimientos elaborado por Chevron en 1993. Posteriormente la Gerencia de Yacimientos de Ecopetrol inició en 1998 una re-evaluación de geología y simulación que terminó en la Evaluación Integrada de Yacimientos del Campo Castilla.

A partir de los resultados de la simulación comprendida en dicho estudio y revisadas durante el estudio de explotación adicional, se presentaron varias acciones de desarrollo para el campo dentro de las cuales se contempla la viabilidad técnica y económica de la perforación de dos pozos verticales y ocho pozos horizontales.

No obstante, dado los buenos resultados de los pozos Castilla 25H y 26, además del tamaño del campo al cual se le ha cuantificado un OOIP de 2600 Millones de Barriles, se plantea un desarrollo adicional que no solo mantenga una tasa de producción de 52000 BIs/día para el campo, sino que aumente de forma significativa el factor de recobro del mismo, por lo que en estos momentos se está interpretando un volumen sísmico 3D adquirido el pasado año en el campo, el cual permitirá conocer de forma mas detallada el yacimiento y así proponer nuevos escenarios de desarrollo; entre los que se vislumbra la posibilidad de perforar 35 pozos verticales y 6 pozos horizontales, alcanzándose una tasa de producción de petróleo cercano a los 80000 barriles de petróleo por día.

La formación productora en este campo como en los campos adyacentes (Chichimene, Suria y Apiay), es la Formación Guadalupe correspondiente al Cretáceo, la cual a su vez se divide en 2 unidades: Unidad Guadalupe Superior (compuesto por una arena delgada, arcillosa, relativamente discontinua y de menor productividad) y Unidad Guadalupe Masivo (conformado por arenas gruesas y limpias, libres de arcillas y es el principal intervalo productor). Estas unidades operacionales corresponden a las unidades K1 y K2 respectivamente. Además encontramos en la parte superior la unidad San Fernando. (Fig. 4.2).

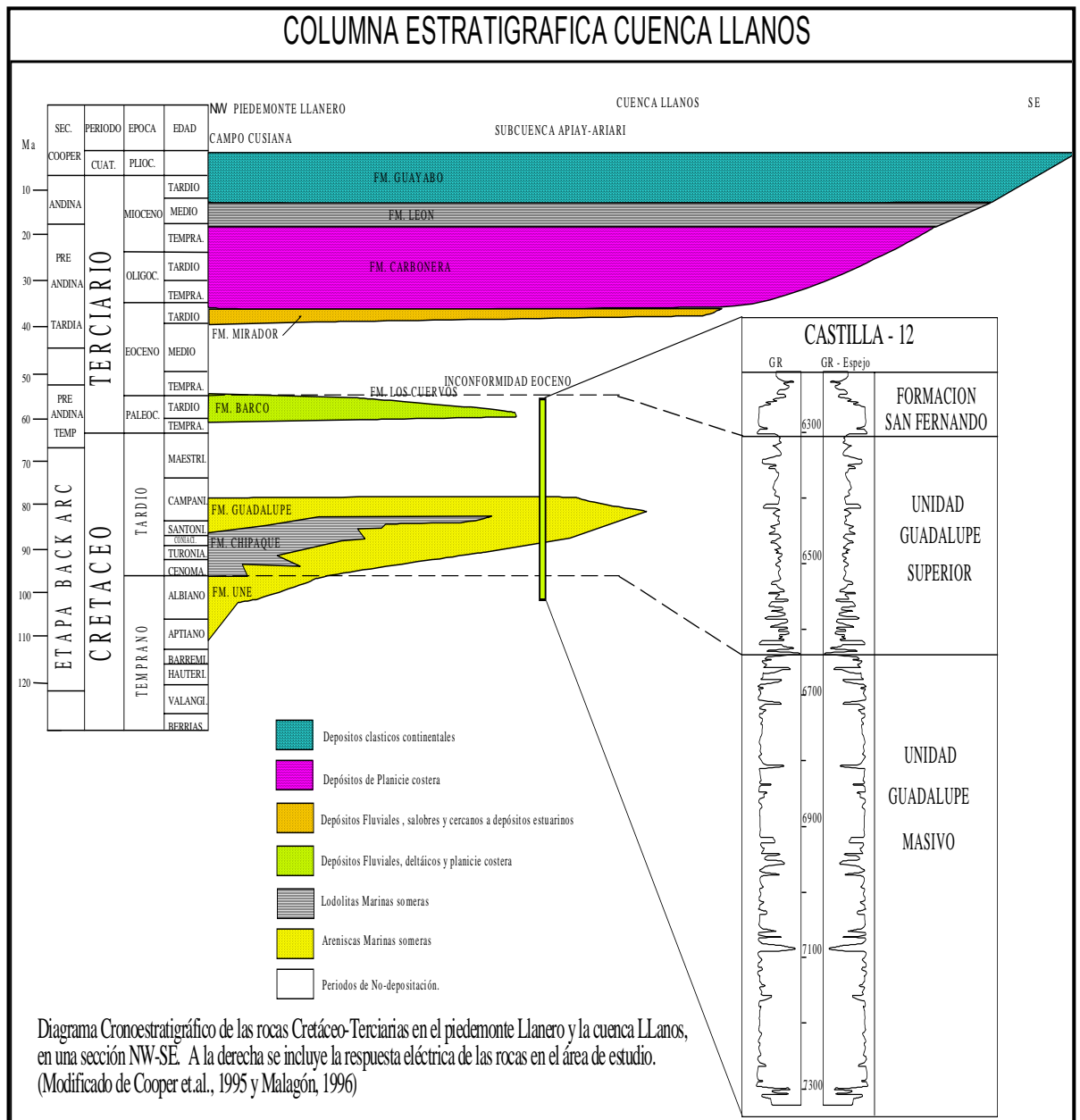


Figura 4.2. Columna estratigráfica

Regionalmente estas unidades representan una secuencia depositacional de larga duración en el cretáceo superior, separada por inconformidades a la base y al tope, con numerosas secuencias de menor orden dentro de esta.

Se reconocen tres patrones de sedimentación, que corresponden a la forma en que se depositaron en la cuenca y su apilamiento vertical. Estos se ven reflejados en electrofacies características para cada evento. (Fig. 4.3).

Inicialmente ocurre la depositación agradacional de las areniscas fluviales de la Unidad Guadalupe Masivo en condiciones de incremento en el espacio de acomodación y aporte de sedimentos en una relación equilibrada, hacia el final se incrementa el espacio disponible para sedimentos, o el abastecimiento de estos disminuye, ocurriendo, una superficie transgresiva marina, que marca el tope de la Unidad Guadalupe Masivo. La respuesta eléctrica de este intervalo en los registros de pozos es de forma cilíndrica o de bloque. Posteriormente se continúa en condiciones de incremento en el espacio de acomodamiento con relación al abastecimiento de sedimentos, dando como resultado la depositación de areniscas y shales inter-estratificados transgresivos de la Unidad Guadalupe Superior inferior, con un patrón de apilamiento retrogradacional, hasta alcanzar una superficie de máxima inundación (MFS), la cual es de carácter regional. La respuesta eléctrica en los registros de pozo, para este intervalo es de campana.

Por último, suprayaciendo esta superficie de máxima inundación se presenta un apilamiento progradacional de shales y areniscas resultado de una disminución en el espacio de acomodamiento en relación con el abastecimiento de sedimentos. La respuesta eléctrica en los registros de pozo, para este intervalo es de embudo. Este último evento es suprayacido por una inconformidad principal que corresponde al tope de la Unidad Guadalupe Superior y que está representada por la base de areniscas fluviales de la Unidad San Fernando.

Para el análisis detallado de las Unidades, se cuenta con la descripción de aproximadamente 1150 pies de corazones de los pozos Castilla-3, Castilla-12 y Castilla-18.

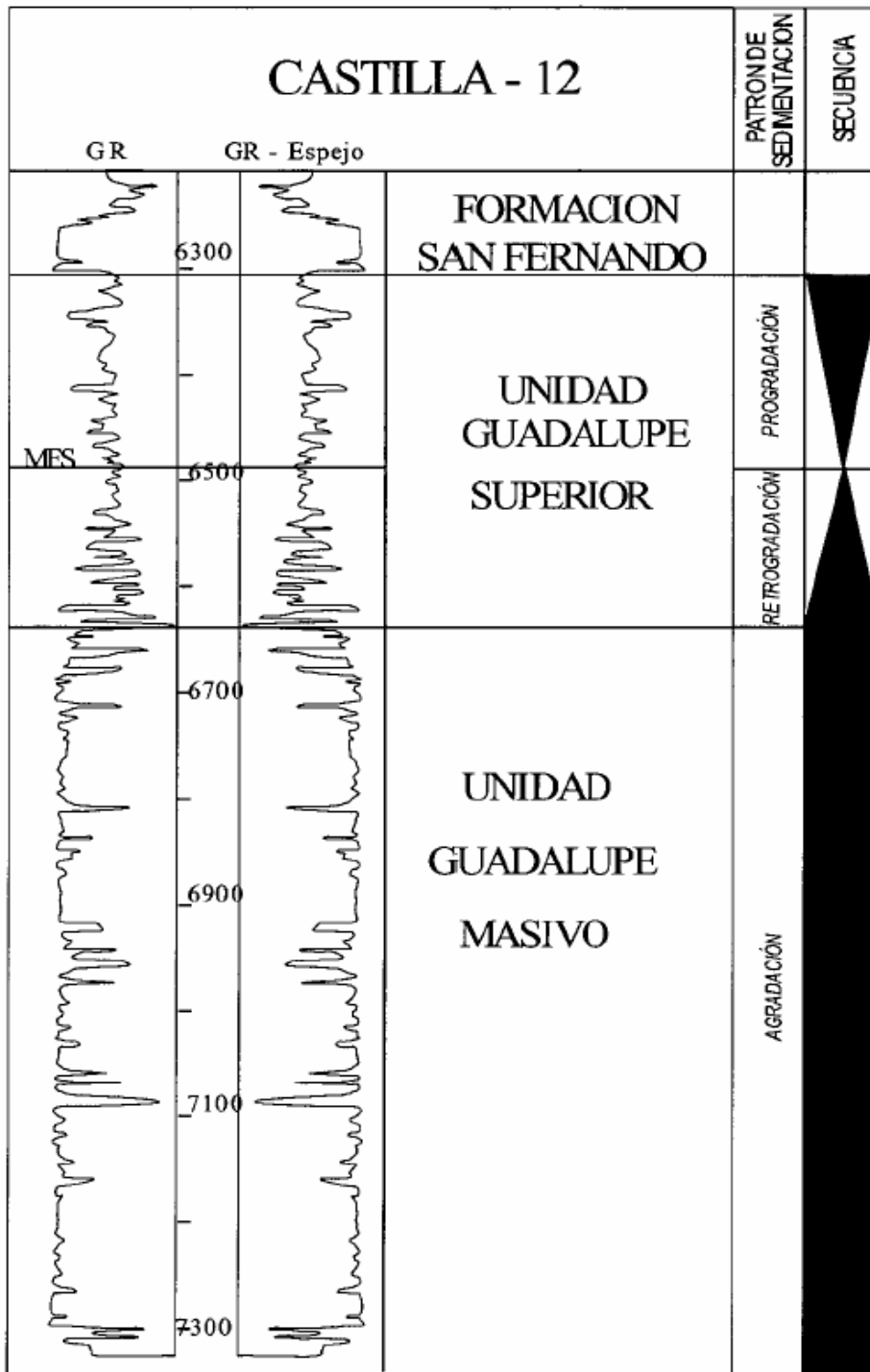


Figura 4.3 Secuencia depositacional del Campo Castilla

Para este trabajo se seleccionó la unidad Guadalupe Masivo, por ser esta de carácter fluvial y por ser el principal intervalo productor del campo; sobre esta área seleccionada se hizo el estudio estratigráfico y la descripción petrofísica.

4.1.1. Unidad Guadalupe Masivo

Corresponde a areniscas de grano grueso a medio y gravas arenosas, con estratificación cruzada depositadas masivamente o como ciclos grano decrecientes, dentro de una extensiva llanura de costa, interceptada por grandes y permanentes corrientes fluviales trenzados (braided). Ocasionalmente entre estos cuerpos de arena se presentan intercalados shales y limos, correspondientes a depósitos de rebose de canales (overbank), o de llanura de inundación (floodplain) y que no presentan mucha continuidad lateral (Fig. 4.4). Al parecer la fuente de los sedimentos para ésta Unidad es el escudo de la Guyana situado al Oriente de la cuenca.

Esta unidad está infrayacida discordantemente por depósitos Paleozoicos y suprayacida por la *Unidad Guadalupe Superior*. El espesor total varía entre 680 y 750 pies.

Las principales litofacies arenosas identificadas en los núcleos corresponden a subambientes tales como: base de canales de carga de fondo, canales fluviales sinuosos a entrelazados (braided), barras longitudinales, transversales y laterales de canales de cargas de fondo entrelazados, relleno de barras y relleno de canales. Estas litofacies se encuentran apiladas verticalmente, conformando gruesos paquetes arenosos de hasta 100 pies de espesor, con una respuesta de registro gamma ray de cilindro muy homogénea y limpia, por lo cual es difícil hacer diferenciación interna de unidades con base en este registro.

Intercalando estas litofacies arenosas se presentan delgadas capas de shales limosos o limosos arcillosos de hasta 8 pies de espesor que corresponden a depósitos de llanura de inundación y/o de rebose de canal (overbank). Estas litofacies separan los gruesos paquetes arenosos descritos anteriormente y representan de un 5 a 10% del espesor total de esta unidad. Aunque se presentan en todos los pozos, no hay intervalos en los cuales exista continuidad lateral, sino que se presentan a manera de lentes discontinuos tanto lateral como verticalmente. Sin embargo se consideran que actúan como importantes barreras de permeabilidad vertical.

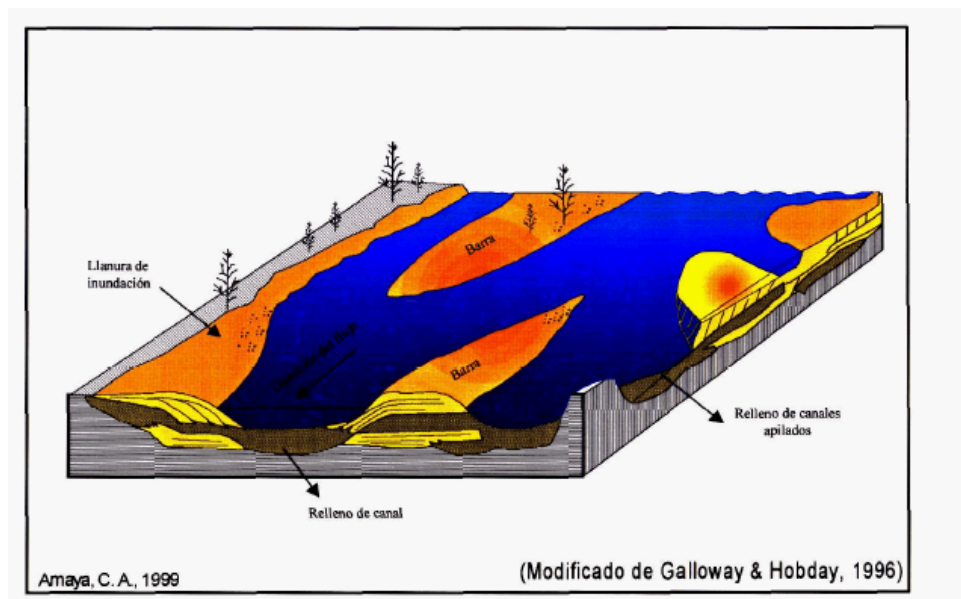


Figura. 4.4 Ambiente de deposición

De acuerdo con análisis petrográficos efectuados por Chevron en esta unidad, tales como análisis de secciones delgadas, de microscopía electrónica de barrido y de rayos x, se determinó que el componente principal de las areniscas es cuarzo, con otros componentes detríticos menores tales como caolinita y micas y

como componentes antigénicos caolinita y pirita. Son clasificados como cuarzoarenitas.

El cuarzo se presenta principalmente como granos detríticos en contactos unos con otros, la presión ha causado superficies irregulares entre granos, aunque pocas veces produce contactos saturados o cóncavo convexos. La sílice disuelta es precipitada en espacios porosos adyacentes como sobrecrecimientos euhedrales, aunque estos también pudieron formarse de sílice derivada de arcillas y feldespatos. En intervalos conglomeráticos el tipo de cuarzo es compuesto, lo cual evidencia origen de rocas metamórficas. El contenido total de cuarzo es de 80 a 95% en peso de toda la roca.

La caolinita se presenta en varias formas: como cuerpos irregulares detríticos y como placas pseudo hexagonales autigénicas, con predominio del primero. El porcentaje en peso de la caolinita es de 5 a 18 %. La pirita ocurre como un mineral diagenético.

La historia diagenética es: compactación, disolución y precipitación de sílice, luego cementación por caolinita y por último formación de pirita. Los sobrecrecimientos de sílice, caolinita autigénica y desarrollo de pirita, aunque son muy comunes en todas las muestras, no han taponado los poros significativamente. La figura 4.5 muestra una sección delgada típica de estas areniscas.

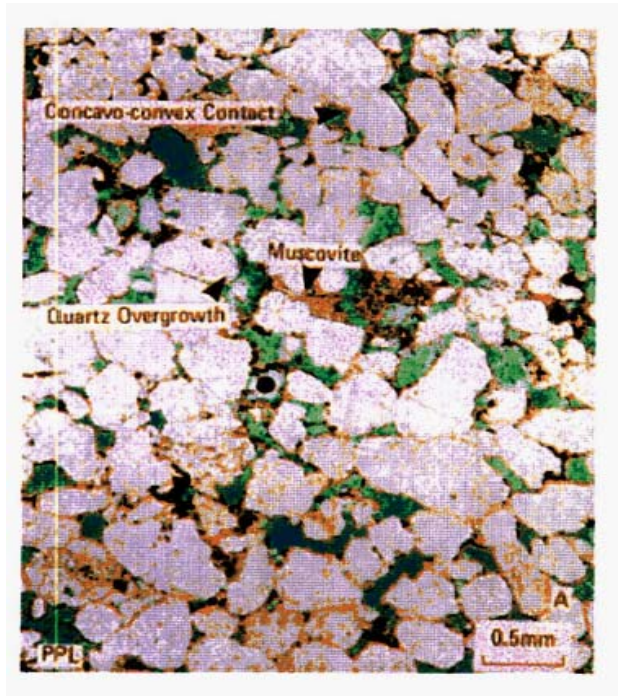


Figura 4.5 Sección Delgada Unidad Guadalupe Masivo

4.2. MARCO ESTRUCTURAL DEL CAMPO

Geológicamente el Campo Castilla hace parte de la Cuenca Llanos, su estructura corresponde a un anticlinal cuyo eje presenta una orientación N 60° E, de 4 Km. de ancho y 10 Km. de largo, limitado al Este por una falla inversa. Se encuentra afectado por una serie de fallas normales e inversas con dirección NE-SW, algunas asociadas a la falla principal. Las fallas normales presentan saltos que varían de 30 a 120 pies y las fallas inversas con saltos mucho menores que varían entre 30 y 50 pies. Al parecer las fallas internas que afectan la estructura no presentan carácter sellante, ya que el contacto agua-aceite no presenta cambios abruptos al lado y lado de las fallas, mostrando en general a lo largo del campo una misma inclinación.

La figura 4.6 presenta el mapa estructural al tope de la unidad Guadalupe masivo.

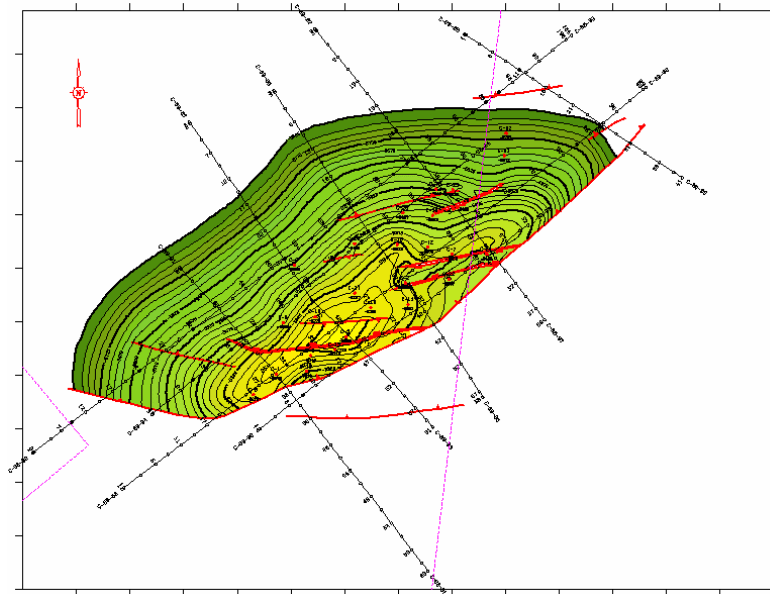


Figura 4.6 Mapa Estructural De La Unidad Guadalupe Masivo

4.3. MODELO PETROFÍSICO

Los pozos del campo castilla cuentan en general con un conjunto básico de registros y como adicional no convencional, algunos pozos con registro EPT (Ver tabla 4.1).

Existe información de análisis básicos y especiales de los pozos castilla-3, castilla-9, castilla-12 y castilla-18. La tabla 4.2 resume los análisis elaborados para cada pozo.

POZO	GR	SP	CAL	LLD	LLS	MSFL	RHOB	PHIN	DT	EPT	ML	CORAZÓN
Castilla-1	X	X	X	*	*	*	X		X			
Castilla-2	X	X	X	X	X		X		X			
Castilla-3	X	X	X	X	X	X	X	X	X			X
Castilla-4	X	X	X	X	X	X	X	X	X			
Castilla-5	X	X	X	X	X	X	X					
Castilla-7	X	X	X	X	X	X	X	X	X			
Castilla-8	X		X	X	X		X	X	X			
Castilla-9	X	X	X	X	X	X	X	X	X		X	X
Castilla-10	X	X	X	X	X	X	X	X	X			
Castilla-11	X	X	X	X	X	X	X			X	X	
Castilla-12	X	X	X	X	X		X	X		X		X
Castilla-13	X	X	X	X	X		X	X				
Castilla-14	X	X	X	X	X	X	X			X	X	
Castilla-15	X	X	X	X	X	X	X	X		X	X	
Castilla-16	X	X	X	X	X	X	X	X		X	X	
Castilla-17	X	X	X	X	X	X	X	X				
Castilla-18	X	X	X	X	X	X	X	X				X
Castilla-19	X	X	X	X	X	X	X	X		X	X	
Castilla-20	X	X	X	X	X	X	X	X				
Castilla-21	X	X	X	X	X	X	X	X		X	X	
Castilla-22	X	X	X	X	X	X	X	X		X		
Castilla-23	X	X	X	X	X	X	X	X		X	X	
Castilla-24	X	X	X	X	X		X	X		X	X	
Castilla-25H	X	X	X	X	X	X	X	X	X			
Castilla-26	X	X	X	X	X	X	X	X	X			
Castilla-27	X	X	X	X	X	X	X	X	X			
Castilla-28	X	X	X	X	X	X	X	X	X			
Castilla-29	X	X	X	X	X	X	X	X	X			
Castilla-30	X	X	X	X	X	X	X	X	X			
Castilla-31	X	X	X	X	X	X	X	X	X			
Castilla-N1	X	X	X	X	X	X	X	X	X			
Castilla-N2	X	X	X	X	X	X	X	X				
Castilla-N3	X	X	X	X	X	X	X	X	X			
Castilla-E1	X	X	X	X	X	X	X	X			X	

- Los pozos Castilla-26, 31 y N-3 poseen registros HLLD/HLLS/MSFL
- Los pozos Castilla-25H, 29, 31 y N-3 poseen registros TNPH
- Los pozos Castilla-25H y 26 poseen registros *arrays* AHT-10,20,30,60,90
- Los pozos Castilla-25H, 26, 27, 28, 29, 30, 31 y N-3 poseen registro PEF

Tabla 4.1 Inventario de Registros de Pozo. Campo Castilla

Pozo	Unidad	Análisis Básicos	Análisis Especiales	Año
Castilla-3	Guadalupe Superior Guadalupe Masivo (sup.-med.)	Porosidad Helio Permeabilidad al aire, saturación, densidad del grano	Sección delgada, análisis SEM, Difracción de rayos X, permeabilidad relativa.	1980
Castilla-9	Guadalupe Superior Guadalupe Masivo	Porosidad Arquímedes Permeabilidad al aire	Permeabilidad relativa, Amott wettability test.	1983
Castilla-12	Guadalupe Masivo (med.-inf.)	Porosidad Helio. Permeabilidad al aire (Kv, Kh)	Presión capilar Permeabilidad relativa	1984
Castilla-18	Guadalupe Superior	Porosidad Arquímedes Permeabilidad al aire, saturación de fluidos, densidad de grano.		1986
Castilla-3 y Castilla-9	Guadalupe Superior Guadalupe Masivo	Análisis de Porosidad y Permeabilidad a dos presiones de confinamiento CMS-300, densidad de grano.		2003

Tabla 4.2. Análisis básicos y especiales Campo Castilla.

El análisis se llevó a cabo en dos fases. La primera en preparar los datos, que comprendió la carga y edición de registros. La segunda fase en el análisis de éstos que consistió en el cálculo de porosidad total, efectiva y determinación de permeabilidades.

4.3.1. Modelo de porosidad. Se empleó el registro densidad (RHOB), presente en todos los pozos del Campo, para determinar la porosidad total o aparente y se correlacionó con los valores de porosidad medidos sobre núcleo, a condiciones ambiente.

La ecuación empleada fue:

$$DPHI = \frac{(RHOMA - RHOB)}{(RHOMA - RHOF)} \quad (4.1)$$

Donde:

DPHI = Porosidad aparente a partir del registro densidad (fracción)

RHOMA = Densidad de matriz de la formación (gr/cc)

$RHOB$ = Respuesta de la herramienta de densidad (gr/cc)

$RHOF$ = Densidad de fluido (gr/cc)

Este modelo tuvo en cuenta, en los pozos que lo poseían, el registro de corrección de densidad ($DRHO$) para identificar intervalos donde la respuesta de densidad no era confiable. El criterio empleado fue ($DRHO < 0.2$) para considerar válido el registro $RHOB$.

Para el cálculo de densidad de matriz ($RHOMA$), se empleó la información proveniente del análisis de núcleos de los pozos Castilla-12, que corazonó la parte media-inferior de K2 y Castilla-3 que corazonó la totalidad de K2. El promedio de densidad de grano, medido en laboratorio para estos dos pozos, se estimó en 2.646 gr/cc considerando un ligero aumento de arcillosidad hacia la parte media-inferior del yacimiento evidenciado en Castilla_12.

Para la determinación de la densidad de fluido ($RHOF$), se consideraron las características de crudo viscoso de Castilla (13.7 °API - 190°F), que corresponden a un hidrocarburo de densidad de alrededor de 0.95 gr/cc muy próxima a la del agua, por lo que se asumió un valor de 1.0 gr/cc, como densidad de fluido para todos los cálculos. Los valores de densidad de matriz y fluido fueron corroborados con gráficos $RHOB$ -PHICORE, tal como se presenta en la Figura 4.7.

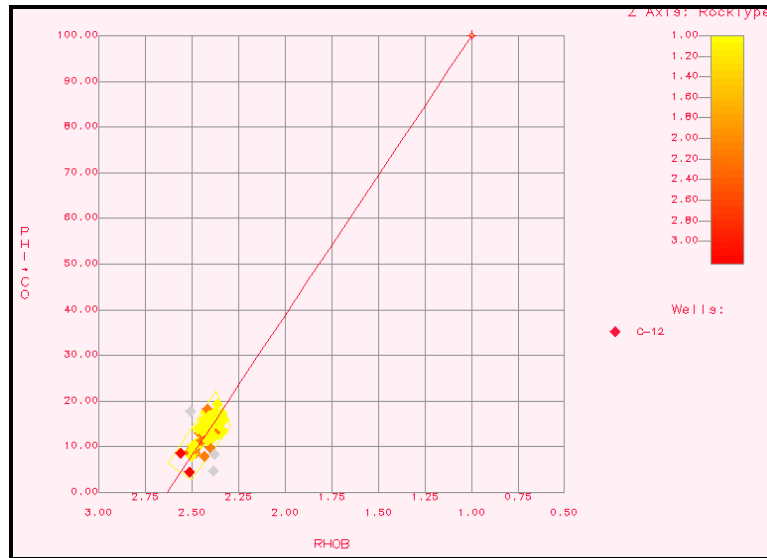


Figura 4.7. Gráfico de RHOB vs PHICORE. Castilla 12. Unidad K2

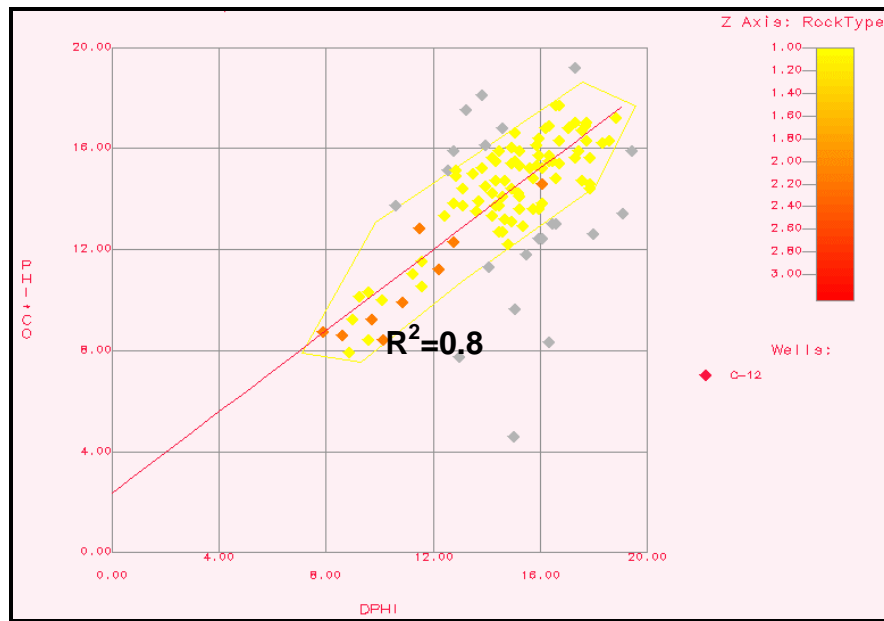
4.3.1.1. Cálculo del Volumen de Arcilla y Porosidad Efectiva.

Por lo analizado anteriormente, se considera que las arenas del K2 tienen mínima corrección por arcillosidad y que, debido a las características de arcilla estructural de la caolinita, ésta no afecta en gran proporción la porosidad efectiva de la arena. Las arenas limpias se pueden diferenciar claramente con el registro litológico Gamma Ray.

4.3.1.2. Comparación de la Curva de Porosidad Calculada con la Información de Corazones.

Una vez obtenido el mejor estimado de porosidad a partir de la interpretación de registros, se procedió a verificar su ajuste con los datos de porosidad de corazones corregidos a condiciones netas de confinamiento o de yacimiento. Para tener un ajuste confiable entre la curva de porosidad y los datos de corazones, fue necesario verificar su respectiva corrección con profundidad. La Figura 4.8

muestra la correlación obtenida entre la porosidad calculada a partir del registro de densidad y la medida en laboratorio para la unidad K2.



**Figura 4.8. Relación entre porosidad calculada y porosidad medida.
Castilla-12 Unidad K2.**

4.3.1.3. Evaluación de los Tipos de Roca

Una vez determinada la porosidad para la unidad K2 y analizado el comportamiento mineralógico de las Formación de interés, el paso siguiente en la evaluación fue determinar los tipos de rocas presentes para asociarlos con una respuesta de permeabilidad. El concepto de tipo de roca esta basado en el hecho de que la relación entre porosidad y permeabilidad es solamente válida dentro de una fracción del yacimiento con características de flujo similares. Por lo tanto, la identificación de los tipos de roca y la determinación de sus valores característicos de FZI es de fundamental importancia en el desarrollo de la relación implícita

entre porosidad y permeabilidad. Mediante una gráfica de Índice de Calidad de Yacimiento (RQI) vs. Porosidad Normalizada (PHIZ), se puede apreciar tres tendencias diferentes en el comportamiento de los datos, que pueden ser corroboradas con la información de núcleos.

Las Figuras 4.9 muestran los diferentes tipos de roca que pueden ser encontrados en la Unidad K2, empleando la información de corazones del pozo Castilla 12.

Las ecuaciones necesarias para el desarrollo de la gráfica son:

$$RQI = 0.0314 * \sqrt{\frac{K}{\phi}} \quad (4.2)$$

$$PHIZ = \frac{\phi}{(1-\phi)} \quad (4.3)$$

$$FZI = \frac{RQI}{PHIZ} \quad (4.4)$$

Donde,

RQI = Reservoir Quality Index (□)

PHIZ = Porosidad Normalizada

FZI = Flow Zone Indicator

K = Permeabilidad (mD)

□ = *Porosidad (fracción)*

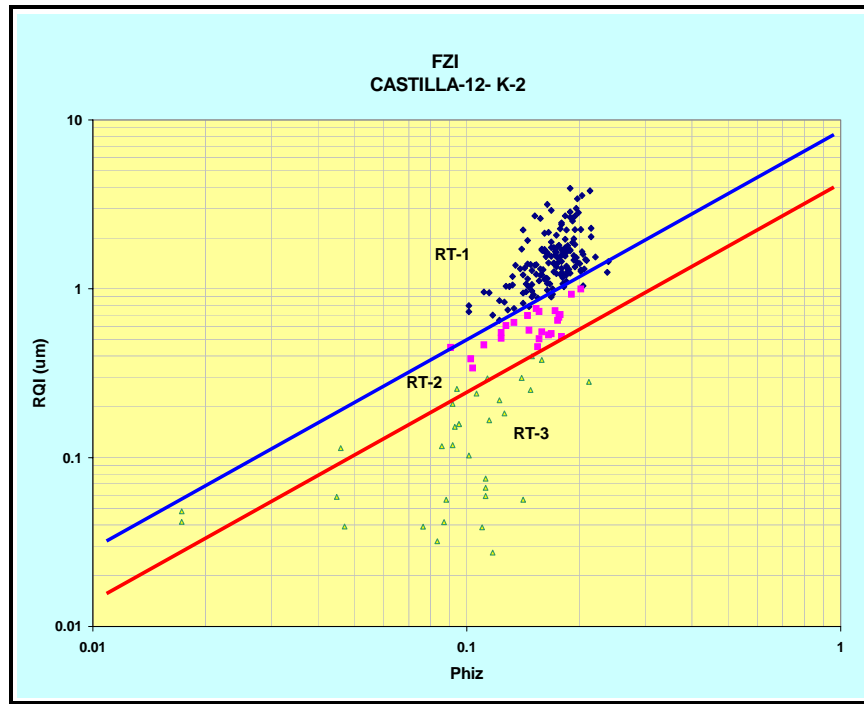


Figura 4.9 Gráfica para identificar tipos de roca para Castilla-12, Unidad K2.

Una vez identificados los tipos de roca con la información de corazones, es necesario encontrar curvas de registro que puedan discriminar adecuadamente estos tipos de rocas y que se puedan extrapolar a todos los pozos del yacimiento. La Figura 4.10 muestra el comportamiento de la curva GR vs. la permeabilidad medida sobre el núcleo del pozo Castilla-12 (Unidad K2); en esta gráfica se aprecian claramente tres tendencias diferentes, dependiendo del comportamiento (calidad) de las arenas. Las rocas denominadas Tipo 1, corresponden a aquellas arenas muy limpias ($GR < 45$ UAPI) con altos valores permeabilidad (> 100 mD) y que se consideran las principales rocas de aporte de flujo de fluidos; las rocas Tipo 2, son aquellas arenas un poco más sucias (GR entre 45 y 90 UAPI) y permeabilidad, en un rango más amplio entre 1 y 1000 mD, pero que por sus características de mayor arcillosidad, se consideran como rocas con

características de flujo más pobre que las primeras, pero con aporte. Por último, las rocas Tipo 3, son las consideradas arcillas, propiamente dichas, con valores de GR mayores a 90 UAPI y valores de permeabilidad por debajo de 100 mD, lo que las convierte en auténticas barreras de flujo.

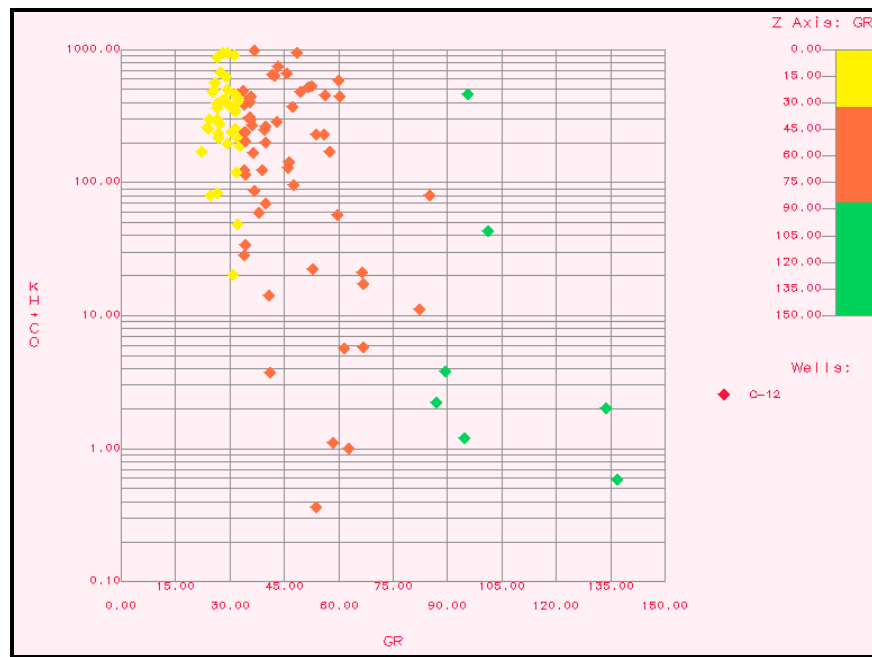


Figura 4.10. Comportamiento de la curva GR con la permeabilidad de corazones. Castilla-12-K2

De esta forma, se determinó la curva GR como el mejor discriminante de la calidad de roca del yacimiento para la Unidad K2 y a partir de ella se pudo generar el modelo de Tipo de Roca más apropiado para el Campo, el cual fue validado con la visualización, descripción y análisis de núcleos.

TIPOS DE ROCA K2

if (GR >90) then "RockType"=3; end if;

if ((GR<=90) AND (GR>=45)) then "RockType"=2; end if;

if (GR<45) then RockType=1; end if;

4.3.2 Evaluación de la permeabilidad. La evaluación de la permeabilidad del yacimiento se obtuvo con base en la ecuación resultante de la relación porosidad/permeabilidad de los corazones para cada uno de los tipos de roca identificados en la Unidad K2.

Las ecuaciones resultantes fueron:

PERMEABILIDAD K2

IF ("RockType"= 1) THEN "PERM"= (-684.482+71.586*"DPHI"); END IF;

IF ("RockType"= 2) THEN "PERM"= (-495.29+53.27*"DPHI"); END IF;

IF ("RockType"= 3) THEN "PERM"=100; END IF;

IF (("RockType"= 1) AND ("PERM"<0)) THEN "PERM"=200; END IF;

IF (("RockType"= 2) AND ("PERM"<0)) THEN "PERM"=100; END IF;

En los modelos anteriores, la porosidad está en porcentaje y para los casos donde su valor es muy bajo producen valores de permeabilidad negativos, los cuales se han corregido.

Los datos de permeabilidad vs porosidad de corazón, para la unidad K2 se pueden observar en la Figura 4.11.

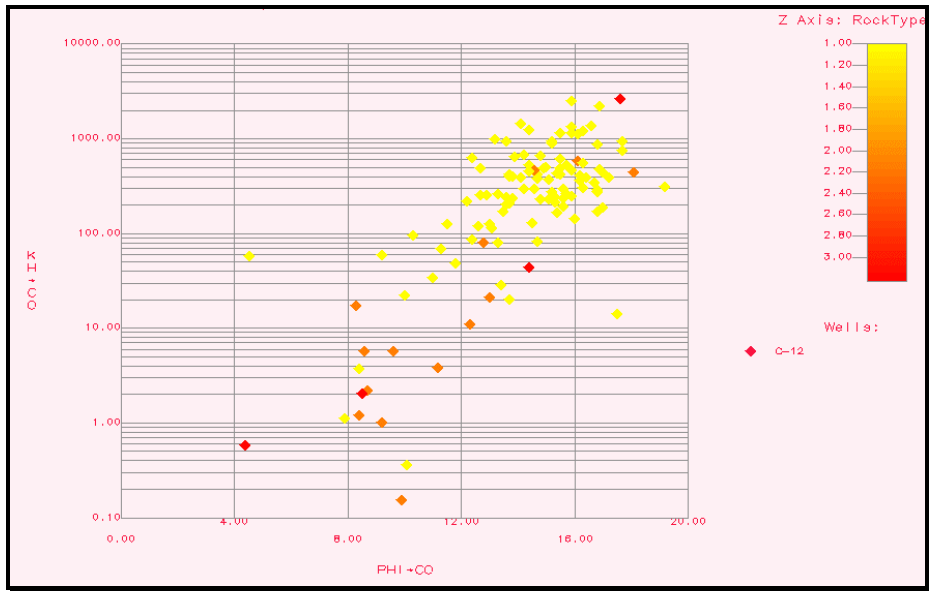


Figura 4.11. Comportamiento de la gráfica de porosidad vs permeabilidad de corazones. Castilla-12-K2.

Con estos modelos de porosidad y permeabilidad generados a partir de corazones y registros se obtuvieron los datos de estas propiedades petrofísicas en los pozos, que serán utilizados para correr el software geoestadístico “sgsim” y así obtener la caracterización entre pozos.

4.4. MODELAMIENTO GEOESTADÍSTICO

El Modelamiento geoestadístico fue realizado con el objetivo de proporcionar la mas cercana representación de la heterogeneidad geológica de la unidad Guadalupe masivo del campo. Se efectuaron simulaciones determinísticas utilizando métodos basados en funciones de probabilidad Gaussianas. Mediante esta técnica se construyeron las distribuciones 3D de las propiedades de la unidad condicionadas al marco estructural vigente.

La aplicación de la metodología fue lograda en tres etapas: Primero, la preparación y carga de datos, segundo, el modelamiento de facies y tercero, el modelamiento de parámetros petrofísicos. Estas etapas se detallan a continuación.

- **Preparación y carga de datos**

Se cargó la información de modelamiento de facies, desarrollada por el grupo de modelamiento de procesos de hidrocarburos (GMPH). Además la información petrofísica de todos los pozos. Posteriormente se creó la unidad Guadalupe Masivo y se generó el grid de simulación.

- **Modelamiento de Facies**

Para el desarrollo de la caracterización se contó con un total de 1663,15 pies de corazones, correspondientes a la formación Guadalupe. Estos pies están distribuidos en un total de 5 pozos, como se muestra en la tabla 4.3.

TOTAL PIES CORAZONADOS	
POZO	PIES
C1	39
C3	532,4
C9	450,25
C12	456,7
C18	184,8
TOTAL PIES	1663,15

Tabla 4.3. Intervalos corazonados para los pozos de Castilla.

Después de tener claramente identificadas la litofacies en los corazones, se analizó la respuesta de los registros eléctricos ante dichas facies identificadas en

los corazones de perforación. En la tabla 4.4 se muestran los pozos sobre los cuales se trabajó la identificación de las electrofacies.

POZO	RT	K1		MFS		K2		TRANS		OWC	
		MD	TVD	MD	TVD	MD	TVD	MD	TVD	MD	TVD
C-1	1330	6224	6222	6367	6364	6589	6586	6844	6841	6940	6937
C-10	1299	6219	6219	6327	6326	6565	6564	6952	6951	7030	7029
C-11	1302	6252	6250	6364	6362	6604	6602	6962	6960	7056	7054
C-12	1303	6295	6292	6437	6435	6692	6689	7040	7037	7170	7167
C-13	1313	6174	6170	6344	6340	6571	6566	6870	6865	6950	6945
C-14	1331	6246	6245	6370	6369	6604	6603	6900	6898	7000	6997
C-15	1315	6281	6279	6448	6446	6678	6676	6825	6822	6900	6897
C-16	1323	6264	6263	6400	6399	6640	6639	7050	7048	7170	7168
C-17	1328	6262	6260	6364	6363	6600	6597	6932	6927	7070	7065
C-18	1316	6475	6475	6608	6608	6864	6864	7158	7158	7314	7314
C-19	1311	6427	6427	6510	6510	6765	6765	7217	7217	7317	7317
C-2	1308	6336	6336	6474	6474	6713	6713	6980	6980	7090	7090
C-20	1322	6407	6407	6539	6539	6797	6797	7174	7174	7266	7266
C-21	1343	6372	6372	6505	6505	6747	6747	7076	7076	7163	7163
C-22	1290	6408	6408	6440	6440	6677	6677	7088	7088	7198	7198
C-23	1308	6515	6515	6591	6591	6831	6831	7185	7185	7310	7310
C-24	1303	6588	6588	6618	6618	6792	6792	7160	7160	7306	7306
C-26	1302	6397	6316	6517	6423	6774	6659	7169	7022	7302	7138
C-3	1313	6128	6127	6263	6262	6490	6489	6905	6904		
C-4	1351	6454	6454	6602	6602	6843	6843	7030	7030	7110	7110
C-5	1317	6223	6223	6380	6380	6607	6607	6870	6869	6924	6923
C-7	1288	6354	6352	6464	6462	6601	6599	7030	7028	7130	7128
C-8	1338	6338	6338	6484	6484	6718	6718				
C-9	1328	6276	6276			6598	6598	6760	6760	6860	6860
C-25H	1303	6438	6361	6596	6498	6946	6737				
C-25P	1303	6438	6361	6590	6498	6847	6736	7283	7144	7424	7276
C-27	1337	6249	6249	6327	6327	6561	6560	6930	6929	7021	7020
C-28	1328	6227	6226	6328	6327	6569	6567	6784	6781	6848	6845
C-29	1337	6277	6277	6405	6405	6647	6647	6872	6872	6956	6956
C-30	1338	6418	6417	6513	6512	6752	6751	6995	6993	7101	7099
C-31	1332	6349	6348	6451	6451	6686	6685	7007	7005	7112	7110
C-E1	1284	6378	6370	6540	6532	6779	6770	7195	7185	7255	7245
C-N1	1300	6564	6561	6724	6720	6975	6971	7350	7344	7460	7454
C-N2	1307	6672	6671	6804	6803	7053	7052	7416	7414	7516	7514
C-N3	1313	6526	6524	6671	6668	6914	6911	7273	7269	7399	7395
C-N4	1274	6770	6628			7241	6990	7630	7250	7803	7356

Tabla 4.4. Topes de la unidades en los pozos trabajados para el campo Castilla.

Primero se ubicaron en profundidad las facies, se realizó la corrección de las profundidades ya que la profundidad reportada por el registro no coincide completamente con la profundidad de las facies reportadas en los corazones.

Para el presente estudio se tomo como verdadera la profundidad reportada en los registros, esto debido a que en el estudio el porcentaje de información proporcionada por los registros es mucho mayor que el porcentaje de información proporcionada por los corazones.

Teniendo claramente identificadas las facies en los registros, se procedió a analizar el comportamiento de las curvas como GR y RHOB, estas curvas toman un valor característico para cada facies. Con esta información se analizó la totalidad de las facies presentes en los pozos corazonados.

Teniendo definidos los valores que presentaba cada una de las facies en las curvas de los registros (ver tabla 4.5), se pudo extrapolar la metodología a la totalidad de los pozos y en ellos interpretar la presencia de las diferentes electrofacies.

	Valores del Registro	Algoritmo Utilizado
CANALES	$0 < GR \leq 47$	<i>IF "GR" ≤ 47 AND "GR" > 0 THEN "F - CHANNEL" = 1; ELSE "F - CHANNEL" = 0; END IF;</i>
NATURAL LEVEE	$47 < GR \leq 70$	<i>IF "GR" ≤ 70 AND "GR" > 47 THEN "F - NATURAL" = 1; ELSE "F - NATURAL" = 0; END IF;</i>
FLOOD PLAIN SHALE	$70 < GR < 200$	<i>IF "GR" ≤ 200 AND "GR" > 70 THEN "F - FLOOD" = 1; ELSE "F - FLOOD" = 0; END IF;</i>

Tabla 4.5. Valores de los registros eléctricos y los algoritmos generados.

Para poder visualizar las electrofacies en profundidad fue necesario elaborar una serie de algoritmos, los cuales fueron cargados al software de Petroworks. Al

cargar estos algoritmos se daba la opción al software de graficar una última curva donde mostrara las electrofacies presentes en profundidad, estas electrofacies fueron interpretadas directamente por el Petroworks de acuerdo a los valores que tomara la curva en los registros Ver figura 4.12.

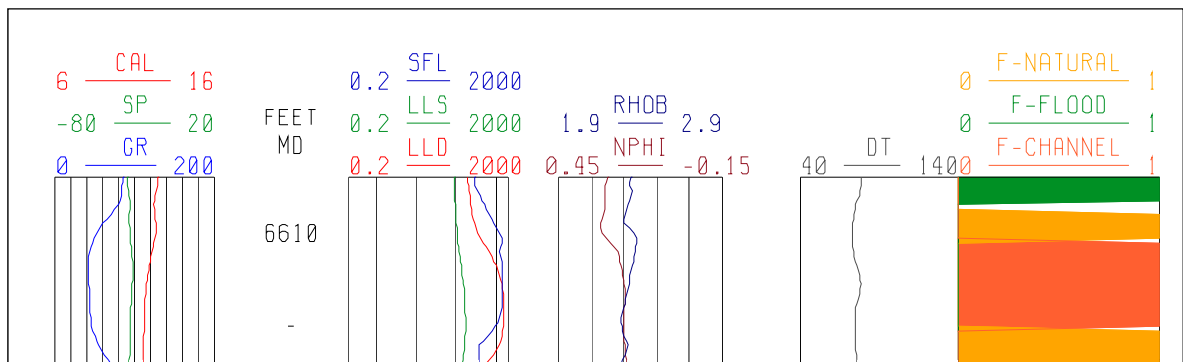


Figura 4.12 Registros eléctricos y a la derecha la curva generada para las electrofacies en profundidad.

A la izquierda y centro de la figura 4.12 se muestran los registros convencionales sobre los cuales se analizó la información y a la derecha de la misma figura se muestra una última casilla que aparece coloreada. En esta casilla se registra el tipo de electrofacies que presentaba cada intervalo.

Los intervalos coloreados de verde pertenecen a Floodplain Shale, los coloreados de amarillo corresponden a los Natural Levee, los intervalos coloreados de naranja corresponden a las arenas de canal y en los intervalos donde había ocurrencia de Crevasse Splays la curva presentaba un comportamiento anómalo con variaciones abruptas de valores muy bajos a valores muy altos y viceversa, para identificar estas facies fue necesario revisar los registros impresos y caracterizarlos de manera manual ya que el cambio continuo y abrupto de los valores en la curva GR no permitió generar un algoritmo con el cual el software identificara directamente este tipo de electrofacies..

Al tener la curva que identificaba en profundidad las diferentes electrofacies, fueron ploteados la totalidad de los registros disponibles para el campo. En estas nuevas sabanas ploteadas se apreciaban los registros eléctricos y la curva de las electrofacies en profundidad. Se hizo la revisión de los diferentes intervalos donde se tuviera algún tipo de incongruencia y se corrigieron.

Se modeló las facies para la unidad Guadalupe Masivo, teniendo como datos de entrada y control la curva de litofacies en cada pozo. En las facies de esta unidad, considerando que es principalmente arenoso (90%), se dejó como “background” ésta litofacies arenosa y se modelaron las litofacies arcillosas como cuerpos lenticulares. Al final se obtiene un volumen de facies integrado de la unidad.

En la figura 4.13 se ilustra una sección horizontal de la unidad Guadalupe Masivo, note la discontinuidad en los shales (color café).

En la figura 4.14 se muestra un corte transversal, en el se muestra algunos pozos con la curva litológica, el color amarillo corresponde a los cuerpos arenosas modelados

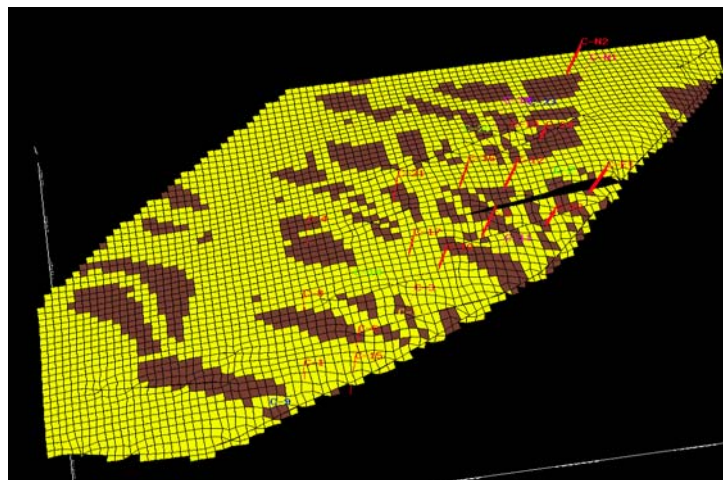


Figura 4.13 Sección horizontal unidad Guadalupe Masivo

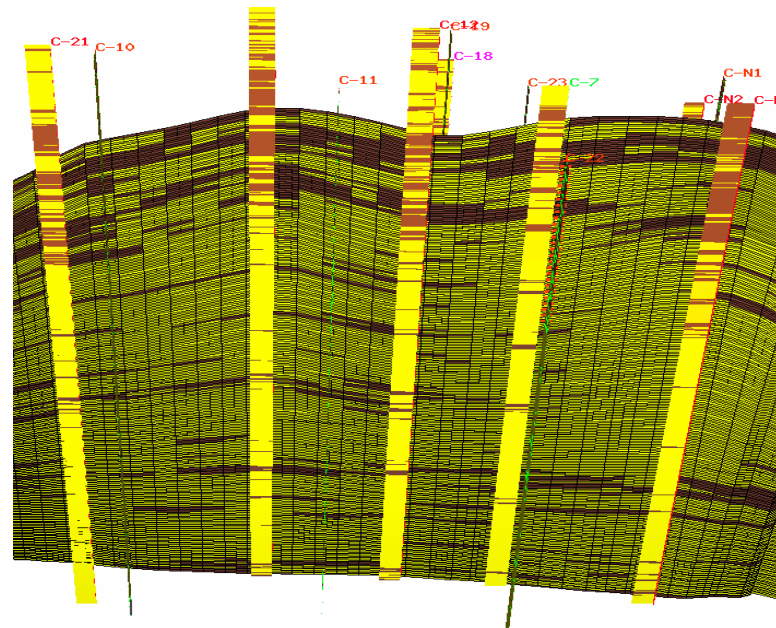


Figura 4.14 Corte transversal unidad Guadalupe Masivo

- **Modelamiento de parámetros petrofísicos**

Teniendo como base el volumen de facies y las curvas de porosidad y permeabilidad de los pozos, se modeló la distribución espacial de estos parámetros, a través de un modelamiento estocástico.

Inicialmente se modeló la porosidad y permeabilidad de manera determinística por interpolación. Este método se basa en los datos del pozo y a partir de estos se interpola mediante una función matemática para asignar valores a las celdas no penetradas por el pozo. El valor promedio de porosidad obtenido para esta unidad es de 19.5% y de permeabilidad de 1384 md.

Adicionalmente se realizó un modelamiento Petrofísico estocástico. Este método determina y predice la distribución de variables espacialmente de acuerdo con el grado de su correlación, dependiendo de la distancia entre ellas.

Inicialmente el procedimiento de este modelamiento estocástico, consiste en el escalamiento de la porosidad y permeabilidad de registros a la escala vertical del grid inicial.

Posteriormente se aplican transformadas para llevar a una distribución normal los parámetros escalados (porosidad y permeabilidad). Luego se crea un estimativo de la varianza contra la distancia entre dos puntos medidos en los datos del pozo. Con el estimativo anterior, se crea una función de variograma, el cual ajusta los puntos estimados anteriormente, en este paso se aplican rangos de anisotropía, luego se crea una matriz de correlación de los parámetros a modelar. Por último se selecciona el método de simulación estocástica, este software (sgsim) presenta varias opciones: Kriging simple, Kriging multivariado y Kriging ordinario; para este trabajo se usó el último, por ser este, el que produce el modelo mas suavizado.

4.4.1. Geometría y convenciones. La ingeniería de yacimientos emplea modelos matemáticos para describir el comportamiento de los fluidos dentro de la roca. Estos modelos requieren de la definición de un enmallado tridimensional, donde las propiedades promedio de la roca son asignadas a cada celda. Cada una de estas celdas en el modelo espacial debe tener la capacidad de representar el componente geológico y las propiedades petrofísicas específicas del yacimiento.

Para que el modelamiento de las propiedades involucradas tenga la capacidad de conjugar la geología y la ingeniería de yacimientos es necesario definir tres enmallados desde el inicio del estudio.

Grid de Referencia

El grid de referencia es el enmallado principal empleado en el modelamiento geoestadístico. Es un grid en tres dimensiones de celdas ortorrómbicas dentro del

cual está descrito el yacimiento y en el cual todas las celdas son idénticas en tamaño. El grid está definido por su origen, las dimensiones de las celdas y el número de estas a lo largo de cada dirección. Este grid proporciona un sistema de referencia de coordenadas para los cálculos efectuados dentro y fuera del paquete de modelamiento geostadístico (Fig. 4.15). Allí se encuentran las coordenadas de origen y final, así como también las dimensiones de cada una de las celdas.

Para el proyecto, las dimensiones para el grid de referencia serán:

Nx = 63	Dx = 0.5	Zx = 10
Ny = 161	Dy = 0.5	Zy = 10
Nz = 215	Dz = 0.5	Zz = 10

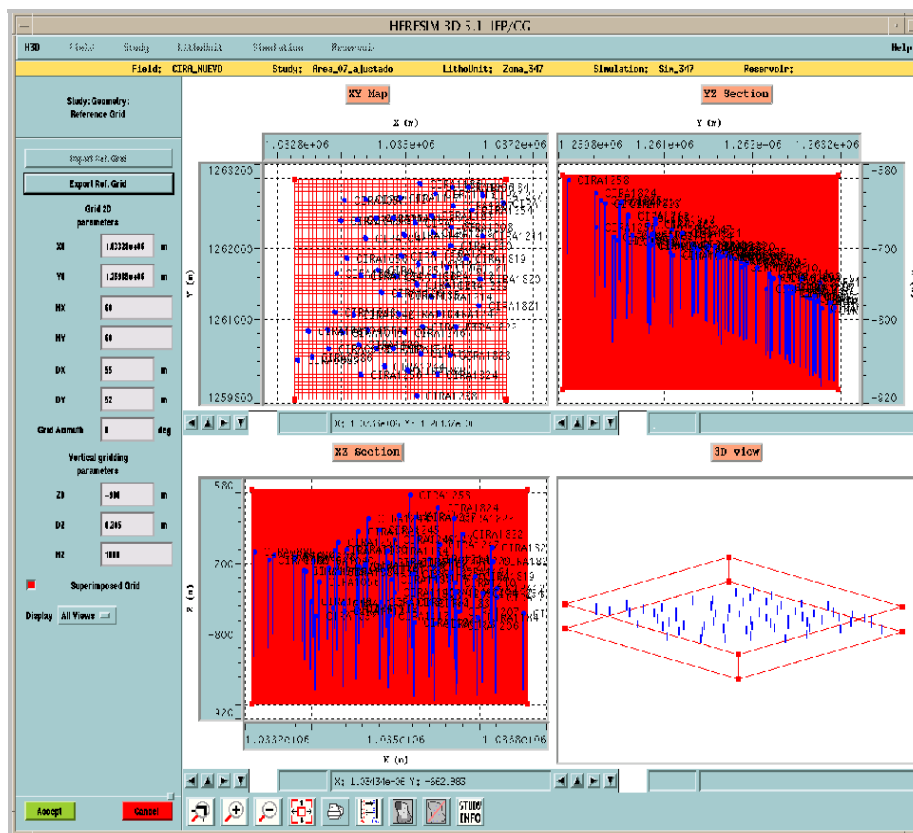


Figura 4.15. Grid de Referencia

La simulación resultante (calculada sobre el grid geoestadístico) se enmarca dentro de este grid de referencia.

Grid Geoestadístico

Este *grid* es construido para todo lo referente a cálculos estadísticos. El *grid* Geoestadístico y el *grid* de referencia tienen la misma sección XY. La diferencia entre estos dos enmallados es el significado físico de las capas y la distribución vertical.

Para el proyecto, las dimensiones para el grid geoestadístico serán:

Nx = 63	Dx = 0.5	Zx = 10
Ny = 161	Dy = 0.5	Zy = 10
Nz = 215	Dz = 0.5	Zz = 10

Cálculos tales como el enmallado vertical de los pozos, las curvas de proporción vertical y horizontal, variogramas y simulaciones, son efectuados dentro de este *grid*. Las capas de este enmallado son consideradas como líneas de correlación, estas representan el modo de depositación de la unidad litoestratigráfica. Dependiendo del modelo de depositación, compactación y erosión de la arquitectura interna de la formación, el enmallado geoestadístico se define como proporcional o paralelo. Para la ejecución de este proyecto se definió como proporcional debido a las inferencias conceptuales del modelo estratigráfico (figura 4.16). La unidad litoestratigráfica es dividida en un número constante de capas, donde cada línea de correlación es un límite de celda.

Un enmallado paralelo se considera para los casos en los cuales no hay ninguna diferencia en la compactación de las capas. En tal caso todas las capas del *grid* de simulación son paralelas a la superficie de referencia.

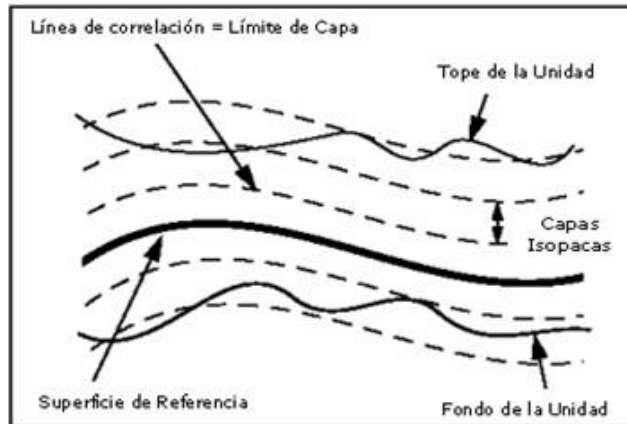


Figura 4.16. Definición del Grid Paralelo

Un enmallado proporcional se usa cuando existe una diferencia de compactación entre las capas. La superficie de referencia no se define y la unidad es dividida en un número constante de capas. Las líneas de correlación son las capas límites (tope y fondo), como se muestra en la Figura 4.17.

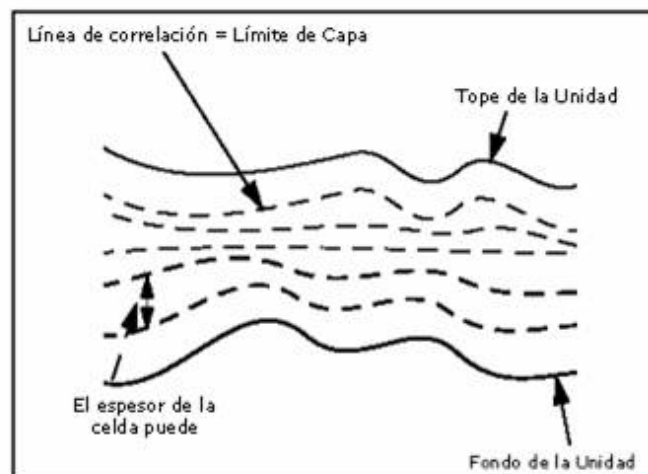


Figura 4.17. Definición del Grid Proporcional

Grid Simulación y Orientación

El *grid* de simulación se usa para el modelamiento de flujo de fluidos. Se considera una versión simplificada del *grid* geoestadístico. Las capas del *grid* geoestadístico son combinadas con las unidades de flujo, considerando las diferencias verticales de porosidad y permeabilidad. En ocasiones es necesario considerar un *grid* semi-regular con refinamientos locales, debido a la complejidad del yacimiento así como de la ubicación y estado de los pozos.

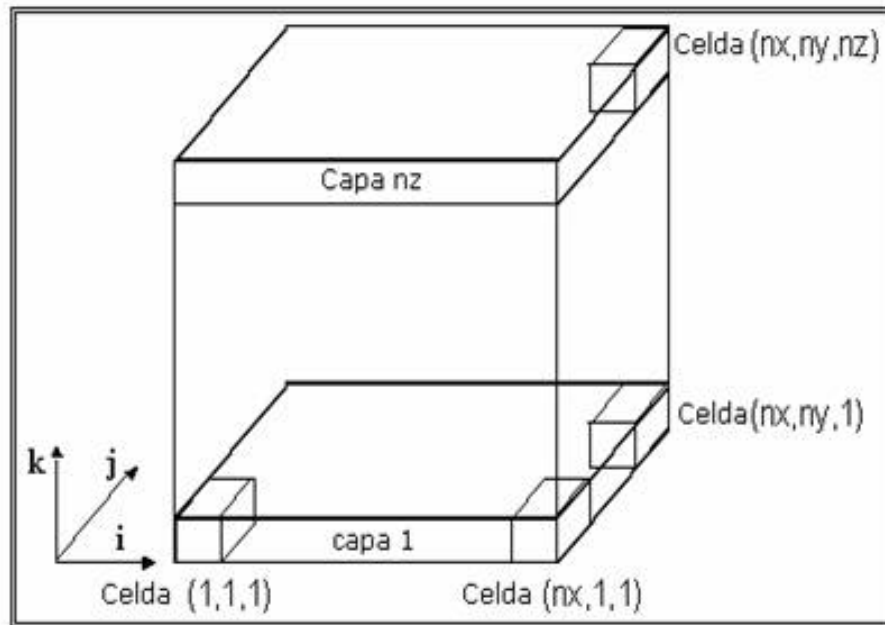


Figura 4.18. Convención de Orientación del Grid

Es muy importante tener en cuenta la orientación del modelo puesto que el conjunto de datos generados están referenciados con base a ésta, presentándose diferencias en la convención de las herramientas usadas en el proyecto.

4.5. ANALISIS GEOESTADISTICO

En esta sección se describe la forma como se procesaron los datos para obtener el modelo Geoestadístico que se utilizó en la generación de las realizaciones del área seleccionada.

Variogramas

Los variogramas se usan para caracterizar el modelo de distribución espacial a lo largo de la unidad de estudio, se deben generar variogramas horizontales y verticales. Los variogramas verticales son calculados a lo largo de la dirección vertical del enmallado de simulación. Los variogramas horizontales son calculados capa por capa, pero el resultado proporciona un variograma promedio para todas las capas del enmallado de simulación incluidas en el dominio de cálculo.

Los variogramas calculados en las simulaciones deben ser lo más aproximados posibles al variograma experimental calculado con los datos reales del pozo. Ya que el algoritmo de simulación requiere un modelo de variograma (una curva teórica con ecuación conocida), es necesario realizar un montaje antes de realizar las simulaciones. Esta operación consiste en elegir el variograma que mejor se ajuste al variograma experimental.

- El método de ajuste es basado en la comparación visual de dos curvas, el variograma experimental y el teórico. Estas curvas deben ser en lo posible similares.

4.5.1. Construcción del modelo geoestadístico. En esta sección, se presentan los variogramas obtenidos. Tal como se mencionó anteriormente, la información

de los pozos, tanto de porosidad como de permeabilidad es cargada y arreglada con el fin de evitar incongruencias con el simulador.

Se construyeron variogramas experimentales con el software “GAMV” en las cuatro (4) direcciones horizontales más importantes (0°, 45°, 90° y 135°) y se utilizó el visualizador “VARGPLT” para obtener la imagen de dichos variogramas, adicionalmente, se hizo uno vertical. A cada uno de estos variogramas experimentales se le hizo sensibilidad en las diferentes variables involucradas (tolerancia, ancho de banda y dip) hasta obtener el modelo con mejor tendencia. A continuación se presentan los variogramas tanto de permeabilidad como de porosidad y la ventana de entrada de los datos del software.

```
1 Parameters for GAMV
2 *****
3
4 START OF PARAMETERS:
5 Poro0_1_ny_1_K2.dat \file with data
6 1 2 3 \ columns for X, Y, Z coordinates
7 1 5 \ number of variables, column numbers
8 -1.0e21 1.0e21 \ trimming limits
9 gamvverticalPO.out \file for variogram output
10 10 \number of lags
11 5.0 \lag separation distance
12 3.0 \lag tolerance
13 5 \number of directions
14 0.0 45.0 30.0 60.0 30.0 50.0 \azm, atol, bandh, dip, dtol, bandv
15 90.0 40.0 60.0 60.0 60.0 60.0 \azm, atol, bandh, dip, dtol, bandv
16 0.0 22.5 10.0 90.0 30.0 10.0
17 45.0 20.0 20.0 45.0 50.0 22.5
18 135.0 40.0 20.0 45.0 60.0 30.0
19 1 \standardize sills? (0=no, 1=yes)
20 1 \number of variograms
21 1 1 1 \tail var., head var., variogram type
```

Figura 4.19. Software GAMV

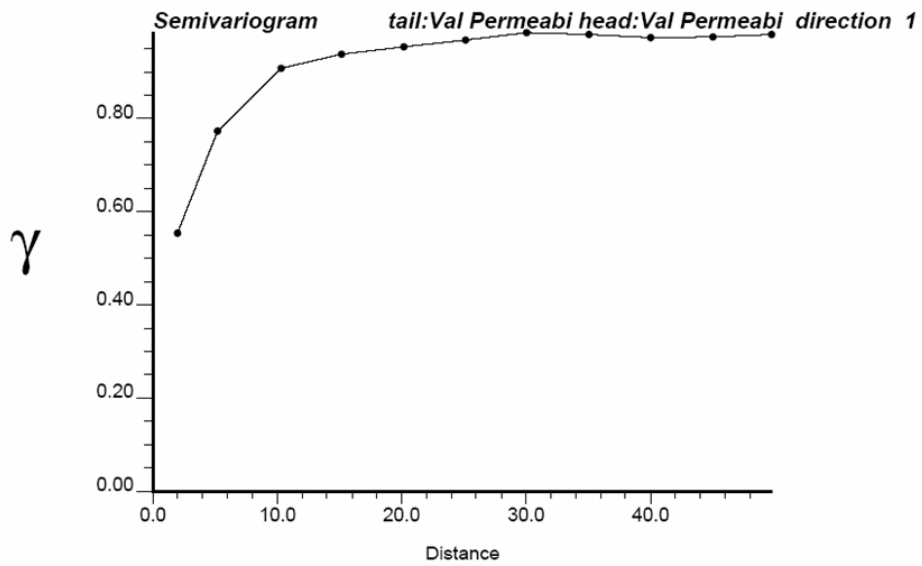


Figura 4.20. Variograma Horizontal De Permeabilidad En Dirección 0°

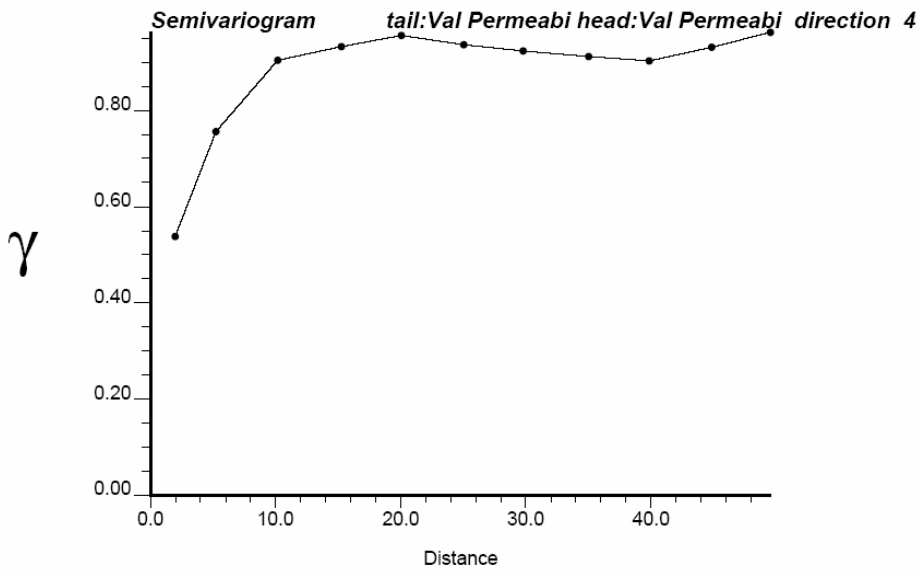


Figura 4.21. Variograma Horizontal De Permeabilidad En Dirección 45°

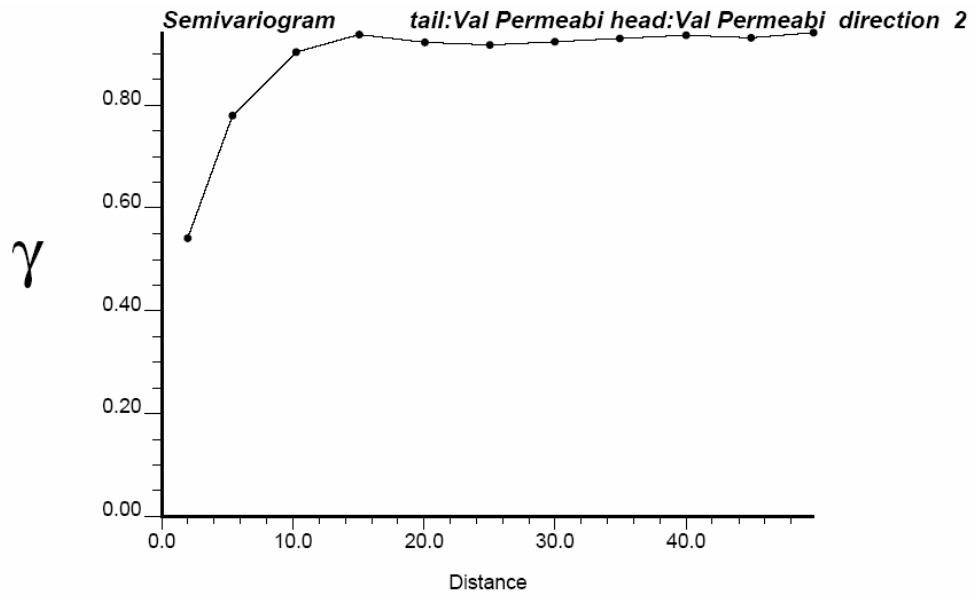


Figura 4.22. Variograma Horizontal De Permeabilidad En Dirección 90°

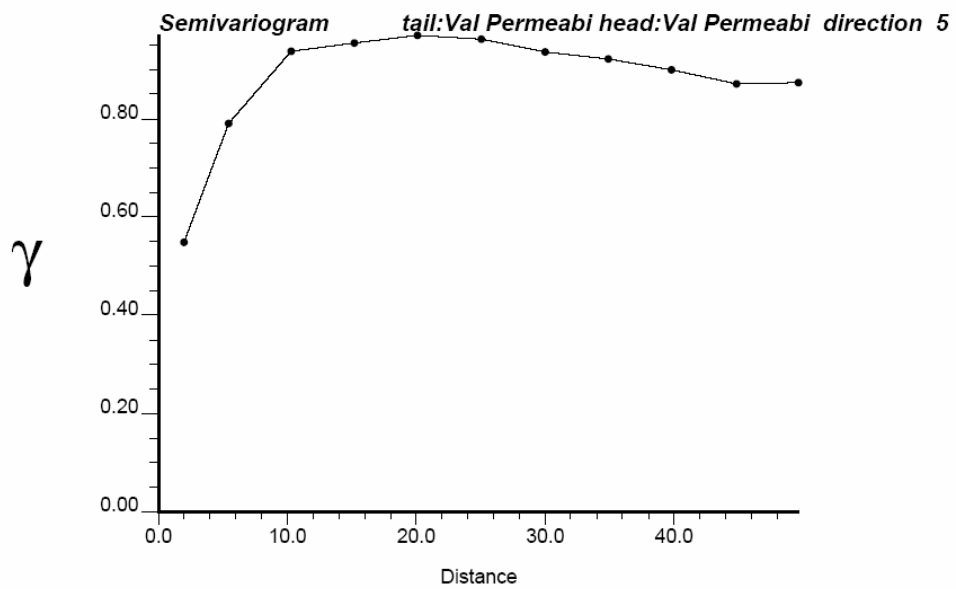


Figura 4.23. Variograma Horizontal De Permeabilidad En Dirección 135°

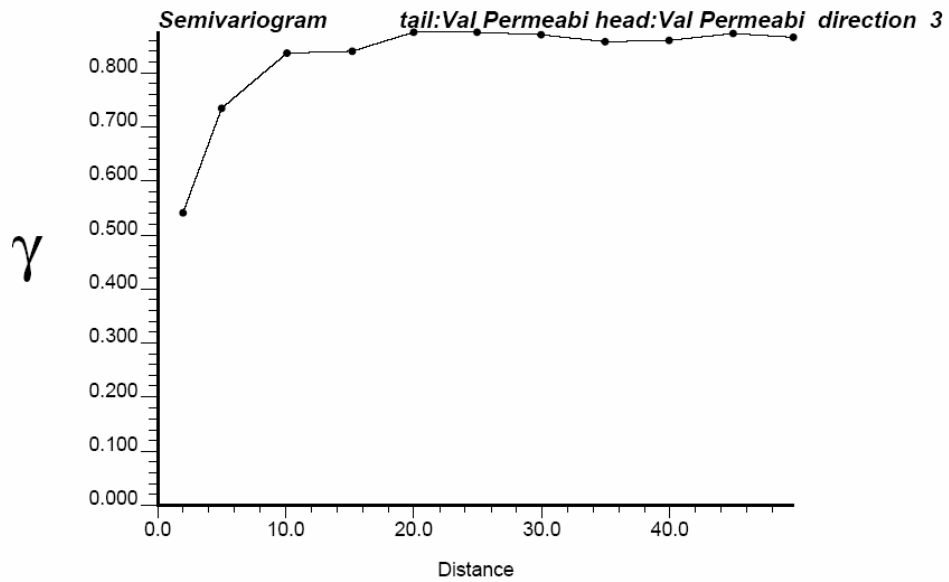


Figura 4.24. Variograma Vertical De Permeabilidad

La figura 4.25 agrupa los cinco modelos de permeabilidad en una sola grafica, en la cual se aprecia la presencia de anisotropía Geométrica.

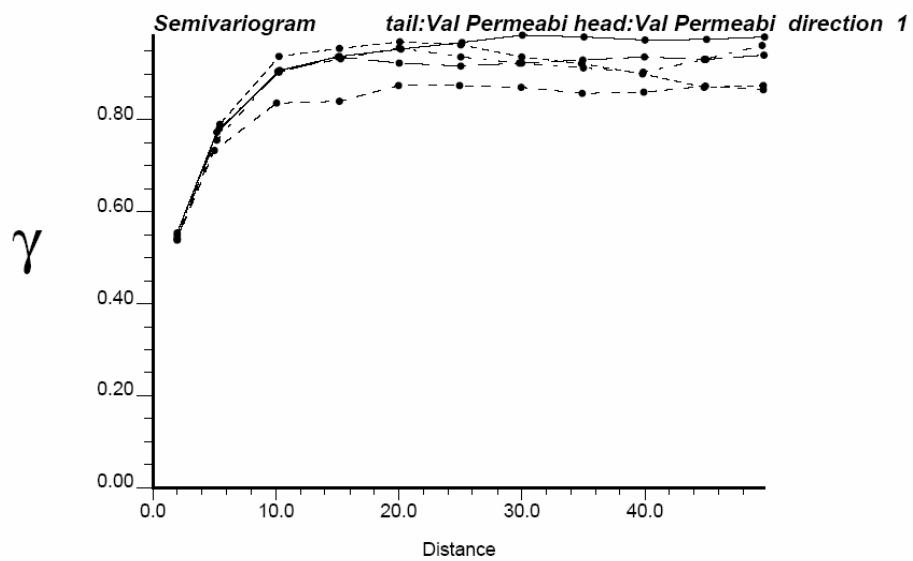


Figura 4.25 Variogramas De Permeabilidad

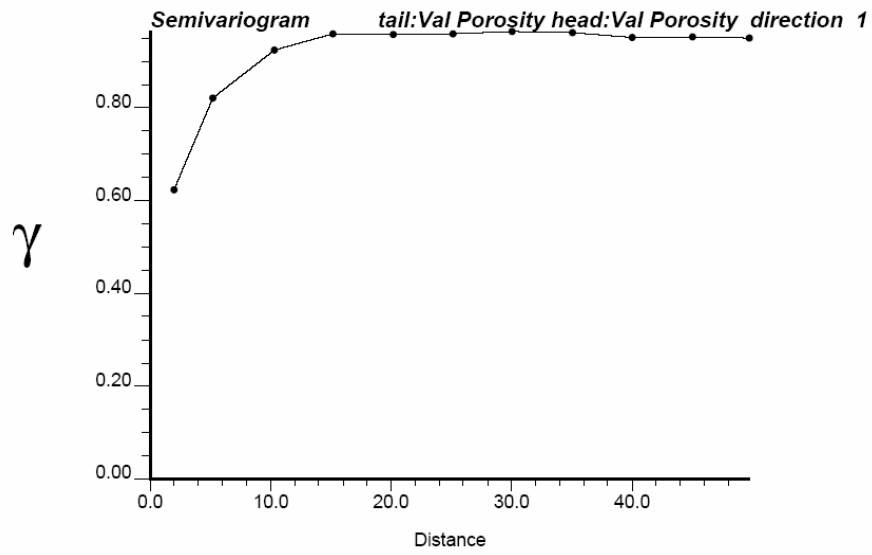


Figura 4.26. Variograma Horizontal De Porosidad En Dirección 0°

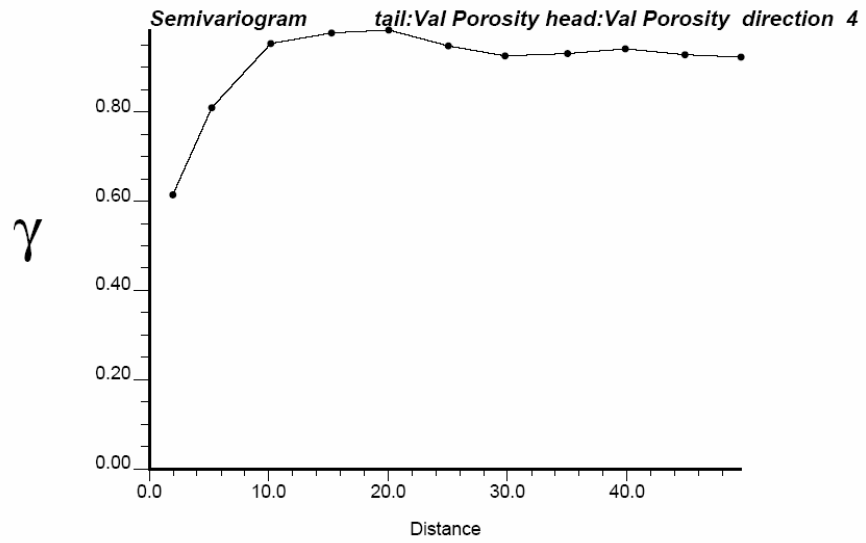


Figura 4.27. Variograma Horizontal De Porosidad En Dirección 45°

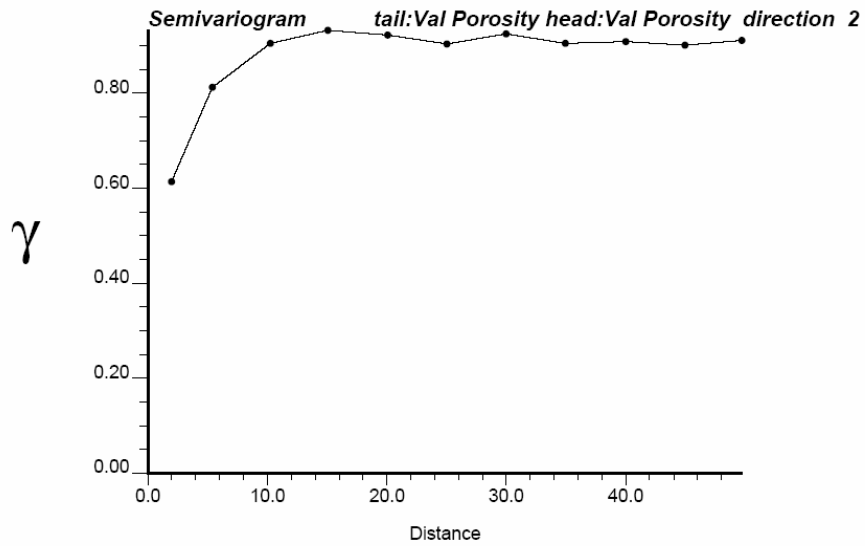


Figura 4.28. Variograma Horizontal De Porosidad En Dirección 90°

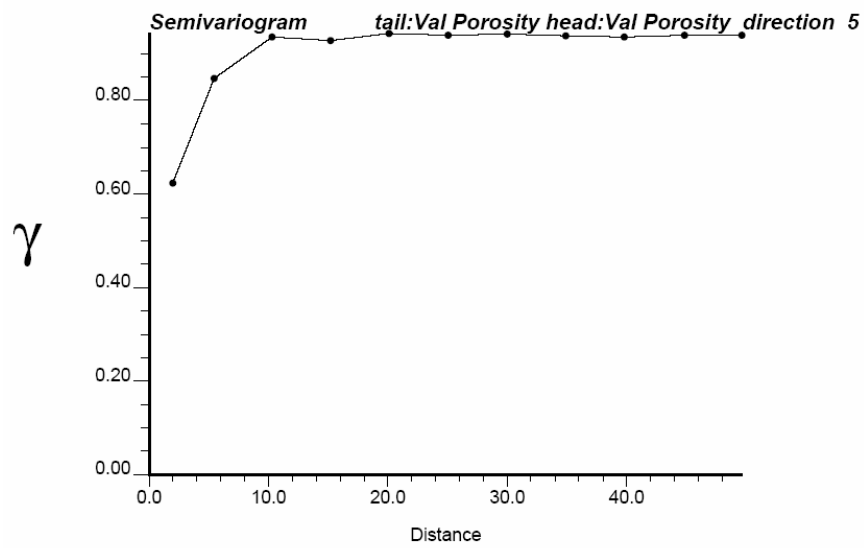


Figura 4.29. Variograma Horizontal De Porosidad En Dirección 135°

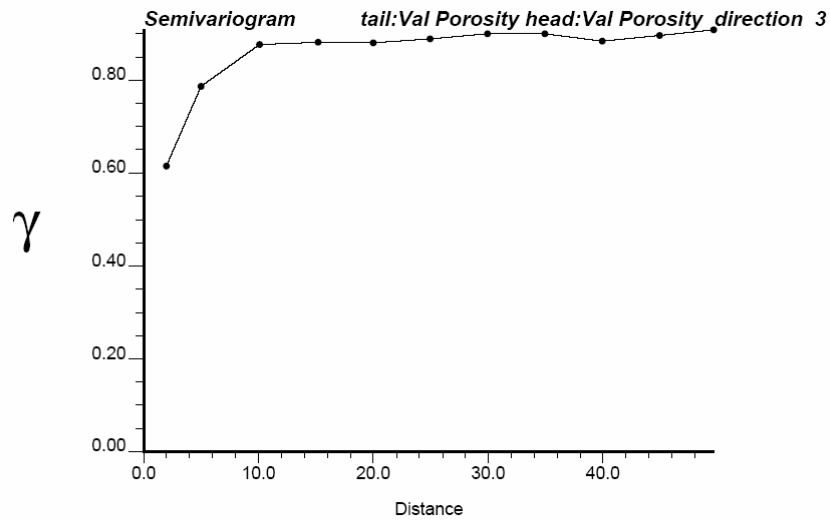


Figura 4.30. Variograma Vertical De Porosidad

De igual manera que en el caso de permeabilidad, la figura 4.31 agrupa los cinco modelos de porosidad en una sola gráfica, en la cual se aprecia la presencia de anisotropía Geométrica a menor escala. Lo cual indica una mejor tendencia para la porosidad.

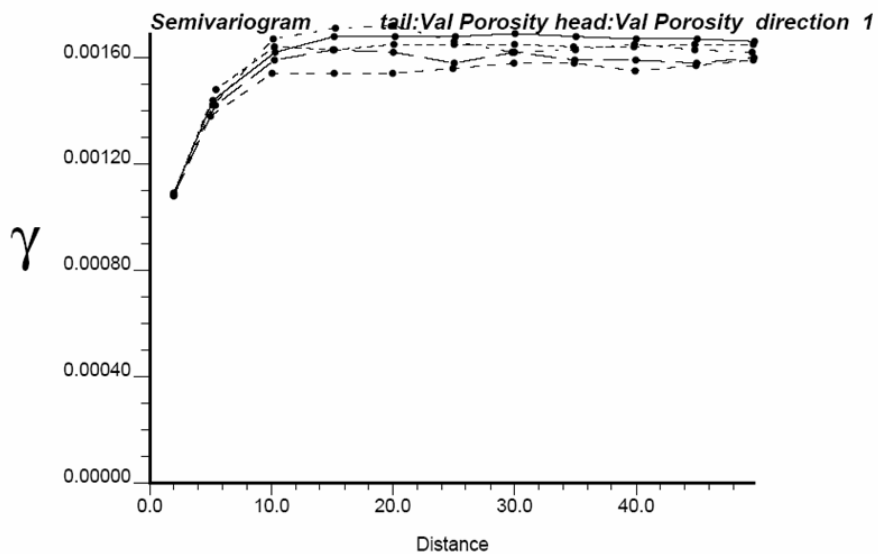


Figura 4.31 Variogramas De Porosidad

Con la información de rangos, sill y efecto nugget de los diferentes variogramas se construirá el elipsoide de anisotropía. Luego de obtener dicho elipsoide se correrá el “SGSIM”.

Los valores relacionados en la Tabla 4.6 y 4.7 se graficaron para evaluar la anisotropía del modelo, de esta manera se obtuvieron los elipsoides de permeabilidad y porosidad que se presenta en la Figura 4.35 y 4.36 respectivamente.

VARIOGRAMA PERMEABILIDAD			
Anisotropía	Dirección	Rango	Sill
	0°	20	0.95
	45°	20	0.95
	90°	25	0.9
	135°	45	0.87
	Vertical	20	0.86

Tabla 4.6. Interpretación Variogramas Permeabilidad

Los resultados de esta evaluación se pueden observar en la Figura 4.32, donde se evidencia que la dirección de mayor correlación tiene 135° (315° de azimut), la longitud del eje en esta dirección es de 88 m., mientras que el eje de menor correlación tiene una longitud de 35 m; en cuanto el rango vertical, este se mantiene igual al interpretado inicialmente.

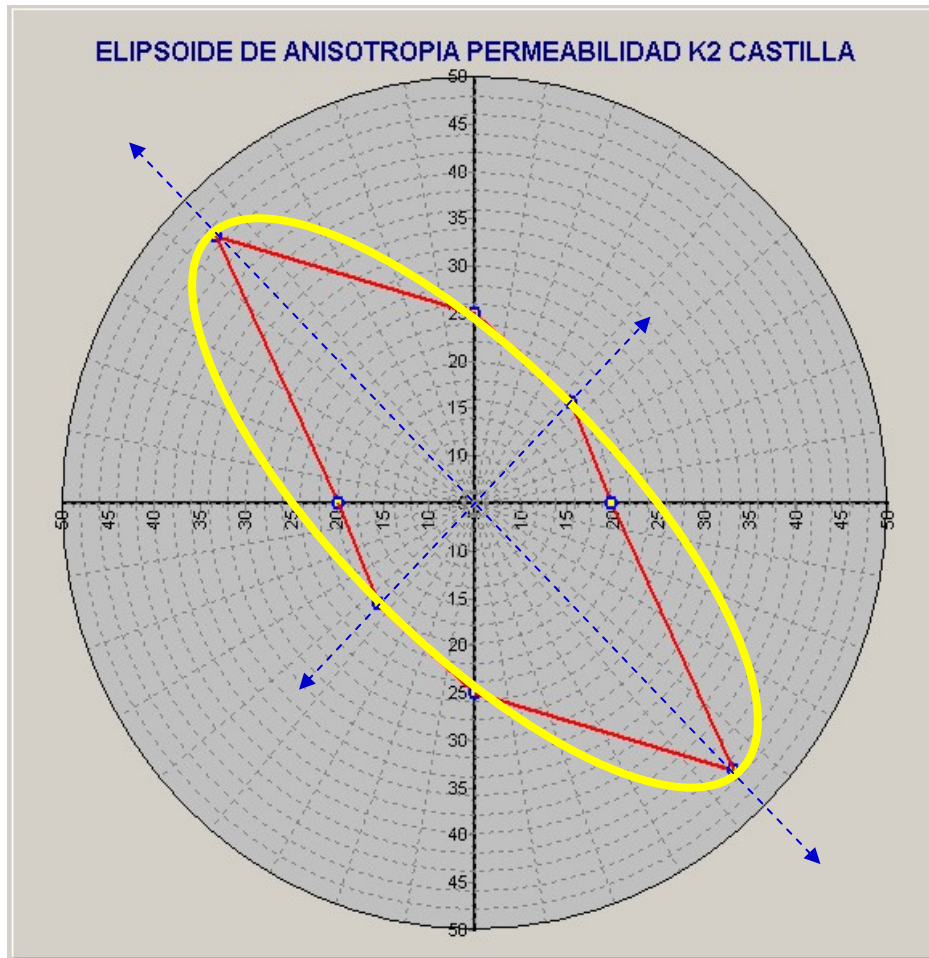


Figura 4.32 Elipse De Anisotropía

VARIOGRAMA POROSIDAD			
	Dirección	Rango	Sill
Anisotropía	0°	15	0.95
	45°	30	0.93
	90°	34	0.9
	135°	20	0.92
	Vertical	10	0.87

Tabla 4-7. Interpretación Variogramas Porosidad

En la Figura 4.33 se evidencia que la dirección de mayor correlación tiene 70° (20° de azimut), la longitud del eje en esta dirección es de 78 m., mientras que el eje de menor correlación tiene una longitud de 34 m; en cuanto el rango vertical, este se mantiene igual al interpretado inicialmente.

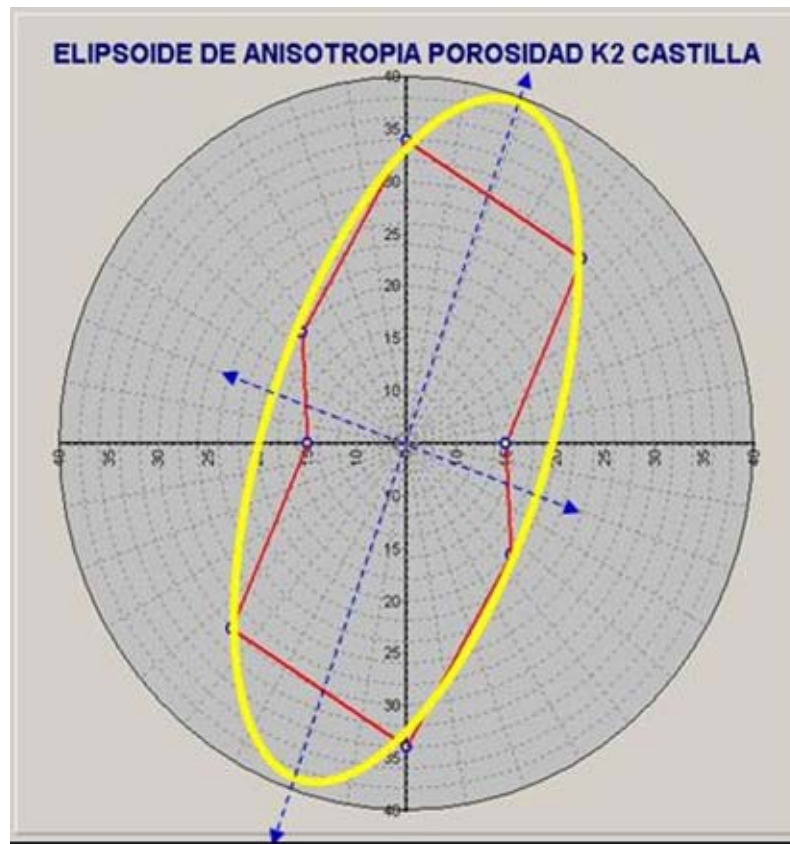


Figura 4.33 Elipse De Anisotropía

4.5.2. Simulación petrofísica. Una vez definido el elipsoide de anisotropía, se logró completar la información necesaria para ejecutar el software geoestadístico y así generar la simulación petrofísica. En la Figura 4.34 se muestra los datos de entrada del simulador.

```

1 Parameters for SGSIM
2 *****
3
4 START OF PARAMETERS:
5 Poro0_1_ny_1_K2.dat          \file with data
6 1 2 3 5 0 0                 \ columns for X,Y,Z,vr,wt,sec.var.
7 -1.0      1.0e21             \ trimming limits
8 1                             \transform the data (0=no, 1=yes)
9 sgsim.trn                    \ file for output trans table
10                             \ consider ref. dist (0=no, 1=yes)
11 histsmth.out                \ file with ref. dist distribution
12 5 0                          \ columns for vr and wt
13 0.0001  0.5                 \ zmin,zmax(tail extrapolation)
14 1      0.0001               \ lower tail option, parameter
15 1      0.5                   \ upper tail option, parameter
16 0                             \debugging level: 0,1,2,3
17 sgsim17a.dbg                \file for debugging output
18 simulacionPOROK2.out        \file for simulation output
19 1                             \number of realizations to generate
20 63  0.5  10.0               \nx,xmn,xsiz
21 161  0.5  10.0              \ny,ymn,ysiz
22 215  0.5  10.0              \nz,zmn,zsiz
23 69071                         \random number seed
24 0  8                          \min and max original data for sim
25 12                             \number of simulated nodes to use
26 1                             \assign data to nodes (0=no, 1=yes)
27 1  3                          \multiple grid search (0=no, 1=yes),num
28 0                             \maximum data per octant (0=not used)
29 78.0 34.0 10.0              \maximum search radii (hmax,hmin,vert)
30 45.0 90.0 0.0 135.0         \angles for search ellipsoid
31 1                             \ktype: 0=SK,1=OK,2=LVM,3=EXDR,4=COLC
32 ../data/ydata.dat           \ file with LVM, EXDR, or COLC variable
33 0                             \ column for secondary variable
34 1 0.5                         \nst, nugget effect
35 1  0.95 0.0 60.0 0.0        \it,cc,ang1,ang2,ang3
36      34.0 10.0 10.0         \a_hmax, a_hmin, a_vert

```

Figura 4.34. Software SGSIM

Con la salida del “sgsim” se obtuvieron las distribuciones petrofísicas tanto de permeabilidad como de porosidad para la unidad K2 del campo castilla, mediante el visualizador “PIXELPLT”. Las cuales fueron de poco interés debido a la incapacidad que tiene esta herramienta a la hora de ubicar los pozos. Para tal fin fue necesario trabajar con un nuevo visualizador llamado “VIEW.EXE”, el cual fue

programado en lenguaje Visual C++ y VTK (Visualización Tool Kit), siendo este ultimo un Kit de herramientas de programación grafica tridimensional para la visualización en modalidad realidad virtual de los modelos. .

A continuación se muestran las distribuciones petrofísicas resultantes de permeabilidad y porosidad respectivamente.

5. RESULTADOS

5.1 PERMEABILIDAD

En la Figura 5.1 se muestra un corte en 3D de la permeabilidad de la unidad K2 del campo castilla utilizando la metodología planteada. En esta se aprecia la gran calidad de las arenas con la cual cuenta este campo.

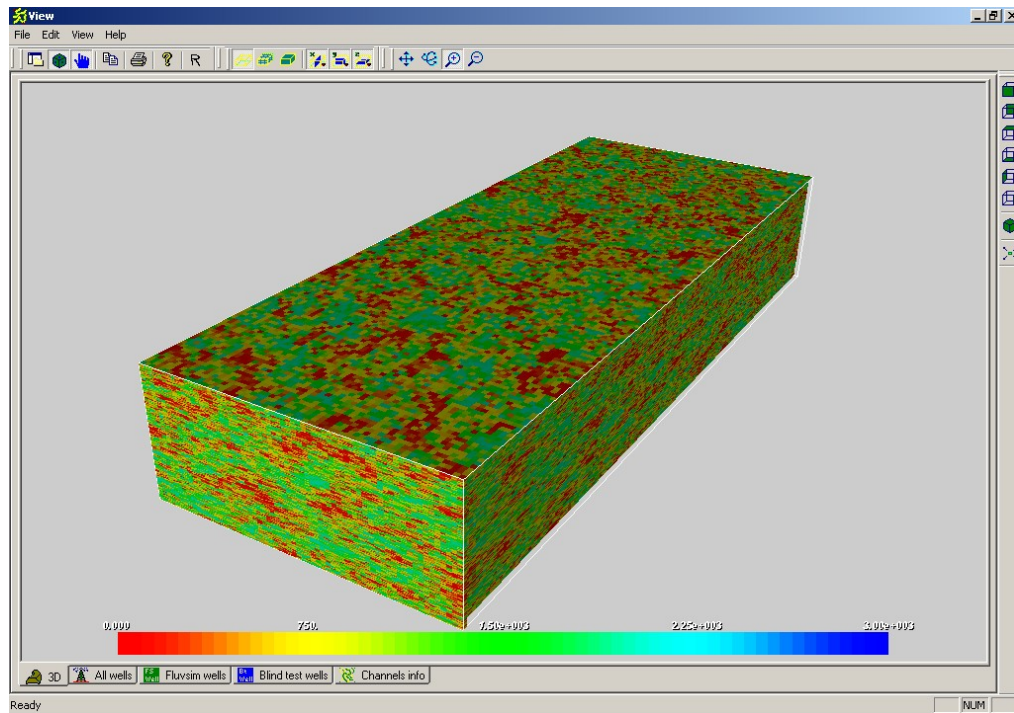


Figura 5.1 Modelo resultante de Permeabilidad

Como se puede apreciar encontramos permeabilidades que van desde cero hasta 2000 md. Pero prevalecen valores mayores a 500 md.

A continuación se muestran varias figuras que representan las diferentes vistas (frontal, lateral y en superficie) obtenidas con el modelamiento geoestadístico, las cuales se comparan con los resultados obtenidos por ECOPETROL en julio del presente año.

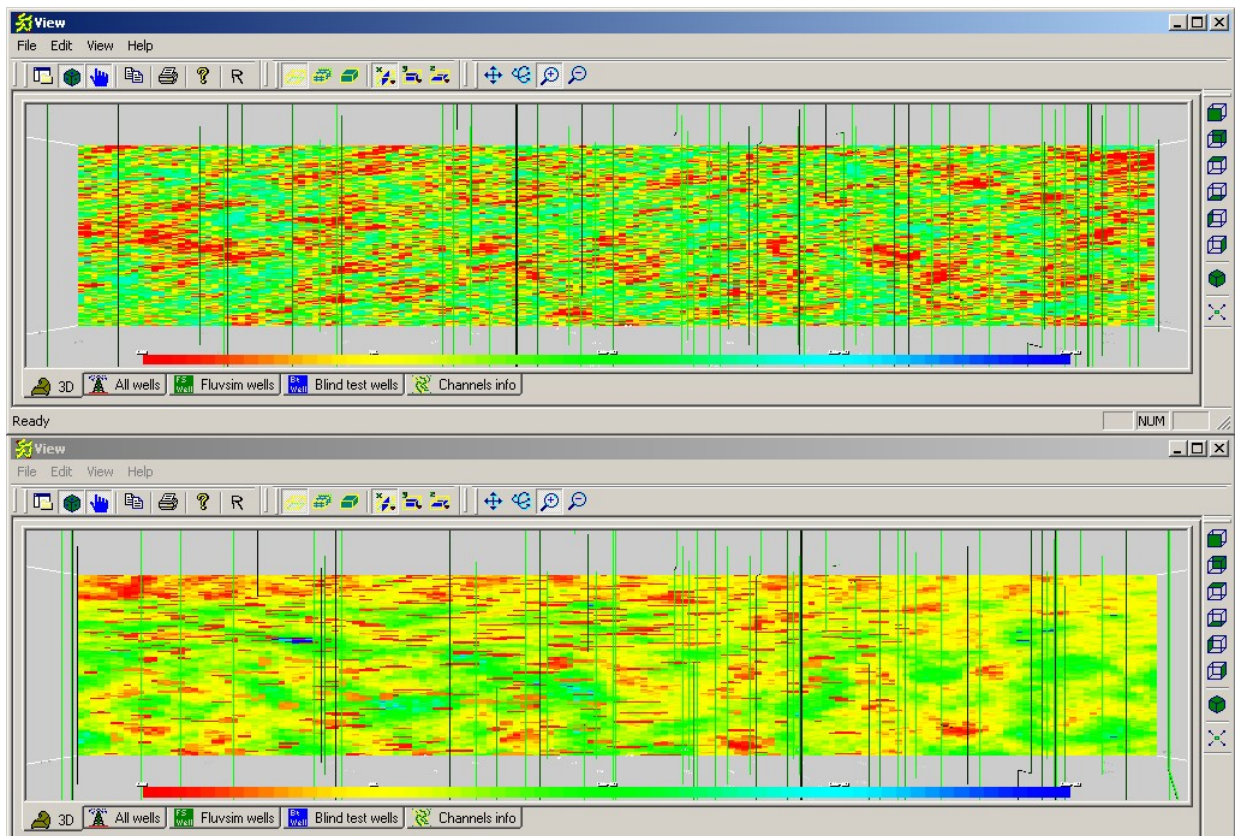


Figura 5.2 Vista lateral de Permeabilidad. Layer 1

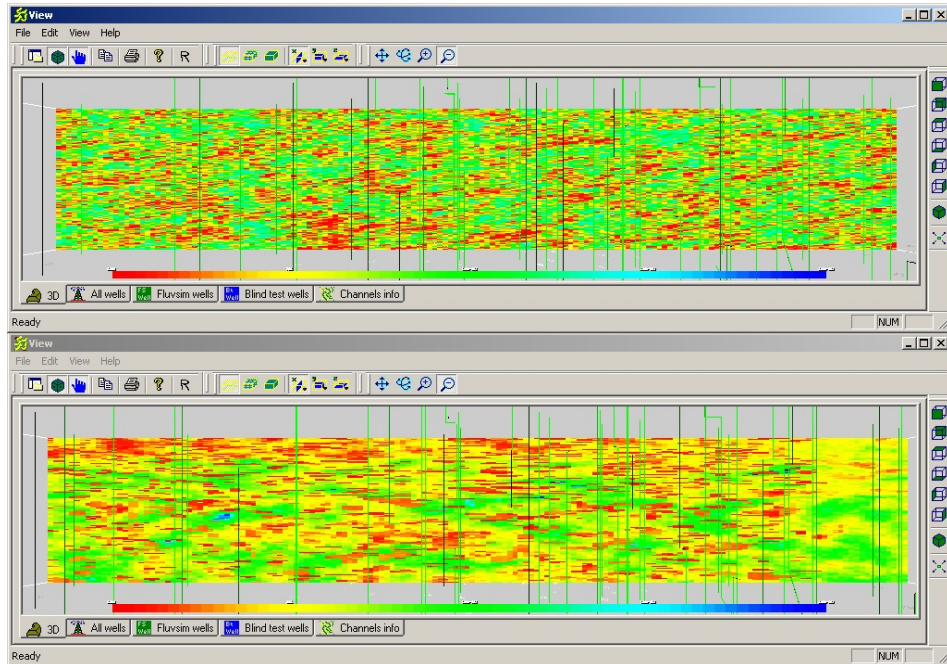


Figura 5.3 Vista lateral de Permeabilidad. Layer 10

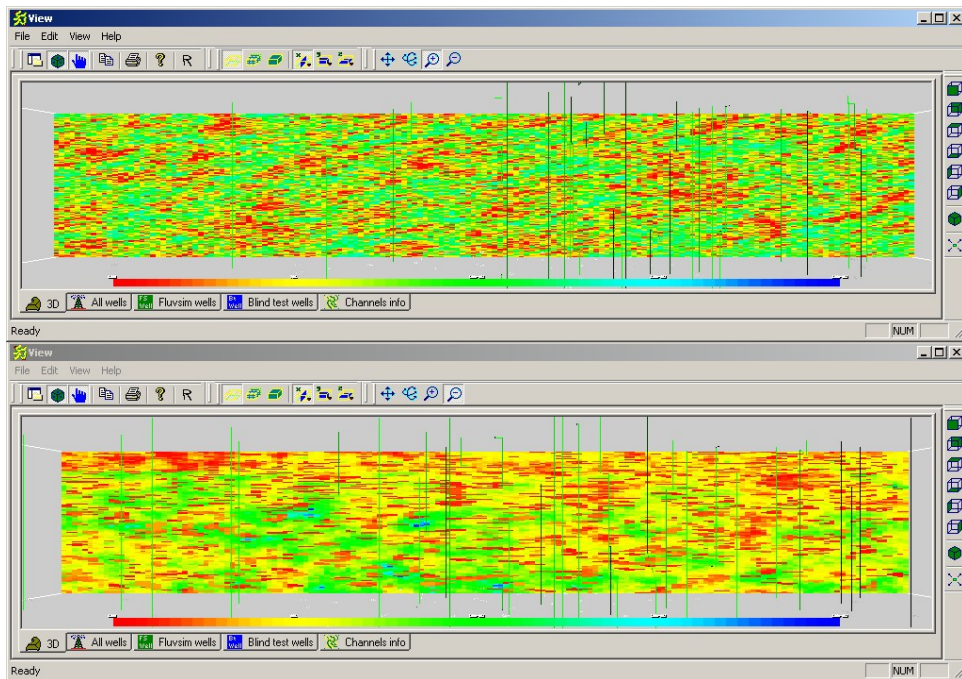


Figura 5.4 Vista lateral de Permeabilidad. Layer 30

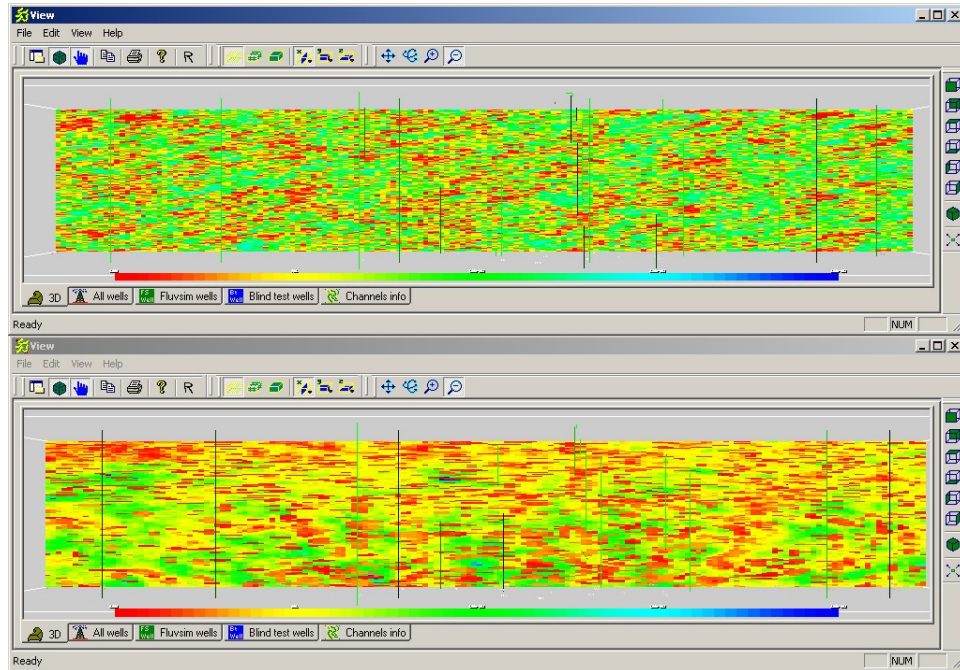


Figura 5.5 Vista lateral de Permeabilidad. Layer 45

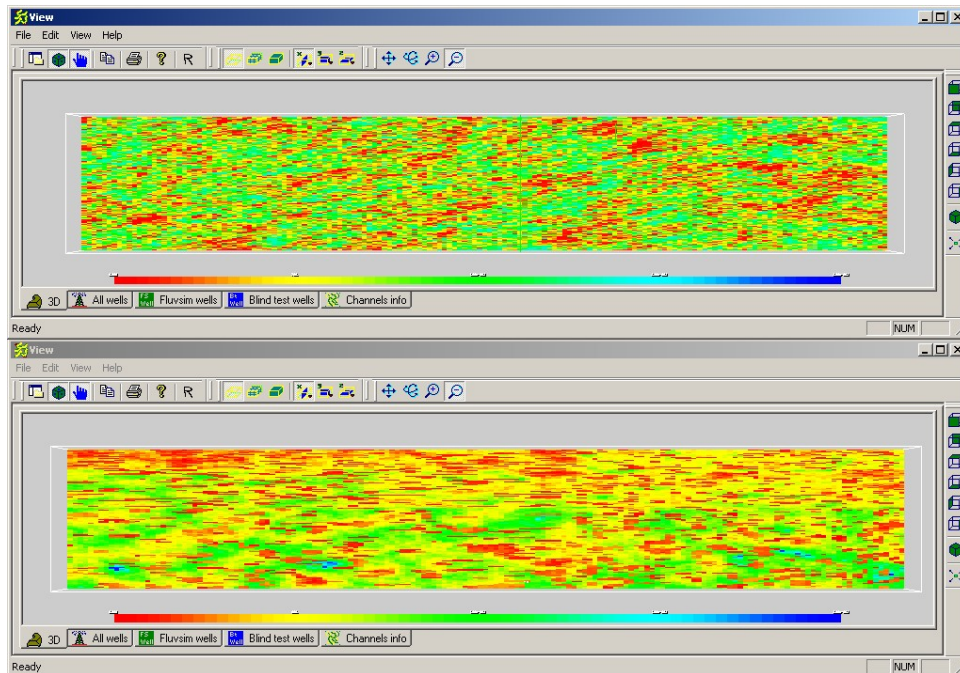


Figura 5.6 Vista lateral de Permeabilidad. Layer 60

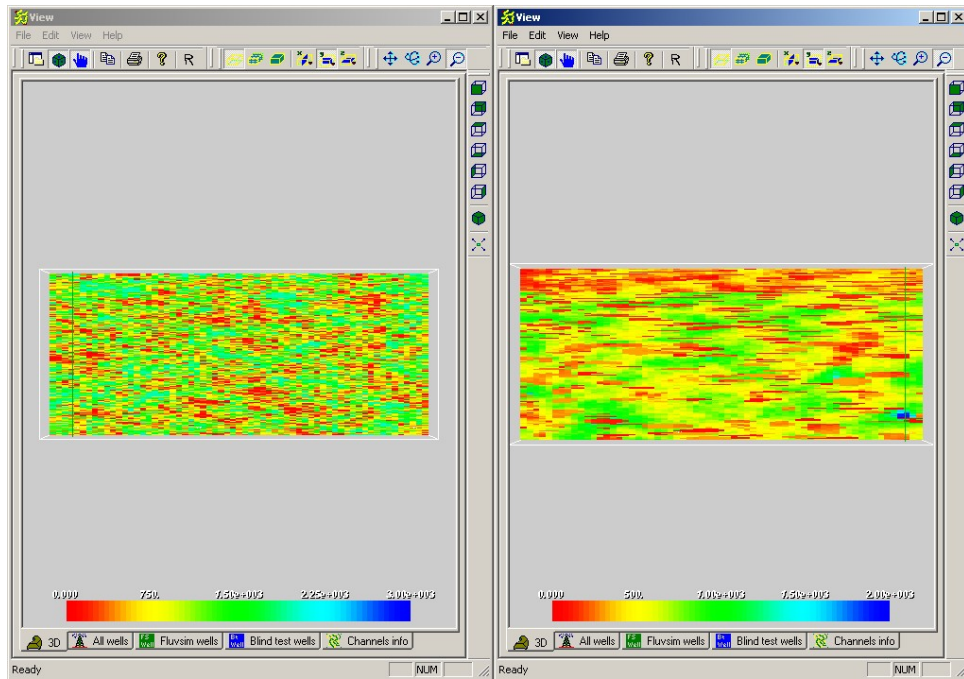


Figura 5.7 Vista frontal de Permeabilidad. Layer 10

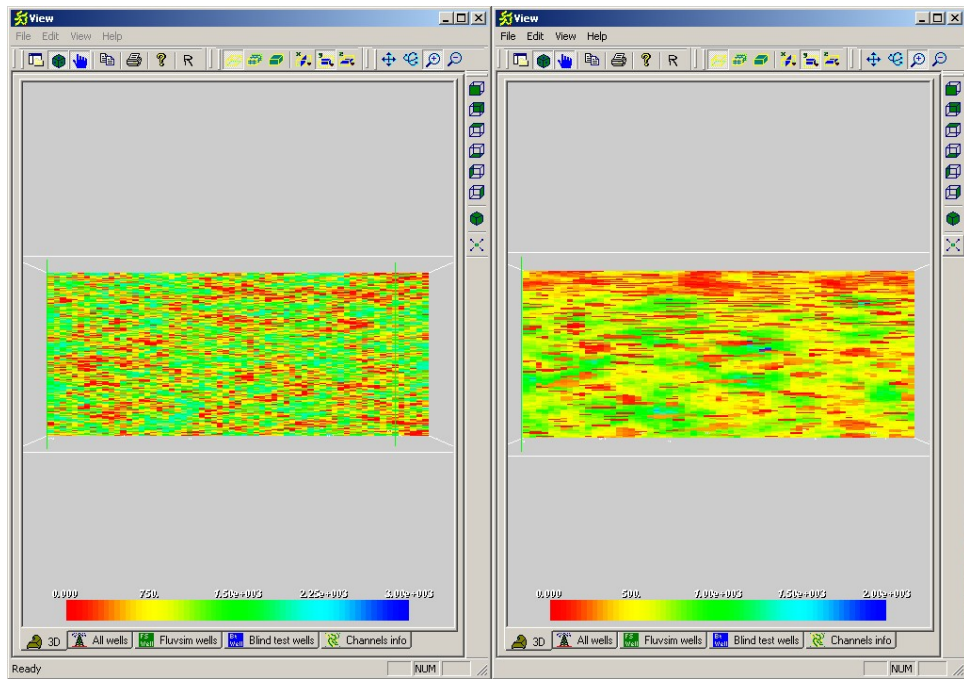


Figura 5.8 Vista frontal de Permeabilidad. Layer 30

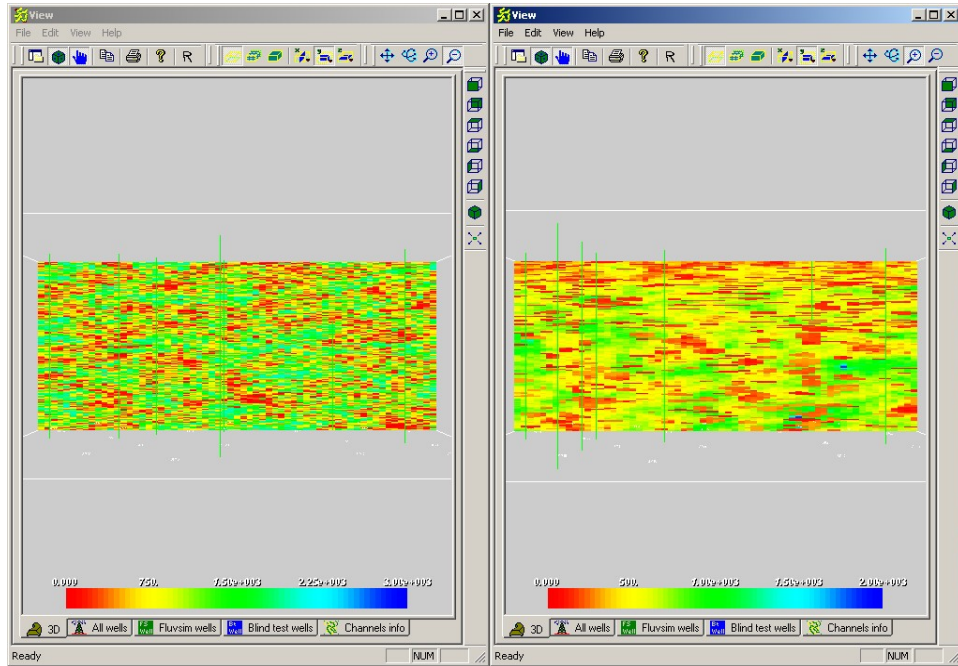


Figura 5.9 Vista frontal de Permeabilidad. Layer 60

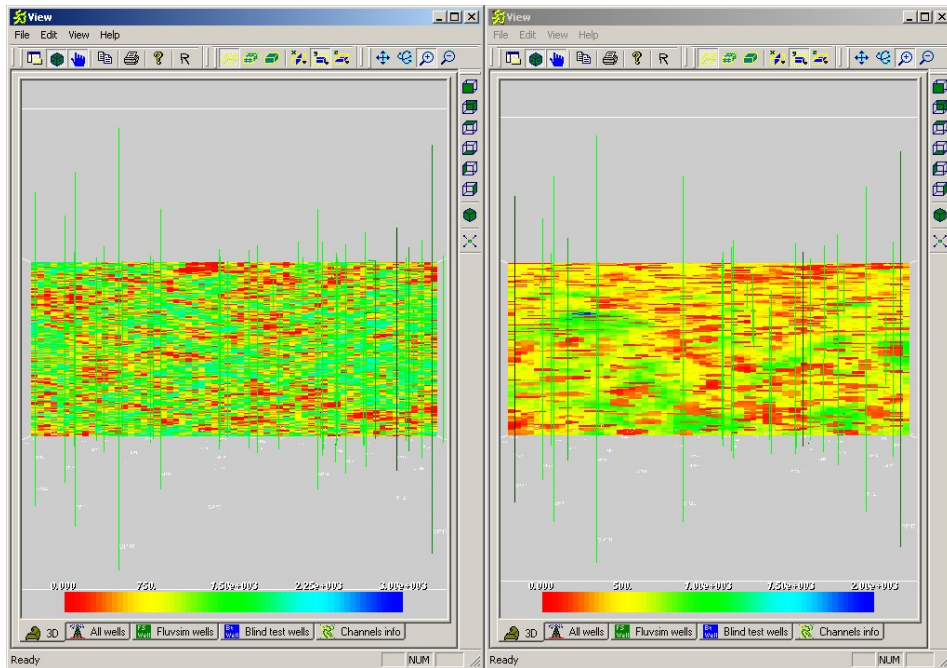


Figura 5.10 Vista frontal de Permeabilidad. Layer 100

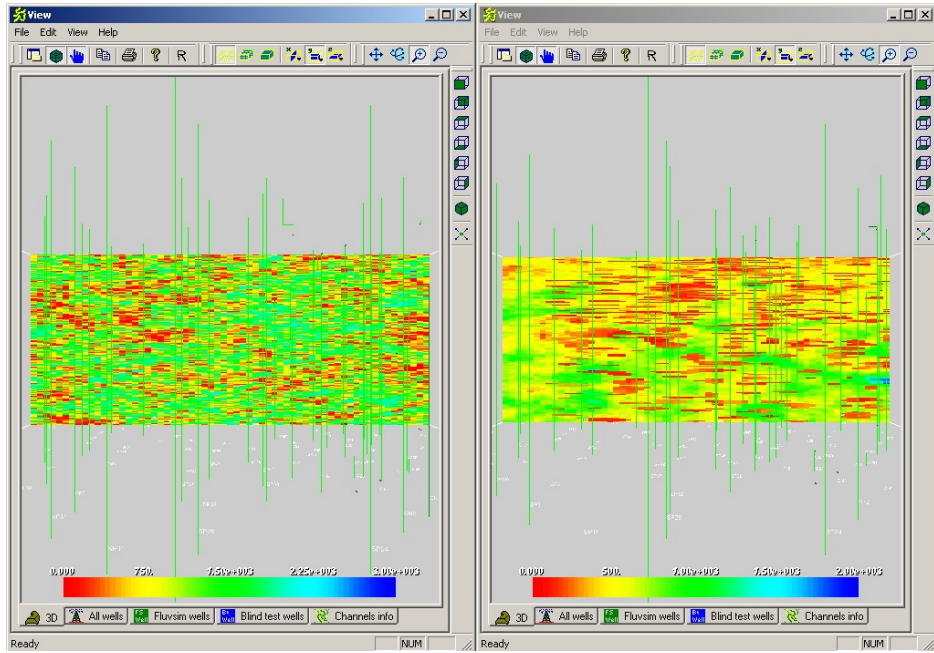


Figura 5.11 Vista frontal de Permeabilidad. Layer 140

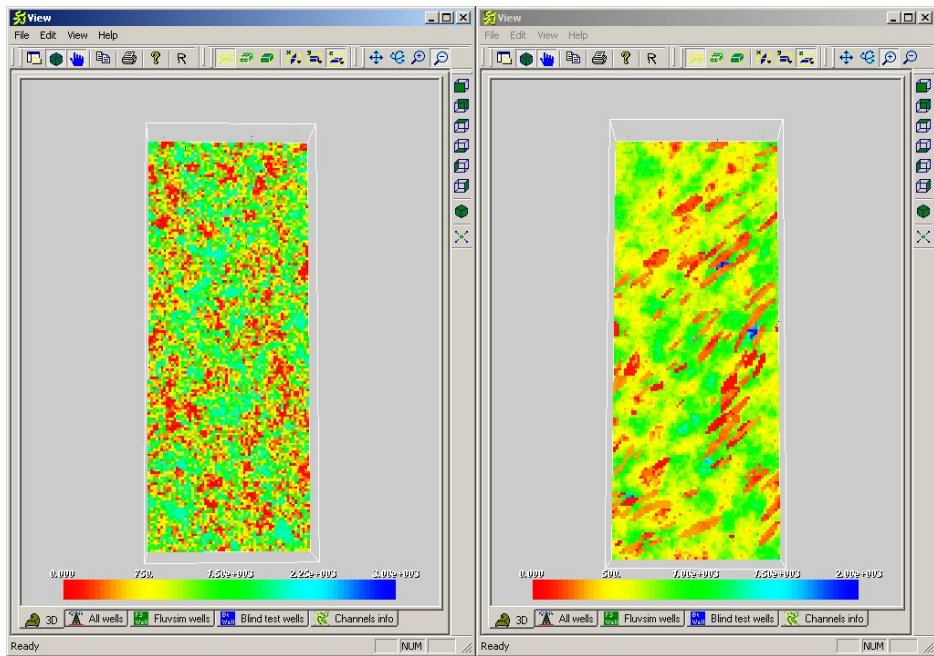


Figura 5.12 Vista en Superficie de Permeabilidad. Layer 1

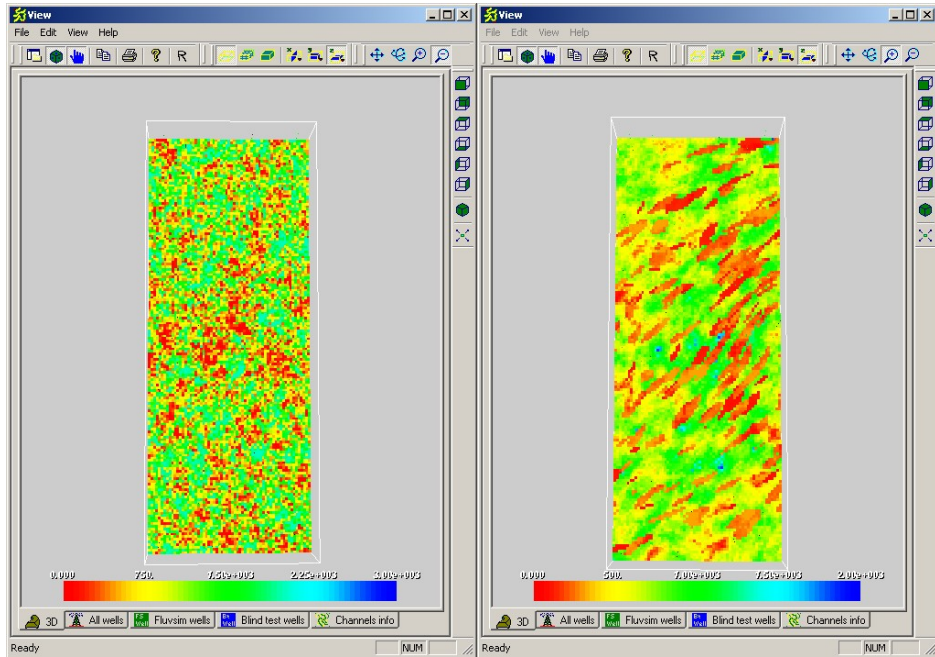


Figura 5.13 Vista en Superficie de Permeabilidad. Layer 15

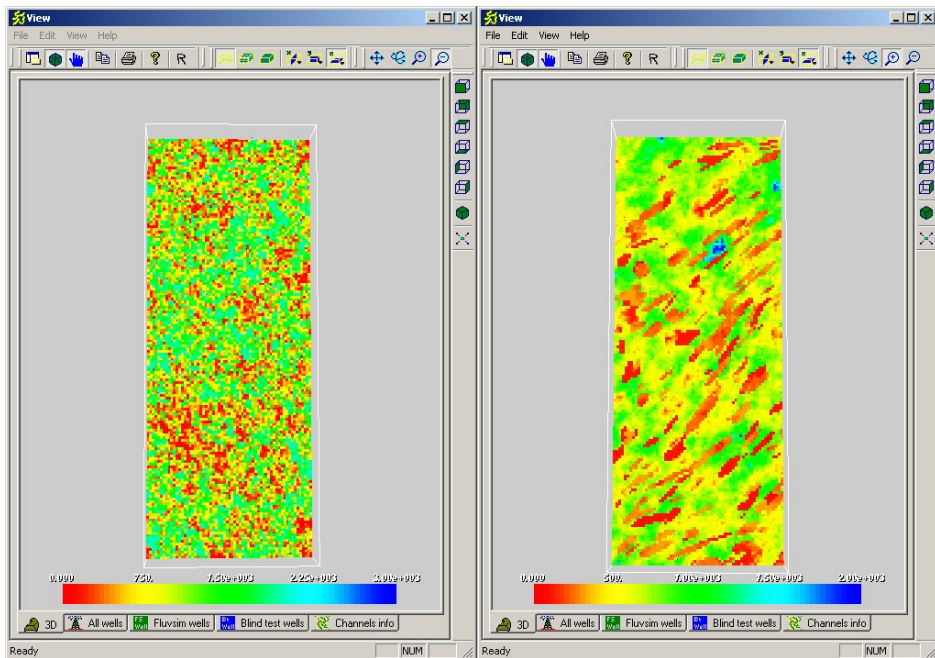


Figura 5.14 Vista en Superficie de Permeabilidad. Layer 50

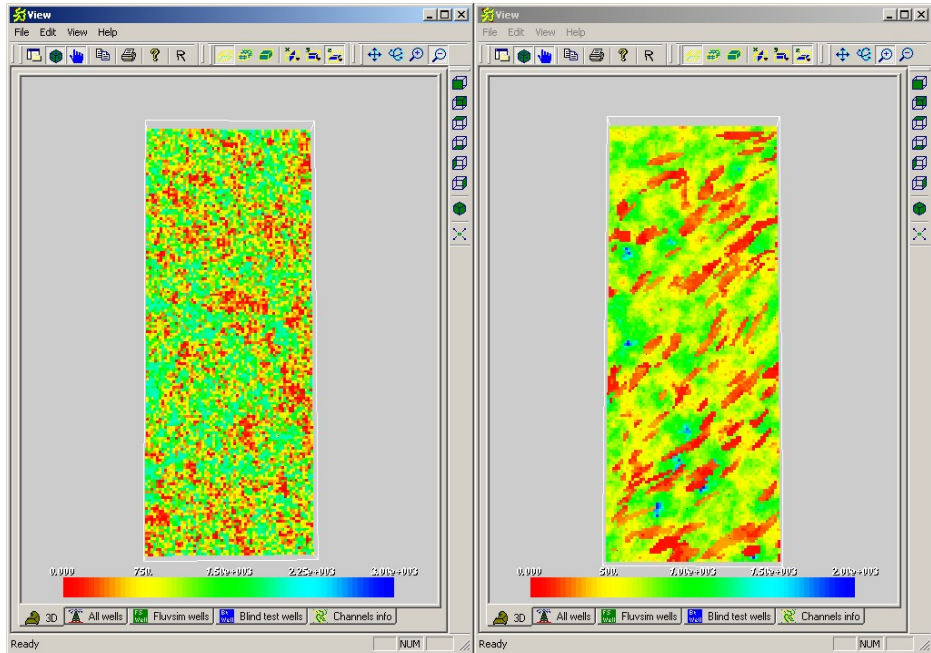


Figura 5.15 Vista en Superficie de Permeabilidad. Layer 120

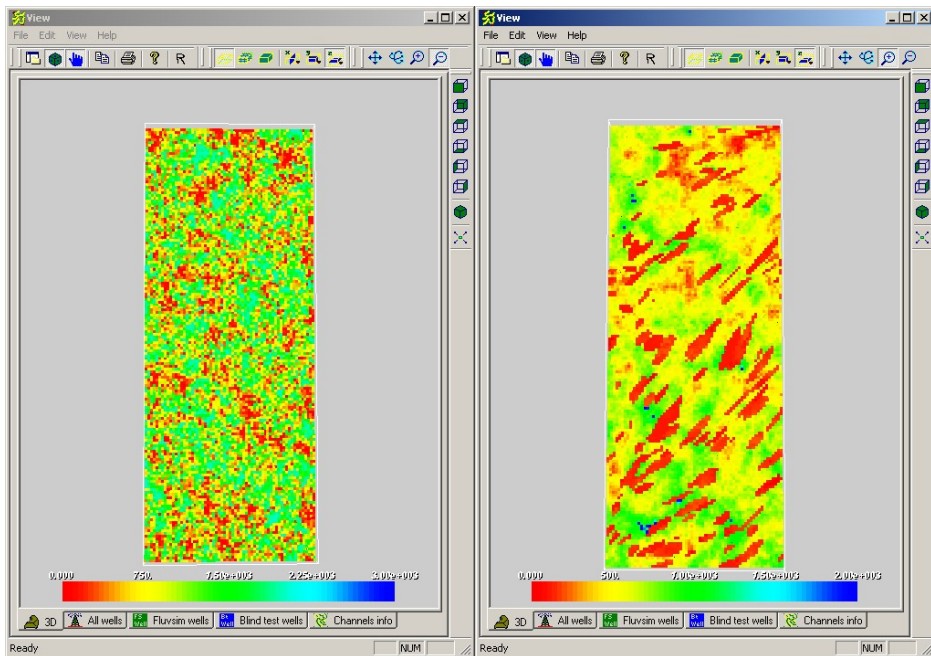


Figura 5.16 Vista en Superficie de Permeabilidad. Layer 161

Como se puede apreciar, ambos modelos coinciden en la orientación de las facies; la diferencia radica en la forma de modelar la estructura. ECOPETROL por un lado hizo el modelamiento basado en objetos, razón por la cual se observa un mejor trazado de canales, mientras que acá se trabajo con geoestadística basada en píxeles.

5.2 POROSIDAD

En la Figura 5.17 se muestra un corte en 3D de la porosidad de la unidad K2 del campo castilla utilizando la metodología planteada. En esta se aprecia que el rango de este parámetro oscila entre 0 y 0.275, predominando valores cercanos a 0.2.

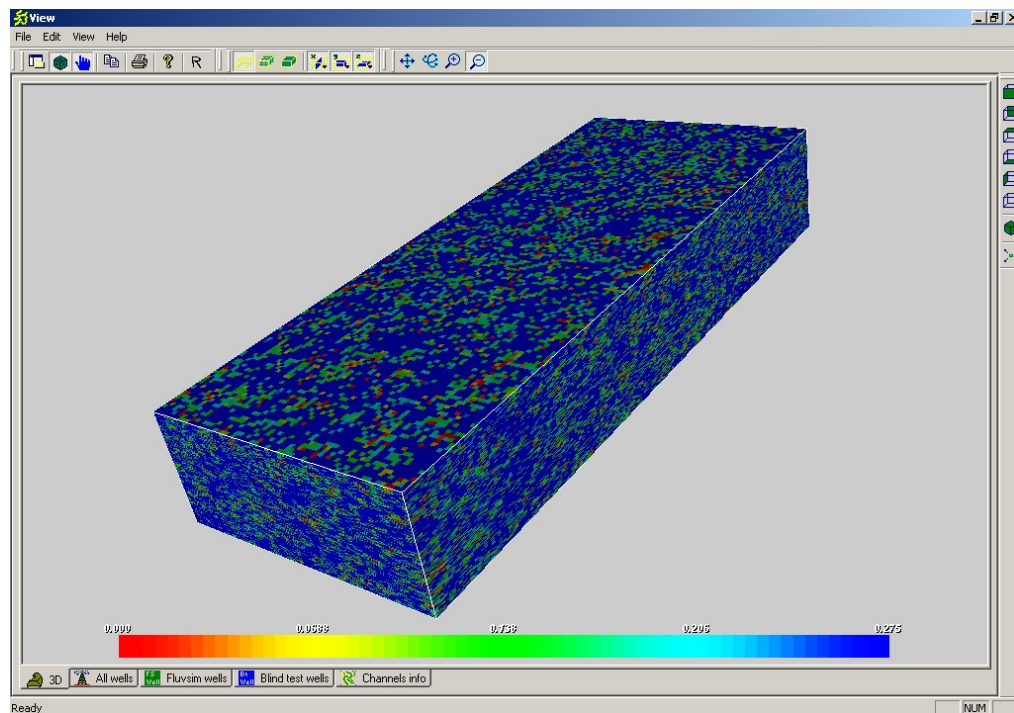


Figura 5.17 Modelo Resultante de Porosidad

De igual forma, se presentan diferentes figuras que ilustran la distribución de porosidad en las vistas enunciadas anteriormente.

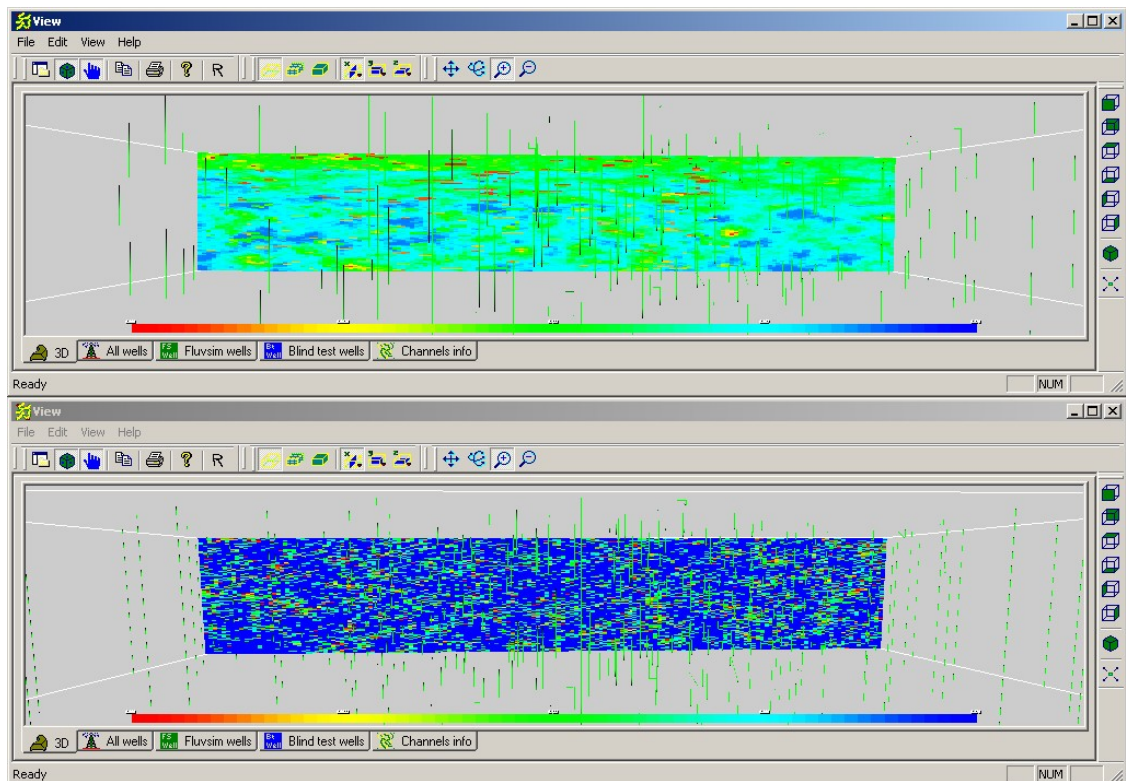


Figura 5.18 Vista lateral de Porosidad. Layer 1

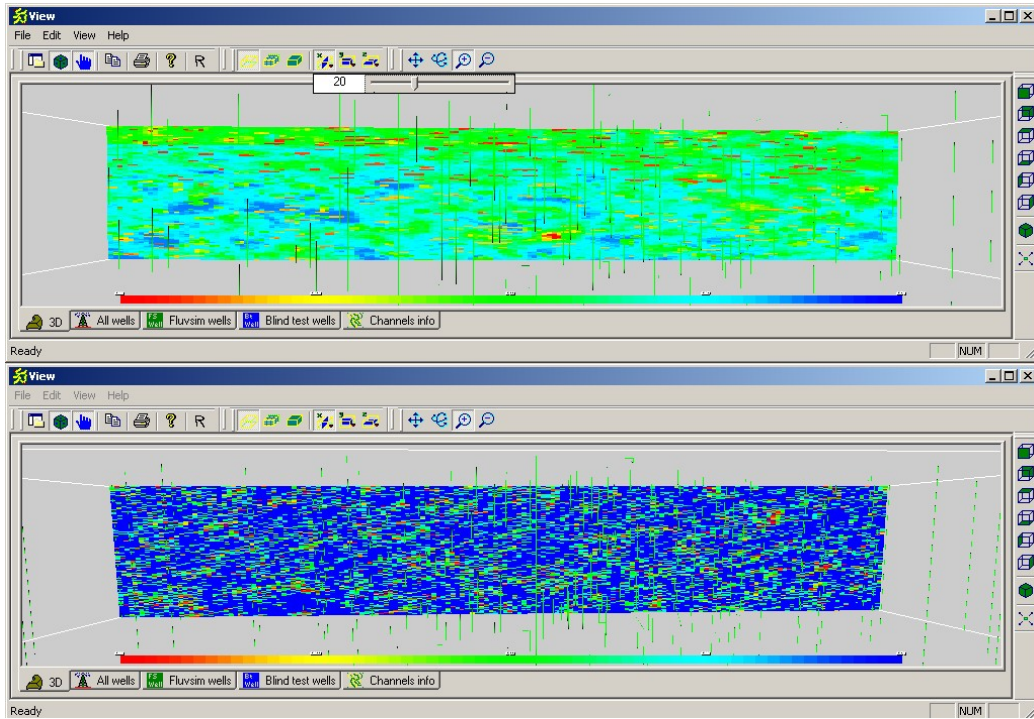


Figura 5.19 Vista lateral de Porosidad. Layer 20

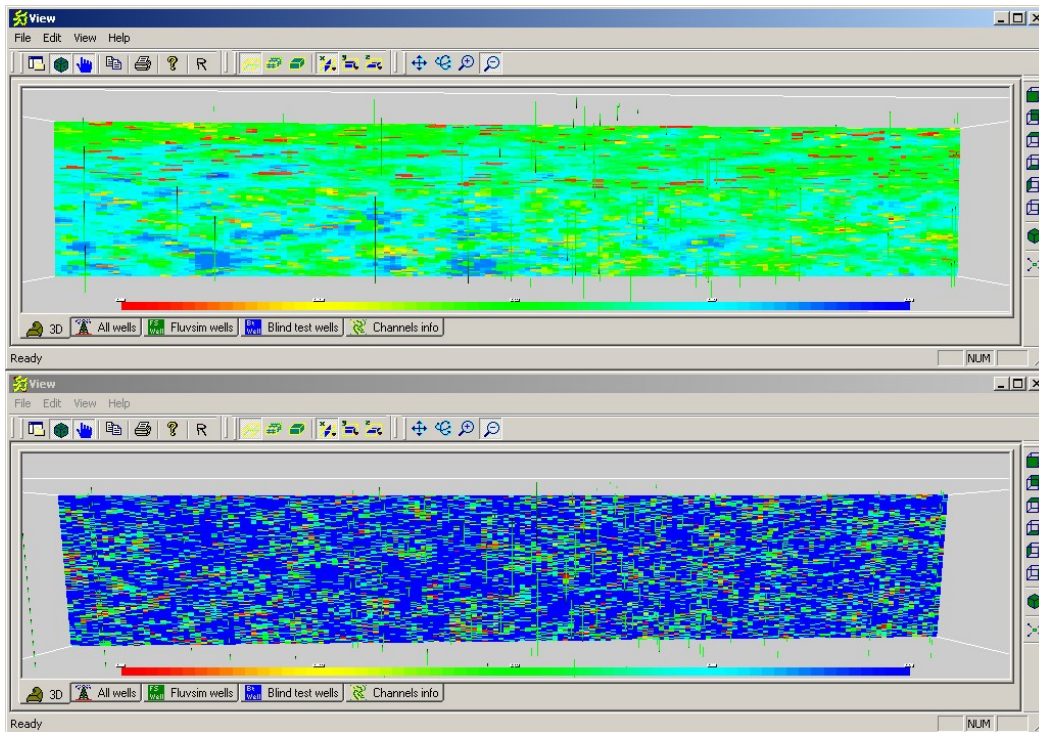


Figura 5.20 Vista lateral de Porosidad. Layer 40

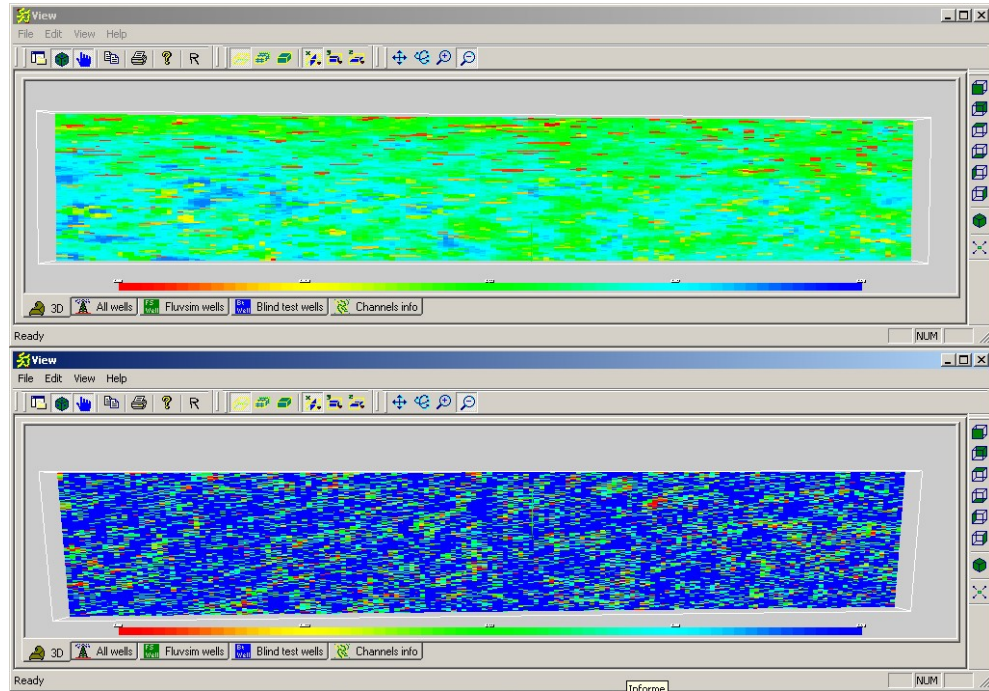


Figura 5.21 Vista lateral de Porosidad. Layer 60

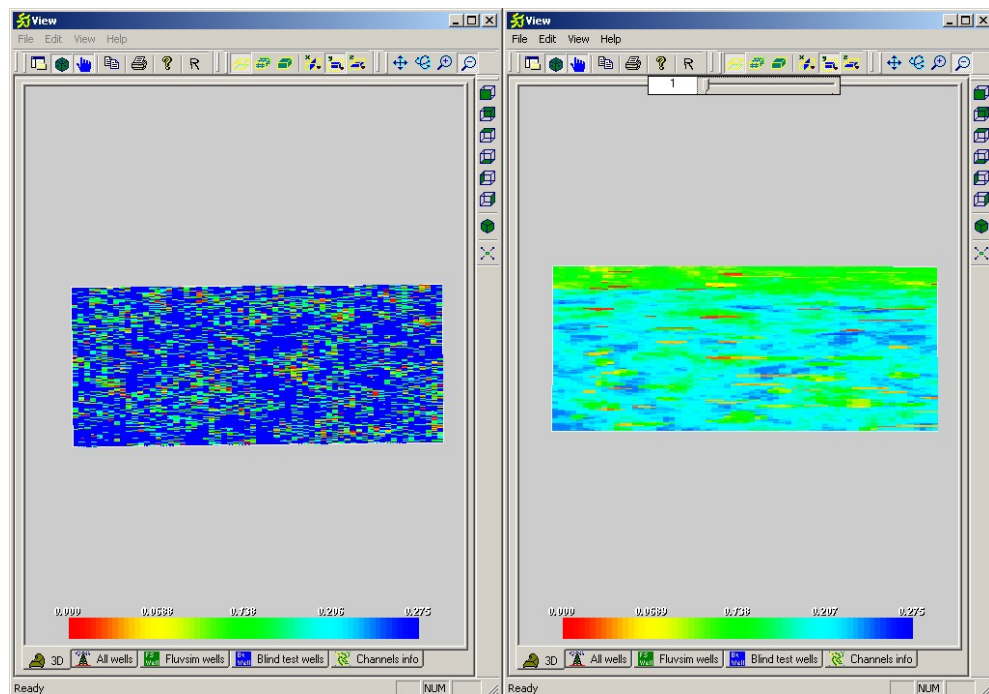


Figura 5.22 Vista frontal de Porosidad. Layer 1

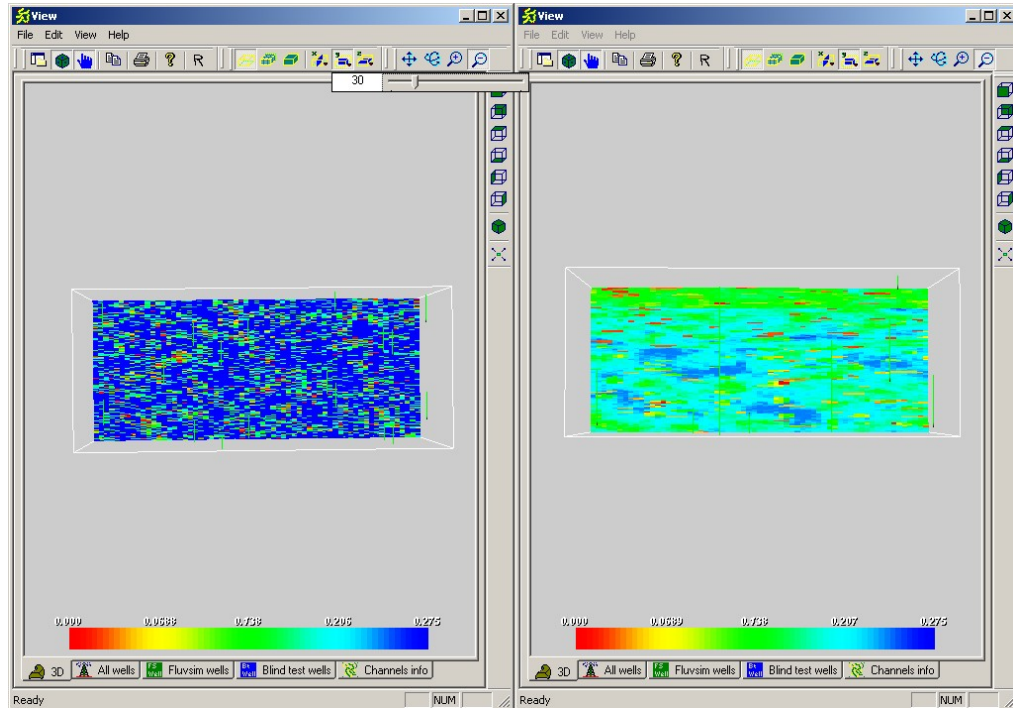


Figura 5.23 Vista frontal de Porosidad. Layer 30

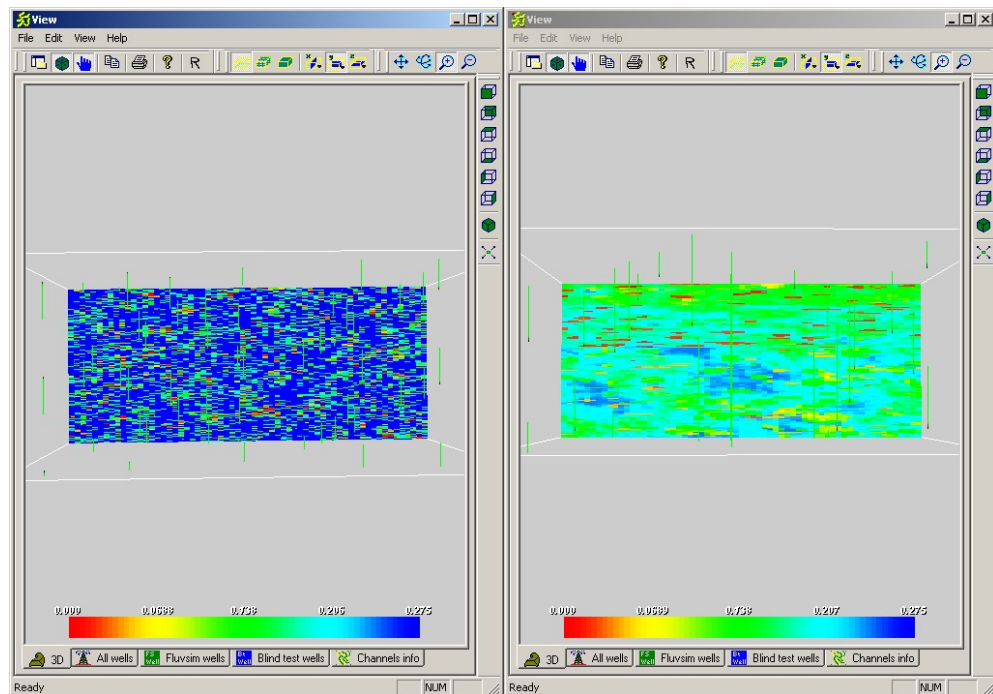


Figura 5.24 Vista frontal de Porosidad. Layer 60

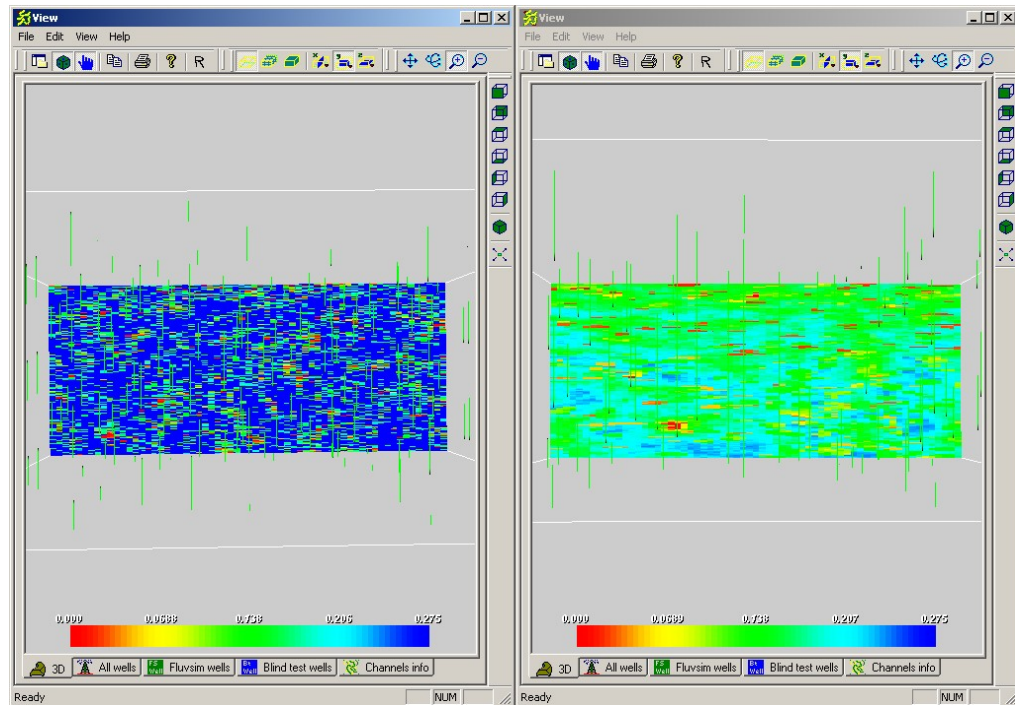


Figura 5.25 Vista frontal de Porosidad. Layer 90

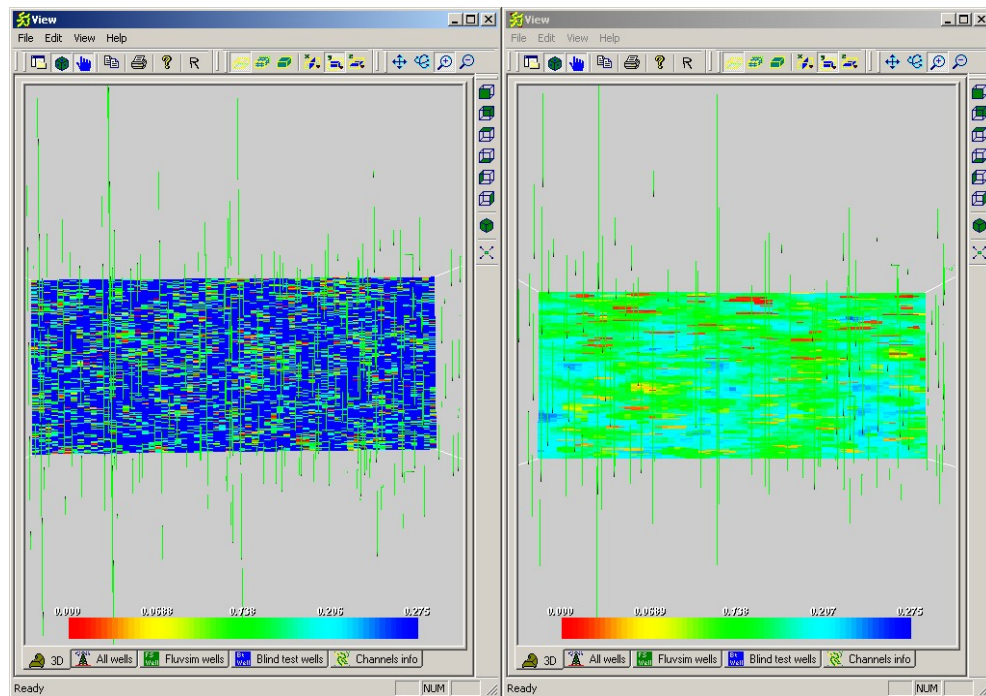


Figura 5.26 Vista frontal de Porosidad. Layer 130

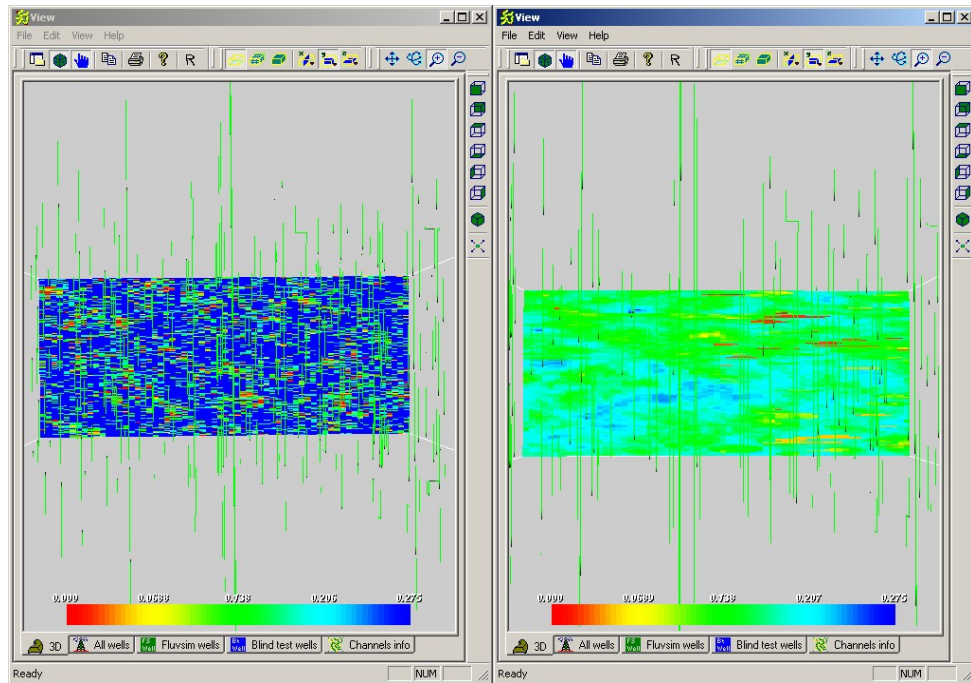


Figura 5.27 Vista frontal de Porosidad. Layer 160

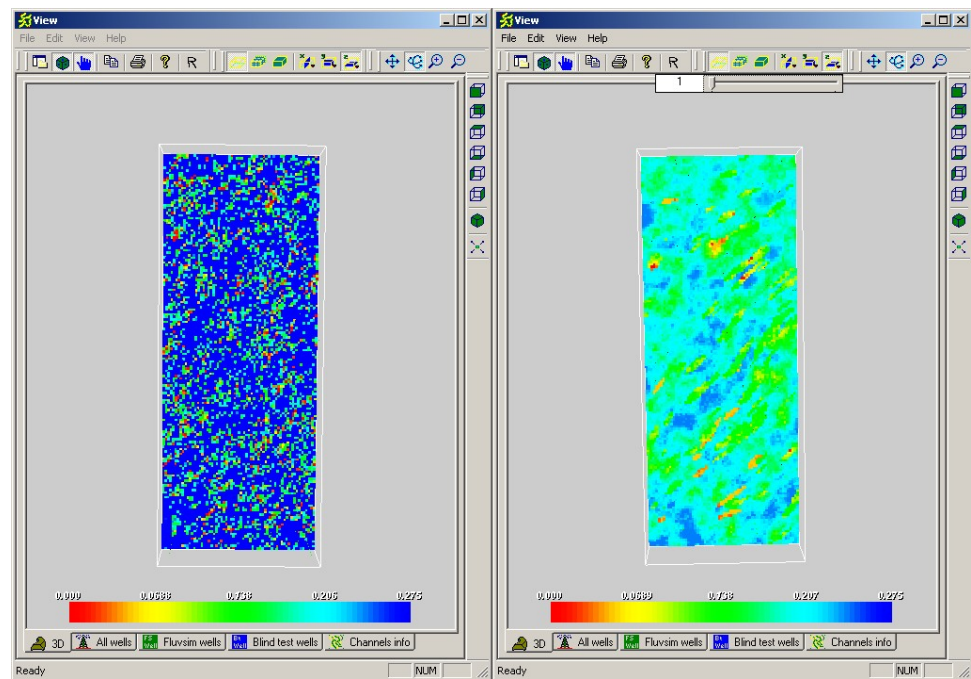


Figura 5.28 Vista en Superficie de Porosidad. Layer 1

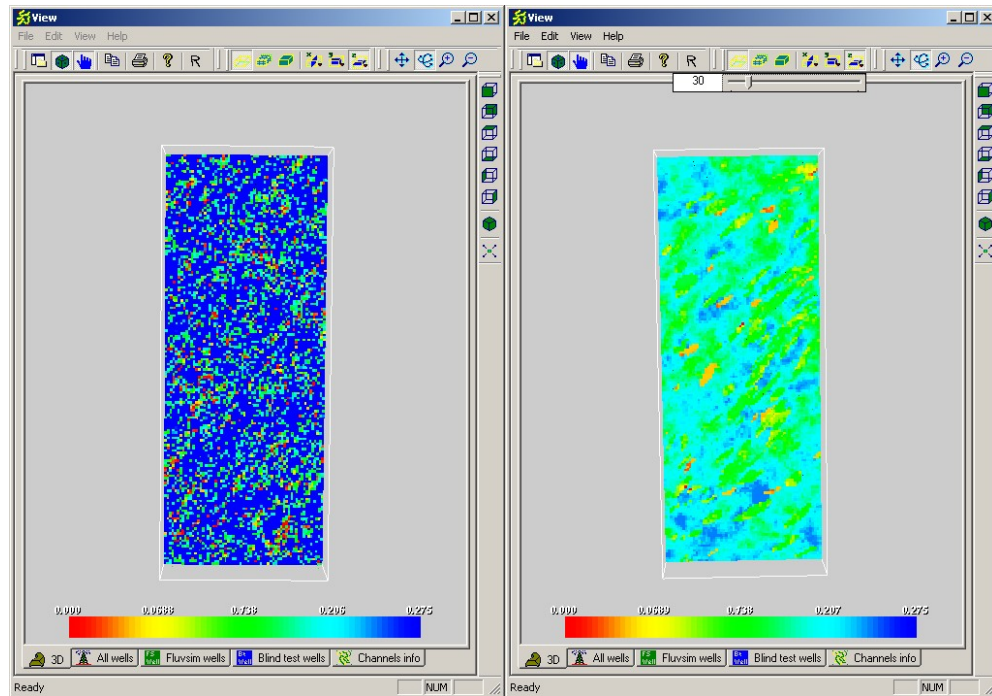


Figura 5.29 Vista en Superficie de Porosidad. Layer 30

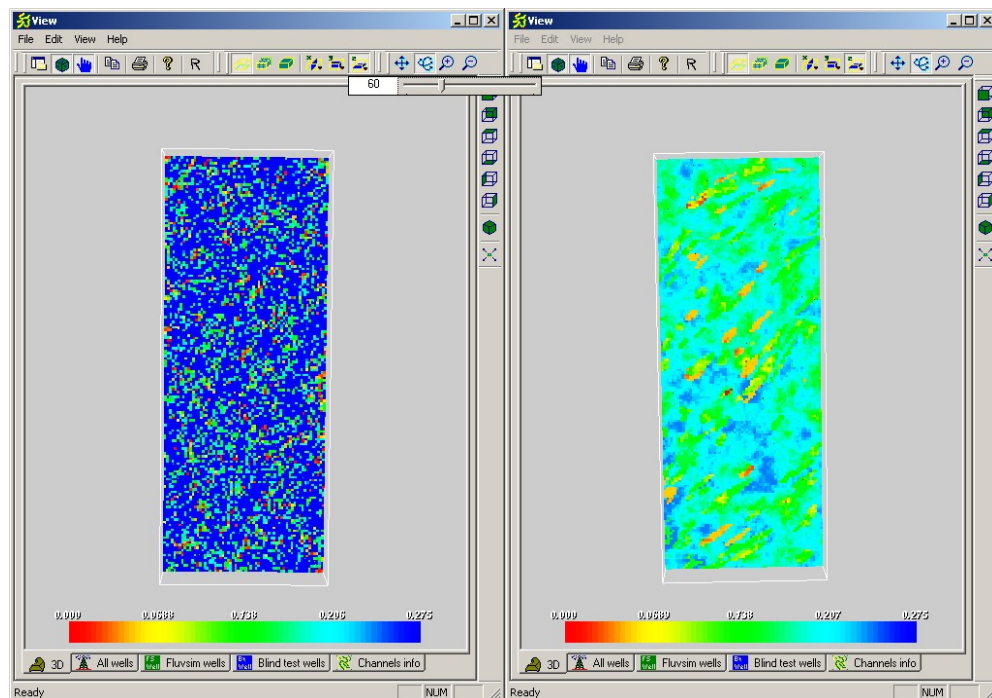


Figura 5.30 Vista en Superficie de Porosidad. Layer 60

Comparando los dos modelos (el de ECOPETROL y el nuestro), podemos decir que ambos coinciden en la orientación y la ubicación de los canales. Lo cual nos permite encontrar las zonas de mejor calidad petrofísica. Pero hay que reiterar que los modelos no coinciden totalmente debido a la forma de modelar la estructura.

Por último, observando los modelos de porosidad y permeabilidad resultantes, encontramos que el primero de estos tiene buena continuidad y por ende un mejor comportamiento, lo cual se ve reflejado en la construcción y delimitación de los canales. Este comportamiento se vio reflejado anteriormente en la construcción de los variogramas.

6. CONCLUSIONES

- Se ha logrado desarrollar una metodología para obtener un modelo de distribución en 3D de las propiedades petrofísicas condicionado a los datos estáticos y dinámicos disponibles para la formación Guadalupe masivo (unidad k2) del Campo Castilla (yacimiento fluvial), con éxito.
- En este trabajo, se muestra como la Geoestadística puede ser aplicada con éxito como herramienta para la caracterización de yacimientos de hidrocarburos condicionando el modelo geológico y estructural a los datos de producción.
- La metodología mostró ser una técnica apropiada para encontrar descripciones de yacimientos que hacen honor a los datos de producción, información de los pozos y a los rasgos geológicos; además es capaz de manejar gran cantidad de parámetros, lo cual es un problema cuando se trabaja en modelos de simulación de yacimientos.
- Esta metodología ayudará, a los ingenieros de yacimientos, a realizar trabajos similares teniendo en cuenta que sirve como guía, lo cual se resume en un ahorro de tiempo y dinero.

- La Simulación Gaussiana Secuencial (modelamiento Geoestadístico) proporcionan una herramienta robusta para ser utilizada en la solución de problemas aplicados a la industria del petróleo. Se genera de esta forma una solución muy económica en términos de costo computacional y eficiencia.
- El éxito de la simulación está dado en que permite mostrar una representación de la realidad, nunca conocida totalmente en la mayor cantidad de problemas prácticos.
- Los modelos de porosidad y permeabilidad obtenidos con los variogramas experimentales se ajustaron al modelo esférico, con el cual se obtuvo la distribución petrofísica a lo largo de la unidad de interés.

7. RECOMENDACIONES

- Aumentar las propiedades petrofísicas en consideración, incluyendo la saturación de agua, el volumen de arcilla como variable; incorporar parámetros adicionales tales como la presión capilar, propiedades PVT, Skin, etc. para obtener resultados mucho mas completos.
- Evaluar la efectividad de la metodología en yacimientos con otro tipo de ambiente de depositación.
- Posterior a este trabajo seria muy importante continuar con el cálculo de reservas para la unidad k2 del campo castilla.
- Para próximos trabajos que incluyan la implementación de métodos geoestadísticos se recomienda la utilización de otro tipo de kriging para determinar cual es el más indicado.
- De igual forma se recomienda el modelamiento basado en objetos para posteriores trabajos donde se implementen métodos geoestadísticos.

8. BIBLIOGRAFIA

- ALLARD, D; Simulation du modèle gaussien seuillé conditionné par des contraintes de connexité : Compte – rendu des journées de Geostatistique. 1993.
- ARMTRONG, M; CARIGNAN, J. Geoestadística lineal aplicada. Escuela de minas de Paris, Francia. 1997.
- BEICIP, FRANLAB; Heresim, Geology-Oriented Geostatistical Reservoir Modeling, User documentation. 1998.
- CHAUVET, P; Memorias de Geoestadística aplicada a minería. Escuela de minas de Paris, Francia. 1994.
- CHICA – OLMO, M; Análisis geoestadístico en el estudio de la explotación de recursos minerales: Universidad de Granada. Tesis Doctoral. 1987.
- CHRISTAKOS, G; PANAGOPOULOS, C: Space transformation methods in the representation of geophysical random fields. 1992.

- CUADOR, GIL; QUINTERO ARELYS. Simulación condicional de variables regionalizadas y su aplicación al comportamiento de la porosidad efectiva en un yacimiento fracturado – poroso. Boletín de la Sociedad Geológica Mexicana. 2001.
- DEUTSCH, CLAYTON; JOURNEL, ANDRE. Geostatistical Software Library and User's Guide "GSLIB". Oxford University, 1998.
- DIETRICH, C.D; A simple and efficient space domain implementation of the turning bands methods. 1995.
- FRANCO CASTRO, FERNANDO; MULFORD MARTINEZ, ALEJANDRO; Geoestadística Aplicada A La Descripción De Yacimientos. Tesis de Grado UIS. 1991.
- GALLI, A; Stochastic models for reservoir characterization: A user friendly review, in 5th American and Caribbean petroleum Engineering Conference and Exhibition: Rio de Janeiro, Brazil. 1997.
- JOURNEL, A.G; Simulation conditionnelle de gisements miniers théorie et pratique. Escuela de minas de paris, Francia. Tesis doctoral. 1974.

- JOURNEL, A.G; HUIJBREGTS, C.J. Mining Geoestatics: London, Academic press. 1978.
- GOMEZ ISIDRO, SULLY, DIR.; MARADEI GARCIA, ZULMA TERESA; ROJAS PUYANA, JUAN GUILLERMO; Distribución Espacial De Variables Hidrológicas En La Cuenca Superior Del Río Lebrija Utilizando Geoestadística. Tesis de Grado UIS. 2002.
- GONZÁLEZ, SANTIAGO; Inversión dinámica de datos para caracterizar un yacimiento fluvial usando Geoestadística y algoritmos genéticos: Campo La Cira. Tesis de Maestría UIS. 2003.
- GUIBAL, D; Esquemas intrínsecos de simulación. Escuela de minas de Paris, Francia. 1972.
- IDROBO, EDUARDO; SANTOS NICOLÁS; Optimización de costos asociados al desarrollo de yacimientos complejos mediante la reducción de la incertidumbre asociada a la determinación de las unidades de flujo. Proyecto de Investigación. Convenio ICP-UIS-COLCIENCIAS. 2004.

- ISAAKS, EDWARD; SRIVASTAVA, MOHAN. An introduction to applied Geostatistics. Oxford University, 1989.
- JIMÉNEZ, EDUARDO; Modelamiento integrado de la formación monserate del campo palogrande – cebú, Neiva. ECOPETROL – ICP, 2000.
- LANTUEJOL, CH; Interactive algorithms for conditional simulations. 1997.
- LE, LOCH; GALLI, A. Truncated plurigaussian methods: Theoretical and practical point of view. 1977.
- MATHERON, G: Aplicación de la teoría de las variables regionalizadas. 1970.
- MYERS, D.E; Multivariable geoestatistical análisis for enviromental monitoring. 1988.
- PEÑA, VICTORIA; FORERO, CÉSAR; ROJAS NELSON; Evaluación Integrada de Yacimientos, Campo Castilla. Empresa Colombiana de Petróleos, Vicepresidencia de Exploración y Producción. 1998.

- PEÑA, VICTORIA; FORERO, CÉSAR: ROJAS NELSON; Modelamiento geológico 3D usando técnicas geoestadísticas y su impacto en la simulación de yacimientos. ECOPETROL – Gerencia de Yacimientos. 2003.
- RIVOIRARD, J; Notas de clases del curso simulación, en el entrenamiento de postgrados, ciclo de formación de Especialistas en Geoestadística. Centro de Geoestadística de la Escuela superior de minas de Paris, Francia. 1998.