

TRABAJO ESPECIAL DE GRADO

EVALUACIÓN DE LA INCERTIDUMBRE EN UN MODELO ESTOCÁSTICO DE YACIMIENTO FLUVIAL MEDIANTE TÉCNICAS DE DISEÑO EXPERIMENTAL

Presentado ante la Ilustre
Universidad Central de Venezuela
Por la Br. Vallés I., María José
Para optar al Título
de Ingeniero de Petróleo

Caracas, Noviembre 2007

TRABAJO ESPECIAL DE GRADO

EVALUACIÓN DE LA INCERTIDUMBRE EN UN MODELO ESTOCÁSTICO DE YACIMIENTO FLUVIAL MEDIANTE TÉCNICAS DE DISEÑO EXPERIMENTAL

TUTORES ACADÉMICOS:

MSc. José R. Villa

MSc. Alexis Gammiero

Presentado ante la Ilustre
Universidad Central de Venezuela
Por la Br. Vallés I., María José
Para optar al Título
de Ingeniero de Petróleo

Caracas, Noviembre 2007

Caracas, Noviembre de 2007

Los abajo firmantes, miembros del Jurado designado por el Consejo de Escuela de Ingeniería de Petróleo, para evaluar el Trabajo Especial de Grado presentado por la Bachiller María José Vallés Iroba, titulado:

“Evaluación de la Incertidumbre en un Modelo Estocástico de Yacimiento Fluvial Mediante Técnicas de Diseño Experimental”

Consideran que el mismo cumple con los requisitos exigidos por el plan de estudios conducente al Título de Ingeniero de Petróleo, y sin que ello signifique que se hacen solidarios con las ideas expuestas por el autor, lo declaran APROBADO.

Prof. (nombre y apellido)
Jurado

Prof. (nombre y apellido)
Jurado

MSc. José R. Villa
Tutor Académico

MSc. Alexis Gammiero
Tutor Académico

Vallés I. María J.

EVALUACIÓN DE LA INCERTIDUMBRE EN UN MODELO ESTOCÁSTICO DE YACIMIENTO FLUVIAL MEDIANTE TÉCNICAS DE DISEÑO EXPERIMENTAL

Tutores Académicos: MSc. José R. Villa, MSc. Alexis Gammiero.

Trabajo Especial de Grado. Caracas, Universidad Central de Venezuela.

Facultad de Ingeniería Escuela de Ingeniería de Petróleo. Año 2007,

118p.

Palabras Claves: Incertidumbre, Diseño Experimental, Simulación de yacimientos, Simulador ECLIPSE 100.

Resumen. La ingeniería de yacimientos tiene como objetivo fundamental la estimación del comportamiento probable de los yacimientos sometidos a procesos de explotación, sin embargo, en todas aquellas actividades involucradas en la predicción del comportamiento de éstos, la incertidumbre es un elemento que debe ser tomado en cuenta. Para considerar este factor se ha venido incluyendo en esta área numerosas técnicas, tales como el diseño de experimentos y simulación Monte Carlo, que han demostrado ser eficientes en el estudio y comprensión de la incertidumbre, una de ellas es el Diseño de Experimentos, técnica que a través del arreglo de unidades experimentales controla la incertidumbre, lo que se traduce en máxima información, precisión y exactitud de los resultados.

Este trabajo presenta el desarrollo de una metodología que permite evaluar la incertidumbre asociada a un conjunto de parámetros, los cuales, permiten caracterizar el yacimiento. Dicha metodología involucra herramientas como la simulación numérica y estocástica, estadística y diseño experimental, todas acopladas en un flujo de trabajo que permite realizar un estudio completo, que incluye análisis de sensibilidad y de incertidumbre, que son los elementos constitutivos de un análisis.

Para llevar a cabo los objetivos propuestos se generaron modelos estocásticos de yacimiento fluvial, mediante técnicas geoestadísticas, y también propiedades inherentes a los fluidos y a la roca presentes en los yacimientos evaluados, utilizando un simulador comercial y un programa de distribución gratuita como herramientas de evaluación y procesamiento de los datos.

El desarrollo de la metodología propuesta permitió concluir que la incertidumbre, como elemento asociado a todos los parámetros que permiten caracterizar un yacimiento, es de gran importancia; y por lo tanto, se debe considerar en el estudio de yacimientos, debido a que posee un gran impacto en la elaboración de pronósticos y en la toma de decisiones.

*"Mi vocación trasciende a mi existencia,
como lo eterno a lo temporal, sin embargo,
ligada por el misterio de la libertad
ha sido formada por mí en colaboración
con la intención Divina; sufre retrocesos,
desviaciones, avances rápidos según mis respuestas
a los acontecimientos y a las manifestaciones de Dios."*

Enmanuel Múnier

Agradecimientos

A Dios Todo Poderoso, por haberme dado la fuerza, voluntad y salud para lograr esta meta.

A la Ilustre Universidad Central de Venezuela, "la Casa que vence la sombra", por haberme albergado durante este tiempo y ser el vehículo para materializar este sueño.

A mis tutores, José Reinaldo Villa y Alexis Gammiero, por haber confiado en mí y apoyarme incondicionalmente; sin ustedes este logro no hubiese sido posible.

A Teresa y Joel, por ser unos padres maravillosos, mis guías, mi apoyo y mi refugio en todo momento, los amo, este triunfo es para ustedes!

A ti Tere, que aunque estás lejos, siempre has sido mi hermana, mi amiga y mi ejemplo, por tus consejos y tu apoyo.

A ti Elisin, por ser un punto de apoyo muy importante para mi, gracias por existir madrinita!

A toda mi familia por sus oraciones, por su apoyo y comprensión, gracias por estar.

A mis amigos: Angel, Eduardo, María Andreína, Ronald, José Armando, Patty, y a mis compañeros de clases por haber compartido conmigo durante todo este tiempo.

A Carlitos, Alexander, Any y a todas aquellas personas que de alguna u otra forma aportaron su granito de arena para que este trabajo fuese posible.

Índice General

Lista de Tablas	VI
Lista de Figuras	VIII
Introducción	1
Objetivo General	3
Objetivos Específicos	3
Justificación	4
Alcances	4
Limitaciones	4
1 Fundamentos Teóricos	5
1.1 Geoestadística y Modelaje Estocástico de Yacimientos	5
1.1.1 Modelaje Estocástico de Yacimientos	6
1.2 Herramientas Estadísticas	9
1.2.1 Conceptos Básicos	9
1.2.2 Modelos de Regresión	15
1.2.3 Prueba de la Hipótesis	19
1.2.4 Análisis de Varianza (ANOVA)	22
1.3 Análisis de Incertidumbre	24
1.3.1 Clasificación de la Incertidumbre	25
1.3.2 Metodologías para el Análisis de Incertidumbre	26
1.3.3 Diseño Experimental	27
1.3.4 Simulación Monte Carlo	38
1.4 Antecedentes	40
2 Metodología	43
2.1 Definición de Parámetros de Incertidumbre	43
2.2 Clasificación de las variables de acuerdo con la incertidumbre asociada	44
2.3 Rangos de Incertidumbre	44

2.4	Generación de los modelos sintéticos de yacimiento	44
2.4.1	Modelado de Facies	45
2.4.2	Modelado de Propiedades	45
2.4.3	Escalamiento	46
2.4.4	Propiedades de los modelos de yacimiento	48
2.5	Esquema de Explotación	48
2.6	Implementación de la Metodología de Diseño Experimental	53
2.7	Flujo de trabajo propuesto	56
2.8	Descripción de las herramientas utilizadas	58
2.8.1	ECLIPSE TM	58
2.8.2	PRiSMa	61
2.8.3	Essential Regression and Experimental Design for Chemists and Engineers	63
3	Análisis de Resultados	66
3.1	Generación de los modelos de yacimiento	66
3.2	Análisis de sensibilidad	66
3.2.1	Discusión de los parámetros de mayor significancia	70
3.3	Análisis de Incertidumbre	75
3.3.1	Superficies de Respuesta	76
3.3.2	Análisis de Incertidumbre y Simulación Monte Carlo de la Su- perficie de Respuesta seleccionada	81
	Conclusiones	85
	Recomendaciones y Trabajos Futuros	87
	Bibliografía	88
	Apéndices	93
A	Correlaciones Empíricas para el cálculo de Propiedades PVT	93
A.1	Presión en el Punto de Burbujeo	93
A.2	Factor de Compresibilidad del Gas (Z)	95
A.3	Relación Gas-Petróleo	95
A.4	Factor Volumétrico de Formación del Petróleo	97
A.5	Gravedad del Gas corregida para condiciones de separador $p = 100\text{psig}$	99
A.6	Factor Volumétrico de Formación del Gas	100
A.7	Viscosidad	100

A.8	Propiedades del Agua	101
B	Compresibilidad de la roca	103
C	Presión Capilar	108
D	Permeabilidades Relativas	111
E	Error en la medición de la presión y de los parámetros PVT	116
E.1	Error en la medición de presión	117
E.2	Error en la medición de la gravedad API del petróleo	117
E.3	Error en la medición de la gravedad específica del gas	117
E.4	Error en la medición de la temperatura del yacimiento	118
E.5	Error en la medición de la relación gas-petróleo en solución	118

Lista de Tablas

1.1	Tabla ANOVA	24
2.1	Características de la malla de alta resolución	45
2.2	Proporciones globales de las facies presentes en los modelos de yacimientos	45
2.3	Parámetros utilizados para la generación de la porosidad	46
2.4	Ecuaciones utilizadas para el cálculo de la permeabilidad	46
2.5	Características de la malla del modelo escalado	48
2.6	Propiedades Generales para los modelos de yacimiento	48
2.7	Descripción de Pozos	51
2.8	Factores Potenciales de Diseño	55
3.1	Parámetros más significativos	69
3.2	Matriz de Diseño 1	71
3.3	Resultados de simulación Matriz 1	72
3.4	Resultados Matriz 1. Modelo de 14 parámetros	72
3.5	Modelo ajustado para 5 parámetros. Matriz 1	73
3.6	Coefficientes de la ecuación para la Superficie de Respuesta 1	76
3.7	Coefficientes de la ecuación para la Superficie de Respuesta 2	77
3.8	Coefficientes de la ecuación para la Superficie de Respuesta 3	78
3.9	Coefficientes de la ecuación para la Superficie de Respuesta 4	79
3.10	Coefficientes de la ecuación para la Superficie de Respuesta 5	80
3.11	Resultados Prueba t para dos muestras	82
3.12	Comparación de modelos	82
3.13	Simulaciones Monte Carlo para distintos n	84
A.1	Coefficientes a_i para p_b de acuerdo con la gravedad API	95
A.2	Coefficientes C_i para R_s de acuerdo con la gravedad API	97
A.3	Coefficientes a_i para el β_0 de acuerdo con la gravedad API	99
A.4	Coefficientes a_i para agua libre de gas	102

E.1	Símbolos adoptados	119
E.2	Unidades	121

Lista de Figuras

1.1	Ejemplo de un modelo de facies en un ambiente fluvial generado con simulación basada en objetos	8
1.2	Esquema General de la Simulación Gaussiana	10
1.3	Distribución Normal	14
1.4	Distribución t-Student	15
1.5	Distribución F-Snedecor	16
1.6	Estudio de la Incertidumbre	25
1.7	Diseño Ortogonal	31
1.8	Diseño Rotable	31
1.9	Cubo de Diseño Central Compuesto	36
1.10	Cubo de Diseño Central Compuesto Inscrito	37
1.11	Cubo de Diseño Central Compuesto con Centros en las Caras[23] . . .	38
1.12	Cubo de Diseño Box-Behnken	38
1.13	Esquema Simulación Monte Carlo	39
2.1	Escalamiento	47
2.2	Esquema de Explotación Utilizado	49
2.3	Sistema Integrado Subsuelo Superficie para una instalación Costa Afuera	52
2.4	Flujo de Trabajo Análisis de Sensibilidad	57
2.5	Flujo de Trabajo Análisis de Incertidumbre	59
2.6	Modelo de un archivo de entrada de ECLIPSE TM	60
2.7	Mallas disponibles para modelar yacimientos	61
2.8	Ventana de Diseño de Experimentos	64
2.9	Ventana de Regresión Lineal Múltiple	65
3.1	Escenario 1, una realización de facies por cinco realizaciones de propiedades	67

3.2	Sección horizontal de los modelos de porosidad y permeabilidad sin condicionar. Realización 1	68
3.3	Sección horizontal de los modelos de porosidad y permeabilidad condi- cionados a las facies. Realización 1	68
3.4	Histograma Superficie de Respuesta 1	77
3.5	Histograma Superficie de Respuesta 2	78
3.6	Histograma Superficie de Respuesta 3	79
3.7	Histograma Superficie de Respuesta 4	80
3.8	Histograma Superficie de Respuesta 5	81
3.9	Histograma. Superficie de Respuesta seleccionada	83
B.1	Diagrama del equipo de laboratorio utilizado par las pruebas de com- presibilidad	105
B.2	Compresibilidad efectiva de la roca yacimiento de acuerdo con el es- tudio de Hall	106
C.1	Presión Capilar	108
C.2	Modelo de sección transversal de un poro con saturación de agua variable	110

Introducción

La ingeniería de yacimientos tiene como objetivo fundamental la estimación del comportamiento probable de los yacimientos sometidos a procesos de explotación. Sin embargo, gracias a los avances tecnológicos en el área de la informática y en las actividades petroleras, la manera de predecir dicho comportamiento ha ido evolucionando, permitiendo aplicar nuevas técnicas e integrar las mismas, con el fin de realizar estudios de mayor complejidad. Técnicas tales como curvas de declinación, métodos de predicción analíticos, modelos físicos e incluso analogías con otros procesos, han sido utilizados hasta llegar a la simulación numérica de yacimientos.

La simulación de yacimientos tiene que ver con la construcción y operación de modelos, donde se integran todos los resultados generados en las diferentes áreas, tales como la geología, geofísica, petrofísica e ingeniería. Estos modelos son capaces de reproducir el comportamiento real del yacimiento, a través de un conjunto de funciones matemáticas, que sujetas a ciertas suposiciones, describen los procesos físicos activos en el yacimiento, tales como: transferencia de masa y movimiento de fluidos en medios porosos. El objetivo fundamental de este trabajo es el de predecir el comportamiento futuro del campo bajo uno o varios esquemas de producción. Los modelos elaborados permitirán obtener resultados aproximados de diferentes esquemas de desarrollo y producción lo cual es el principal insumo en la evaluación técnico-financiera de un proyecto, estos deberán reproducir con la mayor exactitud posible el comportamiento de producción, presión del yacimiento y validar POES / GOES, para garantizar la consistencia de los resultados.

Una de las ciencias que ayuda a que la simulación de yacimientos sea una herramienta de gran versatilidad es la Geoestadística, la cual, a través de sus principios estudia los fenómenos espaciales, con fundamentos matemáticos, las relaciones espaciales de las variables en los modelos creados, ventaja que cobra importancia cuando se dispone de poca información y cuando el conocimiento del fenómeno

sometido a estudio es limitado. Esto se logra gracias al uso de técnicas estocásticas, mediante las cuales se incluyen factores como la incertidumbre de las variables y la aleatoriedad en la distribución de las mismas.

En todas aquellas actividades involucradas en la predicción del comportamiento de un yacimiento, la incertidumbre es un elemento que debe ser tomado en cuenta. Para ello se ha venido incluyendo en el área de la ingeniería de yacimientos numerosas técnicas, que aunque no fueron desarrolladas inicialmente para ser aplicadas en esta área, han demostrado ser eficientes en el estudio y comprensión de la incertidumbre. Una de estas técnicas es el Diseño de Experimentos, la cual través del arreglo de unidades experimentales controla la incertidumbre, lo que se traduce en máxima información, precisión y exactitud de los resultados.[30]

En el presente trabajo se planteará una metodología que integre la simulación de yacimientos, con técnicas de diseño experimental para el análisis de la incertidumbre de las variables relacionadas con modelos de yacimientos generados utilizando técnicas estocásticas y conceptos geoestadísticos.

Objetivo General

Utilizar Diseño Experimental para evaluar y cuantificar la incertidumbre de diversos parámetros sobre el factor de recobro en un yacimiento fluvial generado estocásticamente.

Objetivos Específicos

1. Realizar una revisión bibliográfica.
2. Seleccionar las variables que serán sometidas a estudio.
3. Determinar las variables que poseen mayor contribución sobre la variable de respuesta, es decir, sobre el factor de recobro.
4. Ajustar una superficie de respuesta o modelo, con los parámetros de mayor contribución.
5. Realizar un análisis estadístico de los efectos de la incertidumbre en los parámetros seleccionados sobre el factor de recobro.
6. Generar una metodología que integre los pasos anteriores para determinar la incertidumbre de las variables sometidas a estudio sobre el factor de recobro.
7. Elaborar recomendaciones y conclusiones.

Justificación

Los avances tecnológicos han permitido incursionar en el modelado de yacimientos y de esta forma realizar mejores predicciones al momento de proponer un plan de explotación, ya sea para yacimientos nuevos o maduros. Esta investigación tiene como finalidad implementar una metodología fundamentada en el Diseño Experimental para evaluar y cuantificar la incertidumbre de parámetros relevantes en la vida productiva de todo yacimiento, en el factor de recobro. Entre los parámetros que se analizarán se encuentran la porosidad, permeabilidad, influencia del acuífero, facies sedimentarias, entre otras.

A través del desarrollo del presente trabajo se busca proponer una metodología que permita tomar en cuenta la incertidumbre de una manera sencilla, y que facilite el proceso de toma de decisiones en los estudios de riesgo en la simulación de yacimientos; además de establecer una base de conocimientos para apoyar futuras investigaciones en el área.

Alcance

Esta investigación se centra en implementar una metodología basada en diseño experimental para medir el impacto de la incertidumbre, asociada a las diversas variables de yacimiento, sobre las estimaciones del factor de recobro.

Limitaciones

- En el desarrollo de la metodología sólo se contempló el uso del Diseño Experimental de Tipo Plackett Burman, para generar la matriz de diseño para el análisis de sensibilidad.
- La poca disponibilidad de los paquetes de programas dificultaron la inclusión de ciertos parámetros, como por ejemplo los operacionales.

Capítulo 1

Fundamentos Teóricos

1.1 Geoestadística y Modelaje Estocástico de Yacimientos

La distribución espacial de las propiedades hidráulicas y físicas en los diferentes materiales de la naturaleza es difícil de predecir de manera determinística, para ello existen herramientas como la geoestadística. Esta última permite interpolar propiedades, considerando el efecto que produce la variabilidad en las medidas que se han tomado de poblaciones de estudio. De manera general, las ciencias de la tierra exhiben correlaciones espaciales en mayor o menor grado, de igual manera como crece y decrece la distancia entre dos puntos donde se toma información. La geoestadística es una rama de la matemática aplicada y de la estadística que permite cuantificar y modelar la variabilidad espacial de un atributo dentro de un sistema sometido a estudio, incluyendo las heterogeneidades y la dirección en que se hace la selección de la información.

La geoestadística utiliza variables estacionarias (Regionalized variables), cuyo comportamiento cambia espacialmente de manera continua. Cuando se tiene un grupo de variables correspondiente a puntos cercanos entre si es posible establecer correlaciones para modelar su comportamiento, por lo que se puede decir que éstas describen un fenómeno en función de la distribución geográfica. Sin embargo, aunque el fenómeno exhiba una variación espacial, en ocasiones no es posible tener toda la información deseada, por lo que se deben hacer aproximaciones a partir de las muestras existentes. La distribución, forma y tamaño de los puntos de muestreo son considerados factores de gran ayuda para establecer la correlación de la variable desconocida, pero modificar cualquier parámetros puede originar una correlación

diferente; para ello la geoestadística, a diferencia de las herramientas que proporciona la estadística común, proporciona el medio de cuantificar la correlación espacial de las propiedades, además de aprovechar la información referente a la variabilidad para el empleo de la misma tanto en la interpolación como en técnicas de simulación estocásticas.

De manera concisa se podría definir a la geoestadística como la rama de las ciencias estadísticas que estudia y capitaliza las relaciones espaciales para modelar posibles valores de parámetros que no han sido observados en localizaciones que no han sido probadas.[12]

1.1.1 Modelaje Estocástico de Yacimientos

La simulación estocástica es una técnica de simulación, diseñada para reproducir el histograma de los datos con la mayor aproximación posible, y ser consistentes con el modelo espacial y con datos secundarios, y para evaluar la incertidumbre de un modelo de yacimientos.

Existen muchos métodos de simulación disponibles, la selección de alguno de ellos depende del objetivo y de los datos disponibles para realizar el estudio. Los más utilizados son: la simulación secuencial, matriz de descomposición, simulación basada en objetos, entre otros. De acuerdo con los objetivos del presente estudio, sólo se hará mención de la simulación basada en objetos y de la simulación secuencial de tipo gaussiana.[15]

El modelaje estocástico de las propiedades de yacimiento se realiza en dos etapas: primero se simula la geometría de las facies y luego se simula la distribución espacial de las variables petrofísicas (porosidad y permeabilidad) para cada una de las facies. Este procedimiento permite caracterizar la continuidad y variabilidad espacial de las propiedades de la roca en el yacimiento, integrar información multidisciplinaria con diferentes resoluciones y cuantificar la incertidumbre en la descripción de los yacimientos.[40]

La simulación estocástica en geoestadística se enfoca principalmente en producir mapas realísticos de un fenómeno, más que de minimizar el error de predicción, lo cual, en ocasiones, conduce a mapas suavizados que no son verdaderamente representativos del yacimiento real.[40] Sin embargo, existen cuatro razones principales para utilizar simulación estocástica:[15]

- *Capturar las Heterogeneidades*: Un buen modelo de heterogeneidades implica un mejor entendimiento de la conectividad entre las zonas permeables y las no permeables. Dicho entendimiento se traduce en un mejor conocimiento de la eficiencia de barrido y la producción de fluidos. Por otra parte, aunque, un modelo de heterogeneidades detallado no es requerido para los cálculos volumétricos, si tienen gran impacto en las características de flujo del modelo.
- *Simular las facies, las propiedades de las rocas o ambos*: La arquitectura del yacimiento es, por lo general, el primer paso, y consiste en capturar elementos como fallas, tope y base de la estructura; el segundo paso es identificar las unidades geológicas, teniendo como punto de referencia los principios estratigráficos y definiendo la geometría interna de las capas. El tercer paso, implica el modelado espacial de la distribución de las facies; el paso final es construir el modelo estático para poblarlo con las litofacies, propiedades de las rocas y los fluidos. La principal diferencia entre modelar facies y propiedades de las rocas es que la primera representa una variable discreta y la segunda variables continuas.
- *Consistencia con los datos secundarios*: utilizar métodos estocásticos permite incorporar otro tipo de información que los métodos convencionales no pueden manejar.
- *Evaluar la Incertidumbre*: El estudio de yacimientos siempre tiene incertidumbre asociada, el uso de este tipo de metodología permite producir modelos diferentes, pero que sean consistentes con los datos de entrada.

Simulación basada en objetos

Este tipo de simulación crea modelos de yacimientos basada en objetos, es decir, grupos de cuerpos, que tienen significado genético. Es comúnmente utilizada para reproducir heterogeneidades a gran escala en yacimientos de petróleo, frecuentemente estos cuerpos no están distribuidos de manera uniforme en el dominio de estudio.

El objetivo de esta técnica es reproducir la geometría de cuerpos de manera que estos coincidan con las interpretaciones geológicas; cada cuerpo se considera como un objeto con una geometría dada y la mayoría se disponen en orden aleatorio para llenar las unidades geológicas.

Para hacer uso de ésta es necesario seleccionar una forma básica para cada facie depositacional que describa su geometría, especificar las proporciones de las formas en el modelo final y seleccionar una distribución para los parámetros que describen las formas. Algunos algoritmos tienen reglas que describen como los geocuerpos están dipuestos de manera relativa unos con otros. Una vez realizado estos pasos, se llena el fondo del modelo con alguna de las litofacies (como la lutita), se selecciona aleatoriamente un punto de partida en el modelo y una forma de litofacies para luego dibujarla con un tamaño, anisotropía y orientación adecuada. Posteriormente, se verifica si la forma coincide con cualquiera de los datos de condicionamiento o con otras formas previamente simuladas. Sí esto ocurre se mantiene la forma, sino se rechaza y se regresa al paso anterior. Para finalizar, se verifica si las proporciones globales del modelo se cumplen y sí esto no ocurre se repite el procedimiento a partir de la selección aleatoria del punto de partida en el modelo.

La simulación basada en objetos es una técnica de actual interés en la industria petrolera, ésto es debido a que la propuesta para modelar es particularmente satisfactoria para los geólogos, motivado a que los objetos creados están basados en las estadísticas de las relaciones de formas y facies que han sido medidas y porque las imágenes provenientes de los modelos resultantes de facies depositacionales lucen realistas, ver Figura 1.1¹. [40]

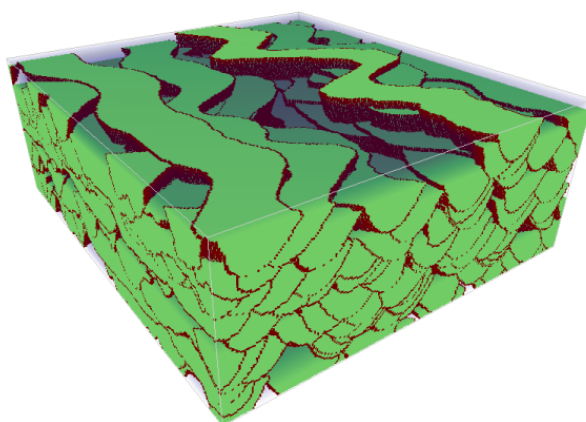


Figura 1.1: Ejemplo de un modelo de facies en un ambiente fluvial generado con simulación basada en objetos

¹Fuente <http://villaj.googlepages.com/prisma>

Simulación Secuencial Gaussiana

La Simulación Secuencial Gaussiana (Sequential Gaussian Simulation SGS) es un algoritmo estadístico ampliamente utilizado en geoestadística para reproducir la distribución espacial y la incertidumbre de las variables de diferentes fuentes, cuando los datos presentan una distribución normal.

Para realizar este tipo de simulación es necesario, como primer paso, determinar si la Función de Distribución de Probabilidad multivariada de la función aleatoria a simular es Gaussiana, y en caso de que no sea, transformarla para que resulte una distribución normal, esto es porque la mayoría de los fenómenos de las ciencias de la tierra no presentan histogramas simétricos, mucho menos gaussianos.

El próximo paso es el proceso de simulación, donde primero, se estiman los parámetros de la función de densidad de probabilidad para cada localización del grupo de datos y luego, aleatoriamente, se genera un valor a partir de su distribución. En la Figura 1.2 se muestra un esquema general para este tipo de simulación.²

El resultado de una simulación es una versión ruidosa de un proceso de estimación, el cual reproduce la estadística de los datos conocidos, dando un aspecto realístico del modelo, pero proporcionando un bajo comportamiento de predicción, sin embargo, si se diseñan múltiples secuencias de simulación, es posible dibujar mapas probabilísticos más confiables.

1.2 Herramientas Estadísticas

1.2.1 Conceptos Básicos

Medidas de Dispersión

La dispersión de un conjunto de observaciones es medida por la desviación de las mismas con respecto a la media del conjunto. Existen diferentes maneras de medir esta dispersión, entre ellas; el rango, la desviación media, el rango semi-cuartil, la varianza, la desviación estándar y otras. Sin embargo, en el área de la ingeniería las de mayor aplicación son: el rango, la varianza y la desviación estándar. Estos términos sirven para cuantificar la variabilidad de una muestra midiendo su dis-

²Fuente <http://132.248.182.159/cursos/gest/Presentaciones/CG8.pdf>

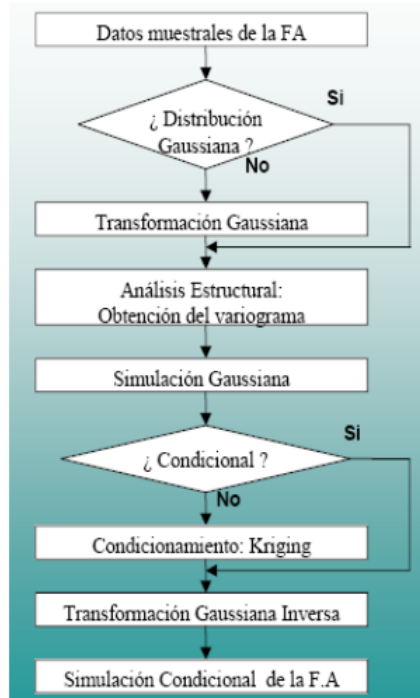


Figura 1.2: Esquema General de la Simulación Gaussiana

persión alrededor de la media.

- Varianza: Es un estimador de la dispersión de una variable aleatoria, y se obtiene como resultado de la división de la sumatoria de las distancias existentes entre cada dato y su media aritmética elevadas al cuadrado, entre el número total de datos.

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (1.1)$$

Algunas propiedades de la varianza son, sea X una variable aleatoria:

1. $V(X) \geq 0$ propiedad que permite que la definición de desviación típica sea consistente.
2. $V(aX + b) = a^2V(X)$ siendo a y b constantes cualesquiera.
3. $V(X) = E[X^2] - E[X]^2$, donde E se define como la esperanza matemática.
4. $V(X + Y) = V(X) + V(Y)$ Si X e Y son variables aleatorias independientes, entonces

5. $P(|X - E[X]| \leq k\sigma) \geq 1 - \frac{1}{k^2}$ Desigualdad de Chebyshev, para cualquier constante k mayor que 1

- **Desviación estándar:** La desviación estándar (o desviación típica) es una medida de dispersión para variables de razón (ratio o cociente) y de intervalo. Es una medida (cuadrática) que informa de la media de distancias que tienen los datos respecto de su media aritmética, expresada en las mismas unidades que la variable.

$$\sqrt{s^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1.2)$$

La ventaja de la desviación estándar sobre la varianza es que se expresa, como la media, en las mismas unidades que los datos. A veces se emplea el coeficiente de variación, que es la razón entre la desviación estándar y la media.

Medidas de Tendencia Central

Las medidas más importantes en el desarrollo de un análisis estadístico son aquellas que indican una posición central, ya que son valores típicos y representativos de un conjunto de datos. Existen varios tipos de medidas de tendencia central, siendo las más comunes: la media, la mediana y la moda. Tienen como propósito general:

- Mostrar en que lugar se ubica el valor promedio de la población.
- Comparar o interpretar datos con el valor promedio o típico.
- Comparar diferentes valores correspondientes a un mismo evento.
- Comparar los resultados medios obtenidos por dos o más grupos de datos.

Las medidas de tendencia central más comunes son:

- **Media:** se define como el promedio de un conjunto de datos, existen diferentes tipos, entre ellas destacan la media geométrica, la media armónica, la media cuadrática, la media ponderada, la media aritmética, la media aritmética geométrica y la media generalizada; siendo la media aritmética la más común, se define como:

$$m = \sum_{i=1}^n (X_i) / n \quad (1.3)$$

- **Mediana:** Considerando los datos de una muestra ordenada en orden creciente $X_1, X_2, X_3, \dots, X_n$ (de menor a mayor), se define como mediana al valor de

la variable que deja el mismo número de datos antes y después que él. De acuerdo con esta definición el conjunto de datos menores o iguales que la mediana representarán el 50% de los datos, y los que sean mayores que la mediana representarán el otro 50% del total de datos de la muestra.

- **Moda:** es el valor que cuenta con una mayor frecuencia en una distribución de datos.

Funciones de Distribución de Probabilidad

Las distribuciones de probabilidad son modelos que describen la variación de los resultados o valores probables de una variable aleatoria o distribuida. Debido a que estas distribuciones tratan sobre expectativas de que algo suceda, resultan ser modelos muy útiles para hacer inferencias y para tomar decisiones en condiciones de incertidumbre[35].

Se pueden clasificar, de acuerdo con el tratamiento que se le da a la serie de datos en: continuas, discretas, limitadas, parcialmente limitadas, paramétricas, no paramétricas y subjetivas. En concordancia con los fines del presente estudio se hará referencia detallada de las funciones de distribución paramétricas y no paramétricas.

- **Distribuciones Paramétricas:** son modelos gráficos que se ajustan a la descripción matemática de un proceso aleatorio que cumple con determinados supuestos teóricos. Los parámetros que definen la distribución en general no guardan relación intuitiva con la forma de la distribución. Son de aplicación cuando: la teoría sobre la que se fundamenta una determinada distribución es aplicable al problema, se acepta que esa distribución da un buen ajuste de la variable aleatoria aunque no haya una teoría para explicarlo, la distribución se ajusta aproximadamente a la opinión del experto y no se requiere mucha precisión. Como ejemplo de este tipo de distribución se encuentran las distribuciones Normal, Lognormal, Exponencial, Beta.
- **Distribuciones no Paramétricas:** son modelos gráficos que representan un grupo en particular de observaciones de la variable aleatoria y que relaciona los diversos valores de la variable con su probabilidad de ocurrencia. Los parámetros que se usan para definir las describen la forma de la distribución. No se apoyan en una teoría que describa el proceso de generación de valores aleatorios. Como ejemplos para este tipo de distribución están las distribuciones: Triangular, Histograma, General, Uniforme, Acumulada.[10]

Una distribución de probabilidad se caracteriza, en general, por medio de tres criterios:

1. El valor central o medida de posición (la media, la mediana o la moda)
2. Una cantidad que expresa el grado de dispersión (e.g., la desviación estándar)
3. La forma de la curva, es decir la forma general de la distribución probabilística

Dentro de las funciones paramétricas, las más habituales se basan en la distribución de probabilidad normal, y al estimar los parámetros del modelo se supone que los datos constituyen una muestra aleatoria de esa distribución, por lo que la elección del estimador y el cálculo de la precisión de la estimación, son elementos básicos para construir intervalos de confianza y contrastar hipótesis, que dependen del modelo probabilístico supuesto. Cuando un procedimiento estadístico es poco sensible a las alteraciones en el modelo probabilístico supuesto, es decir, cuando los resultados obtenidos son aproximados mientras el modelo varía se dice que es un procedimiento robusto.

Distribución Normal : La distribución normal, es posiblemente la distribución más importante en estadística. Se caracteriza por su forma acampanada, es simétrica (Ver Figura 1.3). Hay tres condiciones observables en las variables que siguen la distribución normal: a) El valor probable de la variable es el valor central, promedio, average o media de la distribución, b) La variable podría, indistintamente, tomar un valor por debajo o por encima de la media (Simetría alrededor de la media, desviación estándar (σ)), c) Es más probable que la variable tome un valor cerca de la media que lejos de ella. Esta distribución viene dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-(x-\mu)^2}{2\sigma^2} \quad (1.4)$$

donde: x representa la variable, μ la media y σ la desviación estándar.

Distribución Lognormal : presenta tres características notables: i) la variable puede crecer sin límite, pero no tomar valores negativos, ii) la variable muestra una alta tendencia hacia los valores mínimos, iii) la variable puede ser muy dispersa y sus valores tener diversos ordenes de magnitud, iv) el logaritmo natural de los valores dará como resultado como representación gráfica una curva normal. Esta función está definida por:

$$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp \frac{-(\log(x)-\mu)^2}{2\sigma^2} \quad (1.5)$$

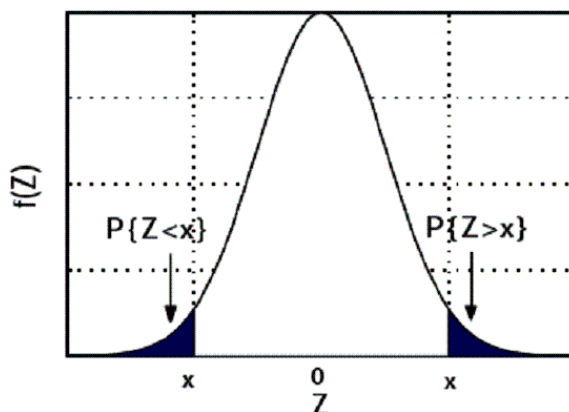


Figura 1.3: Distribución Normal

Distribución t-Student : es una herramienta de gran utilidad cuando la población es reducida (muestras de 30 o menos elementos) y cuando la desviación estándar de la población no se conoce y se tiene que determinar a partir de la data disponible. Se considera un caso especial de la distribución Hiperbólica. Se utiliza para realizar la prueba estadística de significancia entre las medias de dos muestras y para determinar el intervalo de confianza entre las medias de dos poblaciones. Aunque a simple vista es similar a la distribución normal se diferencia de ésta en que es menor en la media y más alta en los extremos que la distribución normal, y tiene proporcionalmente mayor parte de su área en los extremos que la distribución normal, Figura 1.4.

Se utiliza para contrastar hipótesis sobre medias en poblaciones con distribución normal. También proporciona resultados aproximados para los contrastes de medias en muestras suficientemente grandes cuando estas poblaciones no se distribuyen normalmente (aunque en este último caso es preferible realizar una prueba no paramétrica)³.

Matemáticamente, se define como el cociente de una variable aleatoria normal estándar y la raíz cuadrada de una variable aleatoria Chi-cuadrada dividida por sus grados de libertad, cuya expresión es la siguiente:

³Fuente <http://virtual.upct.edu.co/ova/estadistica/docs/libros/tstudent.pdf>

$$t = \frac{Z}{\sqrt{\chi/\nu}} \quad (1.6)$$

Donde:

t : variable que tiene distribución t-Student.

Z : variable con distribución Normal.

χ : variable con distribución Chi-cuadrada.

ν : grados de libertad de la distribución Chi-cuadrada.

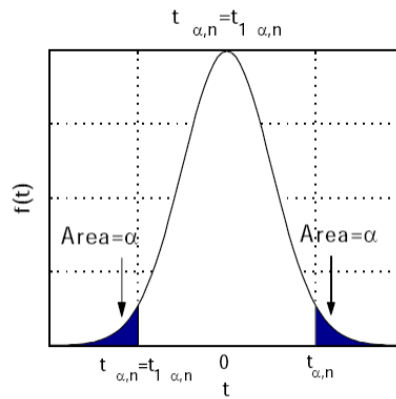


Figura 1.4: Distribución t-Student

Distribución F-Snedecor : Sean χ_1^2 y χ_2^2 dos variables aleatorias independientes Chi-cuadrada, cada una dividida por sus grados de libertad, se dice que la variable aleatoria:

$$t = \frac{\chi_1^2/\nu_1}{\chi_2^2/\nu_2} \quad (1.7)$$

tiene una distribución F. Este tipo de distribución se caracteriza completamente por los grados de libertad ν_1 y ν_2 , por lo que se puede decir que tiene asimetría positiva para cualquier valor de grados de libertad, Figura 1.5. Permite hacer inferencias acerca de la diferencia entre dos medias poblacionales.

1.2.2 Modelos de Regresión

Son modelos que se utilizan para, representar el comportamiento de una variable empírica o respuesta, a partir de una o más variables de entrada, teniendo en cuenta

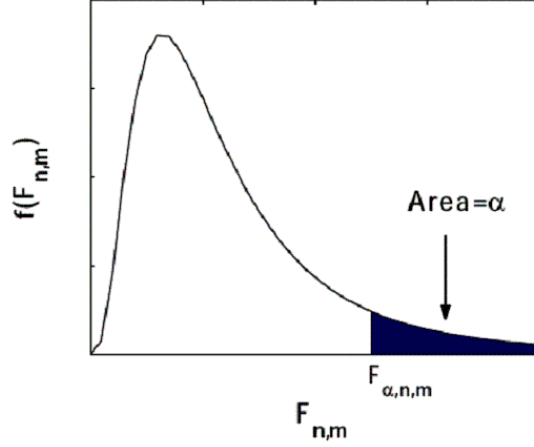


Figura 1.5: Distribución F-Snedecor

la relación funcional que existe entre la variable de respuesta y la(s) variable(s) de entrada. La variable estimada se denomina variable dependiente, y las variables de entrada independientes.[46]

Al hacer mención a un proceso de regresión lineal, se indica que la relación funcional entre las variables de entrada y salida puede ser expresada a través de una ecuación lineal, como la que se muestra a continuación:

$$y_{est} = b_0 + b_1x_1 + b_2x_2 + \dots + b_ix_i + error \quad (1.8)$$

Sin embargo, es importante mencionar que la regresión lineal incluye ecuaciones para modelos que contengan términos de orden superior (cuadráticos, cúbicos, etc) derivados de la relación funcional entre la respuesta (y_{est}) y las variables de entrada (x_i), ejemplo:

$$y_{est} = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_1^2 + b_5x_2^2 + error \quad (1.9)$$

Mínimos Cuadrados

Mínimos cuadrados es una técnica de optimización matemática que, dada una serie de mediciones, intenta encontrar una función que se aproxime a los datos. Se utiliza para estimar los coeficientes b_j de la ecuación 1.8. En este método el término correspondiente al error se define como la diferencia entre la respuesta observada

y la estimada y_{est} para un conjunto x_j dado en cada punto. El error total debe definirse como la suma de todos los puntos o casos.

De manera general, un grupo de datos consiste en n puntos considerados como una muestra de una población entera de datos. Un punto o experimento se define a través de la selección de una variable de entrada o variable independiente $x_1, x_2, \dots, x_i$ y la variable de salida o dependiente para la corrida del experimento dado. Tomando la ecuación 1.8 y considerando una población de n datos se obtiene un conjunto de ecuaciones que conforman un sistema de n ecuaciones lineales, una ecuación por experimento o corrida:

Experimento 1:

$$y_1 = b_{01} + b_1x_{11} + b_2x_{21} + \dots + b_ix_{i1} + error_1 \quad (1.10)$$

Experimento 2:

$$y_2 = b_{02} + b_1x_{12} + b_2x_{22} + \dots + b_ix_{i2} + error_2 \quad (1.11)$$

Experimento n:

$$y_n = b_{0n} + b_1x_{1n} + b_2x_{2n} + \dots + b_ix_{in} + error_n \quad (1.12)$$

Reordenando las ecuaciones para colocar el error del lado izquierdo y facilitar la suma de los mismos se obtiene:

Experimento 1:

$$error_1 = y_1 - b_{01} - b_1x_{11} - b_2x_{21} - \dots - b_ix_{i1} \quad (1.13)$$

Experimento 2:

$$error_2 = y_2 - b_{02} - b_1x_{12} - b_2x_{22} - \dots - b_ix_{i2} \quad (1.14)$$

Experimento n:

$$error_n = y_n - b_{0n} - b_1x_{1n} - b_2x_{2n} - \dots - b_ix_{in} \quad (1.15)$$

Utilizando k como el índice para los experimentos y j como el índice para cada variable independiente i , la suma de los errores elevados al cuadrado, mejor conocida como SSE (Sum of Squared Errors) se obtiene de la suma de los cuadrados de los

términos de la derecha, reordenando las ecuaciones anteriores:

$$SSE = \sum_{k=1}^n (y_k - b_0 - b_1x_1 - b_2x_2 - \dots - b_ix_{ik})^2 \quad (1.16)$$

$$SSE = \sum_{k=1}^n (y_k - b_o - \sum_{j=1}^i b_jx_{jk})^2 \quad (1.17)$$

Para el uso de esta herramienta se establece como punto de partida la siguiente ecuación:

$$y = Xb + error \quad (1.18)$$

Donde, término a término, los elementos constitutivos de la ecuación son:

$$\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1i} \\ 1 & x_{21} & x_{22} & \dots & x_{2i} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{ni} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix} + \begin{pmatrix} error_1 \\ error_2 \\ \dots \\ error_n \end{pmatrix} \quad (1.19)$$

Las n variables dependientes corresponden al vector y de dimensión $nx1$, los regresores o parámetros independientes más la constante o punto de intercepción se representan por el producto de la matriz X , de dimensión $n \times p$, y el vector b $nx1$, y los n errores, para las n observaciones realizadas.

Utilizando el criterio de Mínimos Cuadrados donde $SSE = 0$ y aplicando una serie de transformaciones lineales se tiene:

$$X'Xb = X'y \quad (1.20)$$

ecuación a partir de la cual se estiman los coeficientes b :

$$b = (X'X)^{-1}X'y \quad (1.21)$$

siendo X' la traspuesta de la matriz X y X^{-1} la matriz inversa.

Utilizando la notación de las matrices, el vector de las respuestas ajustadas o predichas es y_{est} , cuya ecuación es:

$$y_{est} = X(X'X)^{-1}X'y \quad (1.22)$$

1.2.3 Prueba de la Hipótesis

La prueba de una hipótesis estadística con respecto a alguna característica desconocida de la población de interés es cualquier regla para predecir si se rechaza la hipótesis nula con base en una muestra aleatoria de la población.[13]

La decisión se basa en alguna estadística apropiada, de acuerdo con la cual se determina si se debe rechazar la hipótesis planteada o si, por el contrario, es conveniente aceptarla, es decir, para ciertos valores de la estadística la decisión será rechazar la hipótesis nula, dichos valores constituyen lo que se conoce como *región crítica de prueba*. Por otra parte, se debe mencionar que para que la prueba de la hipótesis sea robusta es necesario establecer una hipótesis alternativa, donde se refleje el valor posible o intervalo de valores del parámetro de interés si la hipótesis nula es falsa.

Este tipo de prueba se utiliza para probar la significancia del modelo y de los parámetros utilizados para ajustarlo; para el primer caso se realiza bajo la siguiente presunción: *La hipótesis nula se cumple si no existe relación lineal entre alguna de las variables independientes*. Y es equivalente a la siguiente ecuación:

$$H_0 : b_1 = b_2 = \dots b_i = 0 \quad (1.23)$$

$$H_1 = b_j \neq 0 \quad (1.24)$$

La ecuación 1.23 representa la hipótesis nula (H_0) y es la afirmación sobre una o más características de las poblaciones que al inicio se supone cierta; la ecuación 1.24 denota la hipótesis alternativa (H_1), es la negación de la hipótesis nula y es la propuesta por el investigador. Los terminos b_i representan el punto de intersección de la regresión.

Si H_0 es negada, entonces hay al menos una variable independiente que contribuye significativamente al modelo, y por lo tanto, se puede concluir que existe una relación funcional entre la respuesta y al menos una de las variables de entrada.

De manera similar, se puede aplicar la hipótesis a cada uno de los coeficientes del modelo b_j :

$$H_0 : b_j = 0 \quad (1.25)$$

$$H_1 : b_j \neq 0 \quad (1.26)$$

Si H_0 se niega, el coeficiente contribuye significativamente en el modelo, en caso contrario el coeficiente puede ser eliminado de la ecuación del modelo.

Para contrastar la hipótesis y determinar si es verdadera o debe ser rechazada se utiliza la *Prueba F-Snedecor*, a través de la cual se comparan las muestras de las poblaciones, esta tiene la siguiente ecuación:

$$F_0 = \frac{SSR/f_{SSR}}{SSE/f_{SSE}} \quad (1.27)$$

Donde:

SSE/f_{SSE} : es la Suma de Cuadrados de los errores dividida entre los grados de libertad del error

SSR/f_{SSR} : es la Suma de Cuadrados de la Regresión dividida entre los grados de libertad del modelo ajustado.

Cuando se desea probar la significancia de los parámetros utilizados para realizar la regresión se utiliza la Prueba de la hipótesis pero de acuerdo con la función de distribución *t-Student*, en este caso la hipótesis nula establece que al menos uno de los parámetros utilizados como regresor es diferente de cero, sin embargo, es importante mencionar que esta prueba se ve afectada por la adición o eliminación de regresores y que el hecho de tener una gran cantidad de regresores no implica la obtención del mejor modelo.

La ecuación de esta prueba es:

$$t_0 = \frac{b_j}{\sqrt{MSE * C_{jj}}} \quad (1.28)$$

Donde: C_{jj} es el elemento perteneciente al vector b de la ecuación 1.18 obtenida con Mínimos Cuadrados y que corresponde al término b_j ; el término que se encuentra dentro de la raíz cuadrada se considera como el error estándar del elemento b_j .

Los valores obtenidos al aplicar la ecuación 1.28 se someten a un análisis para determinar su significancia, este análisis consiste en calcular el *valor p*, el cual determina la probabilidad de aceptar o rechazar las hipótesis establecidas, se obtiene a través de la función de distribución de Potencia. Este valor se compara con el nivel de significancia (α) seleccionado para el estudio, por lo general se asume $\alpha = 95\%$, y es la

región crítica de prueba, si p es menor al nivel establecido se considera significativo, si es mayor se acepta la hipótesis nula y se dice que el regresor no es importante dentro del estudio.

Otros Conceptos Asociados

El proceso de Regresión, llevado a cabo a través de la técnica de Mínimos Cuadrados, involucra una gran cantidad de definiciones que permiten estimar si el modelo ajustado es el más adecuado y como el comportamiento de cada variable independiente afecta la respuesta. Entre las definiciones más importantes se encuentran:

1. Coeficiente de determinación o R^2 : Sólo puede variar entre 0 y 1, es un indicador de la proporción de variabilidad alrededor del promedio de las respuestas observadas. Un valor cercano a 1 no garantiza necesariamente un buen modelo. La expresión matemática que lo define es la siguiente:

$$R^2 = 1 - \frac{SSE}{S_{yy}} \quad (1.29)$$

2. Suma Total de Cuadrados o S_{yy} (*Total Sum of Squares*): Se define como la variabilidad total en las n observaciones, y su expresión matemática es:

$$S_{yy} = \sum_{k=1}^n (y_k - \bar{y})^2 \quad (1.30)$$

3. Suma de Cuadrados de la Regresión o SSR (*Regression Sum of Squares*): Muestra la variabilidad en la respuesta obtenida por la ecuación del modelo, se define como:

$$SSR = \sum_{k=1}^n (y_{k(est)} - \bar{y})^2 \quad (1.31)$$

4. $R_{adjusted}^2$: Cumple la misma función que el coeficiente de determinación, pero incluye en los cálculos los grados de libertad que existen en el modelo:

$$R_{adj}^2 = 1 - \frac{SSE/(n-p)}{S_{yy}/(n-1)} \quad (1.32)$$

5. Media del cuadrado de los errores o MSE (*Mean Squared Error*): Está relacionado con el error estándar del modelo y es un indicador del error de la varianza:

$$MSE = \frac{SSE}{(n-p)} \quad (1.33)$$

6. Error Estándar (*Standart Error*): Se define como:

$$StandartError = \sqrt{MSE_{x_k}(X'X)^{-1}X_k} \quad (1.34)$$

donde X corresponde a la matriz de variables de entrada.

7. PRESS (*Prediction Error Sum of Squares*): Es la suma de los cuadrados de los residuales; teniendo en cuenta que los residuales son la diferencia punto a punto entre los valores de la respuesta observada y la respuesta predicha por el modelo. Se expresa como:

$$PRESS = \sum_k e_k^2 \quad (1.35)$$

8. $R_{prediction}^2$: Indica el poder predictivo del modelo, y se define como:

$$R_{prediction}^2 = 1 - \frac{PRESS}{Syy} \quad (1.36)$$

1.2.4 Análisis de Varianza (ANOVA)

El análisis de varianza (ANOVA, del inglés Analysis of Variance), es una técnica estadística para el análisis de datos, cuyo objetivo es comparar el impacto entre diferentes factores que poseen múltiples niveles. Este análisis es función de la suma de los cuadrados entre los valores de los factores y la media.[35]

La técnica de análisis de varianza permite hacer inferencias acerca de un conjunto de parámetros relacionados con medias poblacionales. Suponga que se disponen de muestras de tamaño n que provienen de m poblaciones diferentes, entonces el análisis de varianza permitirá probar la hipótesis de que las medias poblacionales son iguales. Cuando la media de una variable aleatoria depende de un sólo factor, generalmente la muestra de la que proviene la variable, entonces se realiza un análisis de varianza de un factor. Una suposición importante que hace esta técnica es que los datos están distribuidos normalmente y con la misma varianza σ^2 (desconocida).

El caso más sencillo de ANOVA, es un modelo lineal para un factor (ver ecuación 1.37), sin embargo, de igual forma es posible determinar si las medias de un conjunto de datos difieren cuando son agrupados en múltiple factores. El modelo para este tipo de estudio depende del número de factores. La Ecuación 1.38 representa un

modelo para tres factores. Un modelo ANOVA puede tener un conjunto completo de parámetros o un subconjunto, pero los términos de interacción complejos no son incluidos a menos que también se incluyan los términos más simples de estos factores.

$$Y_{ij} = \alpha_{.j} + \epsilon_{ij} \quad (1.37)$$

donde:

Y_{ij} : es una matriz de respuesta en la que cada columna representa un grupo diferente.

$\alpha_{.j}$: es una matriz que contiene las media de cada grupo.

ϵ_{ij} : representa la matriz de errores.

$$Y_{ijkl} = \mu + \alpha_{.j} + \beta_{i..} + \gamma_{...k} + (\gamma\beta)_{ij.} + (\alpha\gamma)_{i.k} + (\beta\gamma)_{.jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl} \quad (1.38)$$

Fundamentos del Análisis de Varianza

Supóngase k muestras aleatorias independientes, de tamaño n , extraídas de una única población, con distribución normal. A partir de ellas existen dos maneras independientes de estimar la varianza de la población s^2 :

- Una llamada *varianza dentro de los grupos* (ya que sólo contribuye a ella la varianza dentro de las muestras), *varianza de error* o *cuadrados medios del error*; habitualmente representada por MSE (Mean Square Error) o MSW (Mean Square Within) que se calcula como la media de las k varianzas muestrales (cada varianza muestral es un estimador centrado de s^2 y la media de k estimadores centrados es también un estimador centrado y más eficiente que todos ellos). MSE es un cociente: al numerador se le llama suma de cuadrados del error y se representa por SSE y al denominador grados de libertad por ser los términos independientes de la suma de cuadrados.
- Otra llamada *varianza entre grupos* (sólo contribuye a ella la varianza entre las distintas muestras), *varianza de los tratamientos* o *cuadrados medios de los tratamientos*, está representada por MSA o MSB (Mean Square Between). Se calcula a partir de la varianza de las medias muestrales y es también un cociente; al numerador se le llama suma de cuadrados de los tratamientos (se le representa por SSA) y al denominador $(k-1)$ grados de libertad.

MSA y MSE, estiman la varianza poblacional en la hipótesis de que las k muestras provengan de la misma población. La distribución muestral del cociente de dos estimaciones independientes de la varianza de una población normal es una función

con los grados de libertad correspondientes al numerador y denominador respectivamente, por lo tanto se puede contrastar dicha hipótesis usando esa distribución.

Si en base a este contraste se rechaza la hipótesis de que MSE y MSA estimen la misma varianza, a su vez, se puede rechazar la hipótesis de que las k medias provengan de una misma población. Aceptando que las muestras provengan de poblaciones con la misma varianza, este rechazo implica que las medias poblacionales son distintas, de modo que con un único contraste se verifica la igualdad de k medias.

Existe una tercera manera de estimar la varianza de la población, aunque no es independiente de las anteriores. Si se consideran las kn observaciones como una única muestra, su varianza muestral también es un estimador centrado de s^2 . Se suele representar por MST, se le denomina *varianza total* o *cuadrados medios totales*, es también un cociente, donde el numerador se le llama suma de cuadrados total y se representa por SST, y el denominador $(kn - 1)$, término que corresponde a los grados de libertad.

Los resultados de un análisis de varianza se representan en una tabla como la que se muestra a continuación

Fuente de variación	Grados de Libertad	SS	MS	F
Entre Grupos	k-1	SSA	SSA/(k-1)	MSA/MSE
Dentro de los Grupos	(n-1)k	SSE	SSE/k(n-1)	
Error				
Total	kn-1	SST		

Tabla 1.1: Tabla ANOVA

1.3 Análisis de Incertidumbre

La *incertidumbre* puede definirse, de acuerdo con lo establecido con el Vocabulario de Metrología Internacional, como *un parámetro, asociado al resultado de una medida, que caracteriza el intervalo de valores que puede ser razonablemente atribuidos al medido*. El concepto de incertidumbre refleja duda acerca de la veracidad del resultado obtenido una vez que se han evaluado todas las posibles fuentes de error y que se han aplicado las correcciones oportunas. Por tanto, la incertidumbre caracteriza, de manera razonable, la dispersión de los valores medidos, y puede ser representada a través de la desviación estándar o por un intervalo de confianza

estimado sobre la base de diferentes medidas de la misma.

La cuantificación de la incertidumbre comprende basicamente tres fases, (Ver Figura: 1.6), la primera, donde se cuantifica probabilísticamente cada una de las variables de entrada, la segunda, donde se propaga la incertidumbre de las variables de entrada a través del modelo; y la tercera y última fase donde se cuantifica y se caracteriza probabilísticamente la variable de salida. La cuantificación del nivel de incertidumbre de cada una de las variables se realiza asignando al conjunto de valores disponibles de dicha variable una distribución probabilística que represente el comportamiento de la variable a modelar.

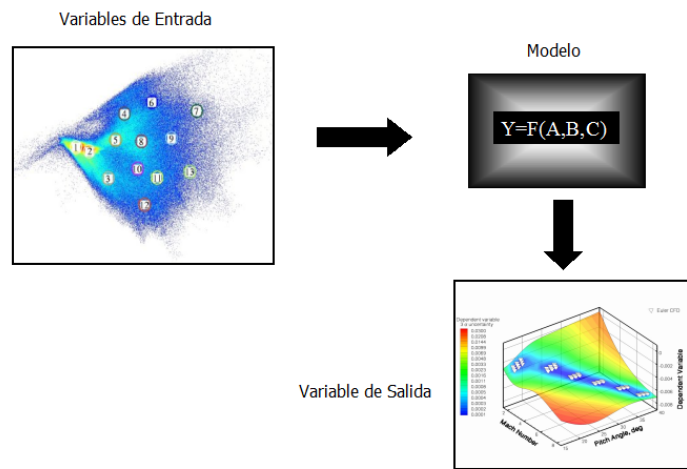


Figura 1.6: Estudio de la Incertidumbre

1.3.1 Clasificación de la Incertidumbre

Para el estudio de la incertidumbre existen diversas metodologías, pero en el siguiente trabajo sólo se analizará aquella que depende del comportamiento estadístico y del status de la misma, la cual sugiere la siguiente clasificación:

De acuerdo con el comportamiento estadístico:

- *Incetidumbre Determinística:* es aquella donde el rango de incertidumbre entre los valores mínimos y máximos es continuo, así como su impacto sobre la variable de respuesta, es decir, la respuesta varía de manera continua como una función de los parámetros de incertidumbre. Comprende coeficientes de

escalamiento, valores promedio de variables como la porosidad y la permeabilidad, correlación de espesores para modelos geoestadísticos, entre otros.

- *Incertidumbre Discreta*: corresponde a los valores que sólo pueden tomar un valor finito, e incluye la incertidumbre asociada a posibles escenarios deposicionales, el comportamiento conductivo o no de una falla.
- *Incertidumbre Estocástica*: describe el impacto no continuo de las variables de incertidumbre sobre la variable de respuesta, en este caso la incertidumbre puede tomar infinitos valores equiprobables, por ejemplo, realizaciones geoestadísticas, modelos de correlación históricos de producción.

De acuerdo con el status:

- *Incertidumbre no controlable*: corresponde a todos aquellos parámetros sobre los cuales no se puede tener control, entre ellos se encuentran los parámetros inherentes al yacimiento como lo son datos geológicos, petrofísicos, entre otros.
- *Incertidumbre controlable*: corresponde a los parámetros cuyos valores son desconocidos pero controlables, como lo son la ubicación de los pozos, tasas de producción e inyección, los cuales pueden ser controlados como parámetros de incertidumbre y también optimizados para maximizar el recobro de hidrocarburos.

1.3.2 Metodologías para el Análisis de Incertidumbre

Existen diversos métodos estadísticos para el análisis de la incertidumbre en el estudio de yacimientos de hidrocarburos, algunos de ellos son el: *Método Monte Carlo*, *Inversión Bayesiana*, *Simulación Nested*, *Método de variación relativa*, sin embargo, aún cuando muestran resultados confiables, requieren de un gran número de simulaciones para manejar eficientemente la incertidumbre asociada a los modelos, lo que aumenta la complejidad cuando se trata de estudios con una gran cantidad de información asociada. Por ello, se han desarrollado metodologías que permiten optimizar el proceso, minimizando el número de corridas y mostrando resultados confiables para cuantificar el impacto de las variables de incertidumbre, entre ellas destaca el Diseño Experimental.

1.3.3 Diseño Experimental

Todas las actividades asociadas con planear y realizar estudios de investigación tienen implicaciones estadísticas que constituyen la base para la estructura de un estudio de investigación. Si la estructura se define de manera adecuada, se obtendrá la información para la que se realizó el diseño, si la estructura presenta fallas, el estudio no funcionará bien y presentará información incompleta o errónea. Los principios estadísticos son los asociados con la recolección de aquellas observaciones que proporcionen la mayor cantidad de información para el estudio de investigación de manera eficiente.

Según Kuehl[30], el diseño experimental es el arreglo de las unidades experimentales utilizado para controlar el error experimental, a la vez que acomoda los tratamientos, todo esto para obtener la máxima información, precisión y exactitud de los resultados con el uso más eficiente de los recursos existentes.

Pautas Generales para realizar Diseños de Experimentos

[36]

- *Identificación y enunciación del problema:* corresponde a la planificación previa del experimento, y es considerada la etapa más importante. En ésta se constata que existe un problema que necesita experimentación y se trata de establecer de manera clara un enunciado para el mismo, lo que contribuye de manera sustancial para realizar una mejor comprensión de los fenómenos sometidos a estudio y la solución final del problema. Es importante tener en cuenta el objetivo global, así como los objetivos específicos que incluyen la confirmación del problema existente, el descubrimiento de una solución y la estabilidad de la misma, además de las cuestiones inherentes al experimento que deberán abordarse de manera específica. En ocasiones, es necesario darle al estudio un enfoque secuencial en el que se utilice una serie de experimentos más pequeños como estrategia para obtener una respuesta confiable.
- *Elección de los factores, los niveles y los rangos:* constituye otra parte de la sección correspondiente a la planificación previa del experimento. Existen numerosos elementos que podrían considerarse dentro de un experimento, sin embargo, no todos son necesariamente influyentes, con la finalidad de solventar esta situación existe una clasificación para dichos elementos donde, de acuerdo con su efecto, son considerados como factores potenciales del diseño o como

factores perturbadores. Para los factores potenciales de diseño una clasificación útil indica la existencia de factores de diseño, factores que se mantienen constantes y otros a los que se les permite variar. Los factores de diseño son los que resultan seleccionados para ser estudiados en el experimento. Los factores que se mantienen constantes son aquellos pueden tener efectos sobre la respuesta, pero que para los fines del experimento no resultan de interés, por lo que se mantendrán fijos en un nivel específico. Los factores que pueden variar por lo general no son homogéneos y se ignora su variabilidad de una unidad experimental a otra, por lo que se confía que la aleatorización pueda compensar este efecto, en algunas ocasiones de acuerdo con la complejidad del experimento este tipo de factores constituye un número reducido de elementos.

Por su parte, los factores perturbadores pueden tener efectos que deben ser tomados en cuenta, aún cuando no haya interés en ellos dentro del contexto del experimento. Se pueden clasificar en controlables y no controlables o de ruido. Un factor perturbador controlable es aquel cuyos niveles pueden ser ajustados por el experimentador. En caso de que un factor perturbador no sea controlable, pero se puede medir su efecto, se considera un factor de ruido; la influencia del mismo puede estimarse de acuerdo con un análisis de covarianza.

Luego que se han establecido los factores de diseño se debe elegir el rango en el cual varíaran, además de los niveles específicos en los cuales se harán las corridas del experimento; para establecer de manera adecuada estos parámetros es conveniente tener un conocimiento previo acerca del proceso. Cuando el objetivo del experimento es realizar un análisis de sensibilidad, por lo general es importante mantener reducido los niveles de los factores, y elegir una región de interés lo suficientemente amplia como para hacer varias los factores y que se aprecie el efecto de dicha variación.

- *Selección de la variable de respuesta:* para realizar esta selección se debe tener la certeza de que la variable elegida proporciona en realidad información útil acerca del proceso sometido a estudio; en algunos casos, el promedio o la desviación estándar será la respuesta, en otros se obtendrán respuestas múltiples. Este paso se considera el último dentro de lo que se refiera a la planeación previa del experimento.
- *Elección del diseño experimental:* este paso implica la consideración del tamaño

de la muestra (número de réplicas), la selección del orden de corridas adecuado para los ensayos experimentales y la determinación de si es necesario considerar la formación de bloques y otros aspectos inherentes a la aleatorización. Es importante tener claros los objetivos experimentales, y enfocarse en qué factores causan diferencias y estimar la magnitud del cambio de respuesta.

- *Realización del experimento:* en esta etapa es vital monitorear con atención el proceso para confirmar que todo se está desarrollando conforme lo planeado. En algunos casos es conveniente realizar corridas piloto para obtener la información acerca de la consistencia del material experimental, una comprobación del adecuado funcionamiento del sistema de medición, una idea aproximada del error experimental y la oportunidad de poner en práctica la técnica experimental implantada.
- *Análisis estadístico de los datos:* se sugiere darle un tratamiento estadístico a los datos para, posteriormente, medir de manera objetiva los resultados obtenidos y no hacerlo de manera intuitiva. También es útil presentar los resultados de varios experimentos en función de un modelo empírico, mediante el cual, con una ecuación que se deriva de los datos, se expresa la relación que existe entre la respuesta y los factores de diseño. El análisis residual y la verificación de la adecuación del modelo son también técnicas importantes.
- *Conclusiones y recomendaciones:* luego de cumplir con todos los pasos anteriores se deben elaborar conclusiones prácticas acerca de los resultados y recomendar acciones; en esta etapa los métodos gráficos suelen ser de utilidad, también se pueden realizar corridas de confirmación del experimento para validar las conclusiones planteadas.

Conceptos Básicos

- *Experimento:* Actividad científica planificada y diseñada para estudiar o comparar varias poblaciones.
- *Factor:* Cada una de las condiciones, que se manipulan en el estudio, capaces de afectar la variable de estudio.
- *Variable de Respuesta:* Variable de estudio.
- *Nivel:* Categorías que tiene un factor, dependiendo del diseño podrían ser nivel alto, medio o bajo.

- *Tratamiento*: Condiciones a las que se somete el experimento, y que se obtiene como combinaciones de los posibles niveles de cada factor.
- *Unidad Experimental*: Objetos o individuos que reciben el tratamiento.
- *Generador*: Es la relación de factores que da como resultado la matriz identidad $ABC = I$.
- *Alias*: Propiedad de las unidades experimentales, que define la capacidad de las mismas de interactuar entre ellas para producir un efecto igual al de otra unidad experimental independiente. Es decir, si se tiene un experimento de 3 factores A,B y C, y es imposible diferenciar entre el efecto de A y BC, se dice que A y BC son alias de acuerdo con la siguiente propiedad:
 $A * I = A * ABC = A^2 * BC$, sabiendo que $A^2 = 1$.
- *Relación de Definición de Diseño*: Se define como el conjunto de todas las columnas que son iguales a la matriz identidad. Si se tiene un experimento con 3 factores A,B y C, se dice que la relación de Definición de diseño es:
 $I = ABC$
donde I es la matriz identidad y ABC la combinación de los factores.
- *Repetición*: Observación estadísticamente independiente y adicional que se hace de un tratamiento.
- *Interacción*: Se produce cuando uno de los factores no produce el mismo efecto en la respuesta con niveles diferentes de otro factor.

Es importante mencionar que, en la ejecución de un diseño experimental existen una serie de factores no controlables en el transcurso del experimento, y que los errores que éstos ocasionan no son aleatorios durante todo el experimento, debido a que los errores en medidas sucesivas están correlacionados. Esto puede hacer que se produzcan errores sistemáticos en los resultados; para superar este problema se utiliza la técnica de aleatorización, que consiste en llevar a cabo los experimentos de forma aleatoria, asegurando de esta manera que los errores en cada nivel no estén correlacionados.

Características del Diseño Experimental

1. *Ortogonalidad*: Se dice que dos vectores son ortogonales si la suma algebraica de sus elementos es cero. Partiendo de dicha propiedad se define un Diseño Experimental como ortogonal si la suma de los elementos de cada fila de la

matriz de diseño es igual a cero. La importancia de esta característica en los diseños es que gracias a ella las estimaciones de los efectos principales y de interacción son independientes entre si. Ver Figura 1.7

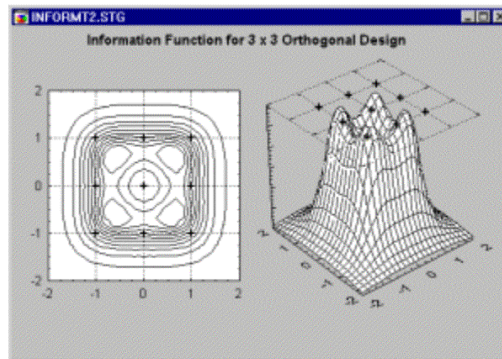


Figura 1.7: Diseño Ortogonal

2. Rotabilidad: depende de la varianza de la respuesta predicha, la cual debe ser constante en todos los puntos que son equidistantes del centro del diseño. En los diseños relativos los contornos asociados a la varianza de los valores predichos son círculos concéntricos. Esta característica permite rotar los diseños alrededor del punto central sin cambiar la varianza. Ver Figura 1.8⁴

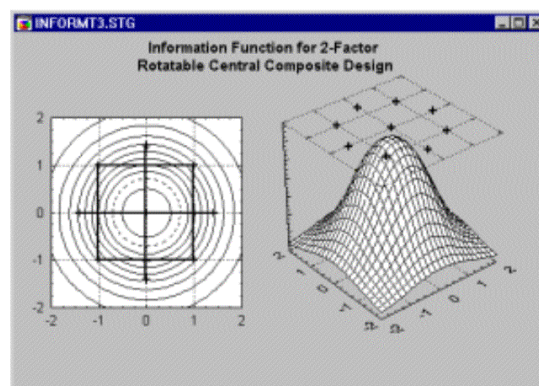


Figura 1.8: Diseño Rotable

⁴Fuente <http://www.mathworks.co.uk/cmsimages/>

Tipos de Diseño Experimental

Existe una gran variedad de diseños desarrollados, que se pueden clasificar de acuerdo con el uso de técnicas estadísticas y a su aplicación, entre ellos se encuentran:

- Diseños por Bloques Aleatorizados
- Cuadrados Latinos
- Diseños de Bloques Incompletos
- Diseños Factoriales (Full Factorial) de dos y tres niveles ($2^k - 3^k$)
- Diseños Fraccionados (Fractional Factorial) de dos y tres niveles
- Diseños Jerárquicos
- Análisis de Regresión Lineal
- Superficie de Respuesta

De acuerdo con los objetivos del presente estudio, sólo se hará breve reseña de los diseños más relevantes, como lo son diseños implementados para análisis de sensibilidad y de riesgo.

Diseños para Análisis de Sensibilidad

El objetivo de este tipo de diseño es determinar, dentro de un conjunto de variables de entrada, cuales tienen influencia en la variable de respuesta. Los más comunes son:

- Diseño Factorial (Full Factorial): Es el tipo de diseño más utilizado cuando es necesario incluir en el estudio varios factores y medir su impacto en la variable de respuesta. Se entiende por diseño factorial como aquel que en cada ensayo o réplica completa del experimento se investigan todas las posibles combinaciones de los niveles de los factores. Cuando los factores están incluidos en el diseño se dice que es un diseño factorial cruzado. El efecto de un factor se define como el cambio en la respuesta producido por el cambio en el nivel del factor, con frecuencia se le llama *efecto principal* porque se refiere a los factores de interés primario en el experimento.[36]

La principal desventaja de este tipo de diseño es que el número de combinaciones se incrementa exponencialmente con el número de variables que se

someterán a estudio. Este tipo de diseño no es recomendado cuando el número de variables es mayor a cinco. Generalmente se utilizan en experimentos donde el propósito primario sea escoger o determinar los efectos principales de las variables de incertidumbre, para establecer cuales son más significativos (análisis de sensibilidad).

Existen 2 tipos de diseños factoriales:

1. Diseño Factorial de dos niveles 2^k : considera sólo dos niveles, alto y bajo, 1 y -1 respectivamente, es completamente aleatorio, y los niveles de las variables pueden ser cualitativos y cuantitativos. Estos diseños utilizan todas las posibles combinaciones de los factores de entrada y el modelo se puede representar por la siguiente ecuación:

$$Y_{ij} = \mu + A_i + B_j + AB_{ij} + \epsilon_{ij} \quad (1.39)$$

Donde todos los términos lineales y de interacción de cada factor son evaluados.

2. Diseño Factorial de tres niveles 3^k : considera k factores con tres niveles cada uno, los cuales son el nivel bajo, medio y alto. Este diseño se propuso con el objeto de modelar las posibles curvaturas en la función de respuesta y para tomar en cuenta el caso nominal de 3 factores. Un tercer nivel para un factor continuo facilita la investigación de una relación cuadrática entre la respuesta y cada uno de los factores.

Los diseños factoriales de tres niveles presentan la misma desventaja que los diseños factoriales de dos niveles, cuando el número de factores a estudiar se incrementa, el número de corridas que se genera con las posibles combinaciones también se incrementa, convirtiendolo en un diseño deficiente.

La ecuación que modela el comportamiento de este tipo de diseño es la siguiente:

$$Y_{ijk} = \mu + A_i + B_j + AB_{ij} + C_k + AC_{ik} + BC_{jk} + ABC_{ijk} + \epsilon_{ijk} \quad (1.40)$$

- Diseño Experimental Fraccionado (Fractional Factorial): Cuando el número

de factores con los que se va a realizar el experimento se incrementa, la cantidad de corridas necesarias para realizar una réplica completa del diseño excede, con rapidez, los recursos disponibles; para solventar este tipo de problemas se desarrollaron los diseños factoriales fraccionados, en los cuales el experimentador supone, razonablemente, que ciertas interacciones de orden superior son insignificantes, para obtener información de los efectos principales y de las interacciones de orden inferior. Este tipo de diseño es de gran utilidad para los casos donde se quiera hacer un tamizado de las variables, es decir, una sensibilidad que permita identificar aquellos factores que tienen grandes efectos.

Los diseños fraccionados tienen como principios fundamentales:

1. *El principio de efectos esparcidos o escasez de efectos*: cuando existen varias variables, es posible que en el proceso sólo influyan algunas de ellas, por sus efectos principales y las interacciones de orden inferior.
2. *La propiedad de la proyección*: los diseños factoriales fraccionados pueden proyectarse en diseños más fuertes (más grandes) que el subconjunto de los factores significativos.
3. *Experimentación secuencial*: se puede combinar las corridas de dos o más diseños factoriales fraccionados para ensamblar secuencialmente un diseño más grande, que permita estimar el efecto de los factores de interés y sus interacciones.

Resolución del diseño

Es una propiedad concerniente a los diseños factoriales fraccionados, establece que un diseño es de resolución R cuando ningún efecto del factor p es alias de otro efecto que contiene $R - p$ factores. A continuación se presenta la definición correspondiente a algunas resoluciones importantes.

- Diseño de resolución III: son diseños en los que ninguno de los efectos principales es alias de ningún otro efecto principal, pero los efectos principales son alias de las interacciones de dos factores y algunas de las interacciones de dos factores pueden ser alias entre si.
- Diseño de resolución IV: son diseños en los que ninguno de los efectos principales es alias de ningún otro efecto principal ni de las interacciones de dos factores, pero las interacciones de dos factores son alias entre si.

- Diseños de resolución V: son diseños en los que ninguno de los efectos principales ni de las interacciones de dos factores son alias de otro efecto principal o interacción de dos factores, pero las interacciones de dos factores son alias de las interacciones de tres factores.

De manera general, mientras mayor sea la resolución del diseño menos restrictivas serán las suposiciones que se requieren para decidir cuán insignificantes son las interacciones de los factores.

- Diseño de Plackett Burman: Dentro de la gran variedad de diseños fraccionados que existen se encuentra el planteado por Plackett y Burman, éste es un diseño factorial fraccionado de dos niveles para estudiar K variables en N corridas, sabiendo que $k = N - 1$ y que N es múltiplo de 4. Este tipo de Diseño tiene una estructura compleja de alias, los efectos principales no están aliados entre ellos, pero si a las interacciones de segundo orden.

Diseños para el ajuste de Superficies de Respuesta

Es una colección de técnicas matemáticas y estadísticas útiles en el modelado y análisis de problemas en los que la respuesta de interés recibe la influencia de diversas variables y donde el objetivo es optimizar esta respuesta. Entre ellas se encuentran la Metodología de Superficie de Respuesta. Es un procedimiento secuencial, cuando se está en un punto de la superficie alejado del punto óptimo, el sistema presenta una curvatura moderada, por lo que se podría ajustar su comportamiento a una ecuación de primer orden, una vez que se ha encontrado la región del punto óptimo puede emplearse un modelo más elaborado como el de segundo orden. Su objetivo es identificar de manera detallada la dependencia entre los diversos parámetros y la respuesta.

La respuesta se modela de acuerdo con la siguiente ecuación:

$$Respuesta = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_1^2 + b_5x_2^2 \quad (1.41)$$

Este tipo de diseño presentan una propiedad particular: no se establecen alias entre los términos completamente cuadráticos de la ecuación de respuesta, pero si se pueden presentar alias con los términos de orden superior. Para ésto es necesario que exista un mínimo de tres niveles, para así acceder a los términos cuadráticos de orden superior.

Diseños para ajustar Superficies de Respuesta

1. Diseño Central Compuesto (DCC) (Circumscribed Central Composite Design): Consta de un diseño factorial de dos niveles (2^k) o de un factorial fraccionado de resolución V, con n_F corridas axiales o estrella, las cuales corresponden a las esquinas del cubo de diseño y dependen del número de puntos usados en la porción factorial del diseño, y n_c corridas centrales; usualmente se usa de manera secuencial donde las primeras corridas se hacen para ajustar el modelo de primer orden y luego se agregan las corridas axiales, para ajustar los términos cuadráticos.

Es importante especificar dos parámetros al momento de usar este tipo de diseño: la distancia α de las corridas axiales al centro del diseño y el número de puntos centrales n_c , donde α depende del número de puntos que tenga el diseño y está asociado a la rotabilidad del mismo. Siendo esta propiedad la que permite que el diseño sea circunscrito. Como medida general se dice que $\alpha = \sqrt{k}$.

Por lo tanto, este tipo de diseño puede presentar simetría circular, esférica o hipersférica, dependiendo del número de factores. Son los diseños que abarcan la mayor superficie de respuesta, debido a que este genera nuevos valores extremos, de esta manera el investigador puede observar si estos nuevos límites son válidos para el proceso que estudia. (Ver Figura 1.9)

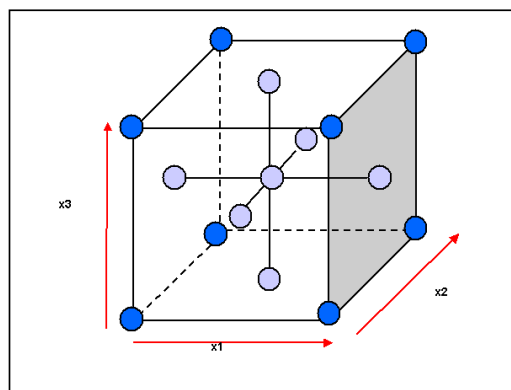


Figura 1.9: Cubo de Diseño Central Compuesto

2. Diseño Central Compuesto Inscrito (CCI)(Inscribed Central Composite De-

signs): La configuración de este diseño viene dada por los puntos axiales, es decir, representa los valores límites, a partir de los cuales se hacen los cálculos para el resto de los valores dentro del rango definido por los límites. Se obtiene dividiendo cada nivel del diseño Central Compuesto entre el valor de α para el diseño. Al igual que Diseño Central Compuesto (DCC) también posee 5 niveles, son rotativos y poseen la menor superficie de respuesta. (Ver Figura 1.10)

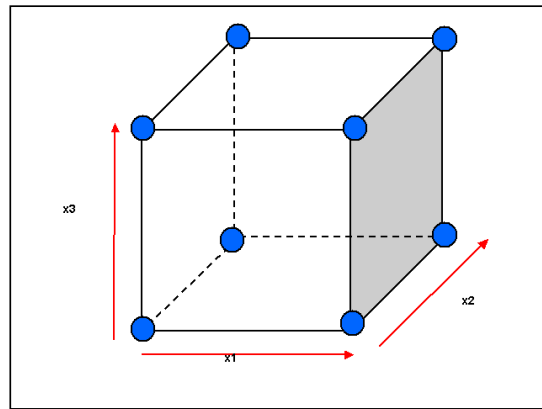


Figura 1.10: Cubo de Diseño Central Compuesto Inscrito

3. Diseño Central Compuesto con Centros en las Caras (CCF) (Faced Centered Central Composite Designs): Este diseño toma en cuenta una región cuboidal, donde $\alpha = 1$, y los puntos axiales o estrella se localizan en los centros de las caras del cubo de diseño. Posee sólo 3 niveles por factor y no es rotable, es una variante del diseño central compuesto, requiere menos puntos central para las n_c corridas. La calidad de las predicciones con este diseño es relativamente alta y no requiere puntos que se encuentren fuera de la región original, sin embargo, la estimación de los términos cuadráticos no es muy precisa. Figura 1.11
4. Diseño de Box Behnken: Es un diseño de tres niveles que se utiliza para ajustar modelos de segundo orden, requiere tres niveles por cada factor y dependiendo de la cantidad de factores puede ser rotable, con todos los puntos ubicados en una esfera de radio $\sqrt{s}$ asimismo, no contiene ningún punto en los vértices de la región cúbica creadas por los límites superior e inferior de cada variable. Esto, disminuye el poder de predicción en las esquinas del cubo de diseño, sin embargo, se considera una ventaja en aquellos casos donde los puntos de los vértices del cubo representan combinaciones de los niveles de los factores, cuya

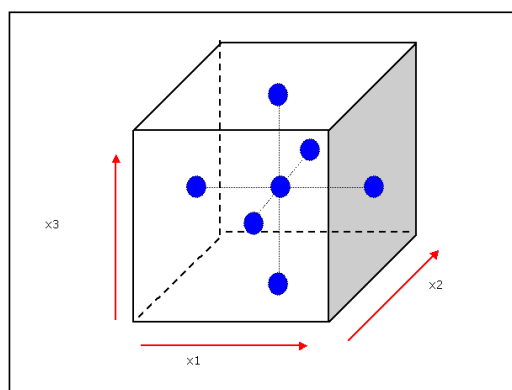


Figura 1.11: Cubo de Diseño Central Compuesto con Centros en las Caras[23]

prueba es muy costosa y en algunos casos imposible por las restricciones físicas del proceso. Puede ser utilizado para experimentación secuencial. Figura 1.12 ⁵

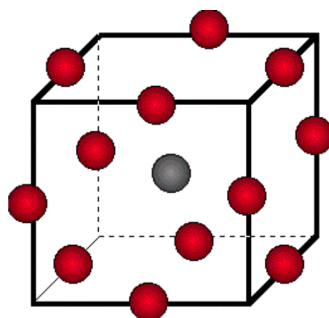


Figura 1.12: Cubo de Diseño Box-Behnken

De acuerdo con los objetivos planteados, para este estudio se seleccionó el diseño Box-Behnken, debido a que permite manejar un número de parámetros superior a 5.

1.3.4 Simulación Monte Carlo

Se define como *Simulación Monte Carlo* a un esquema que utiliza números aleatorios, pertenecientes al intervalo $U(0,1)$, para solventar algunos problemas estocásticos o determinísticos que no pueden ser tratados analíticamente, en los cuales el paso del tiempo no tiene un rol importante. Las simulaciones de este tipo

⁵Fuente <http://www.geocities.com/ResearchTriangle/System/3737/exptlidesrsm29.html>

tienen, por lo general, carácter estático.

La simulación Monte Carlo es uno de los métodos para el estudio de la propagación de la incertidumbre, donde el objetivo es determinar como la variación aleatoria, la carencia de conocimiento acerca del fenómeno, o las fuentes de error afectan la sensibilidad, ejecución o fiabilidad de un sistema que esta siendo modelado. Se clasifica entre los métodos de muestreo porque los datos de entrada son generados aleatoriamente a partir de una función de distribución de probabilidad para simular el proceso de muestreo desde una población real. La selección de la función de distribución debe ser seleccionada de manera tal que los datos de entrada cotejen lo más posible con la función, o que represente de mejor manera el estado de conocimiento que se tenga acerca del fenómeno.

Los datos generados a partir de la simulación pueden representarse en funciones de distribución de probabilidad o convertirlas gráficos de barras que reflejen el error, la fiabilidad de las predicciones, zonas de tolerancia e intervalos de confianza⁶. Ver Figura 1.13

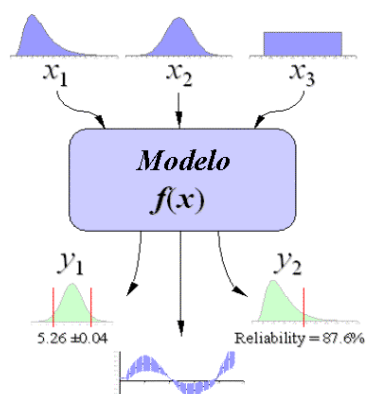


Figura 1.13: Esquema Simulación Monte Carlo

Para el desarrollo de una simulación Monte Carlo se utilizan los siguientes pasos:

- Paso 1: crear un modelo paramétrico del modelo a simular $y = f(x_1, x_2, \dots, x_q)$.
- Paso 2: generar un conjunto de variables de entrada aleatorias, $x_{i1}, x_{i2}, \dots, x_{iq}$.

⁶Fuente <http://www.vertex42.com/ExcelArticles/mc/MonteCarloSimulation.html>

- Paso 3: evaluar el modelo y almacenar las respuestas como y_i .
- Paso 4: repetir los pasos 2 y 3 desde $i = 1$ hasta n .
- Paso 5: analizar los resultados utilizando histogramas, intervalos de confianza, entre otros.

1.4 Antecedentes

En las últimas décadas, el diseño experimental y la metodología de superficie de respuesta han sido ampliamente utilizados en la ingeniería para el desarrollo de procesos óptimos. En el área de la ingeniería de yacimientos estas herramientas estadísticas se han convertido en una alternativa a los métodos tradicionales para el estudio y evaluación de diversos parámetros de incertidumbre.

La aplicabilidad de esta metodología para evaluar la incertidumbre asociada a los parámetros de yacimiento ha sido demostrada a través de un gran número de investigaciones y publicaciones. En estas se infiere, que es un método efectivo y además permite obtener resultados óptimos de una manera rápida, características de gran importancia en la realización de un estudio de yacimientos, y que permite clasificarla como un método robusto y versátil.

Entre los numerosos estudios realizados destacan:

Djean and Blanc, (1999) : En este estudio se sugiere la importancia de gestionar las fuentes de incertidumbres en la ingeniería de yacimientos, y clasificar las variables de la siguiente manera: a) los parámetros físicos en la descripción de yacimientos, considerarlos como variables de incertidumbre incontrolables; b) los parámetros referentes al desarrollo de yacimientos considerarlos como controlables. De acuerdo con esto, describir las fuentes de incertidumbre que realmente afectan la producción de petróleo, plantear un esquema de producción previsorio y optimizado. Para este fin plantean el uso de técnicas de diseño experimental, metodología de superficie de respuesta y simulación Monte Carlo, las cuales, integradas permiten a los ingenieros cuantificar la incertidumbre en la producción de petróleo, condicionada a las variables de incertidumbre en el modelado de yacimientos. Los resultados obtenidos al aplicar dicha metodología en un caso real, fueron de utilidad para construir intervalos de confianza para la producción máxima acumulada de petróleo y para optimizar

la selección de localizaciones futuras. Se demostró que las principales ventajas son: reducción de los costos computacionales y se provee de las herramientas necesarias para identificar las variables de incertidumbre de mayor importancia, construir modelos que relacionen la variable de respuesta con las variables de incertidumbre, seleccionar modelos de carácter predictivo.[19]

Chewaroungroaj et al.,(2000) : Reconocen que el pronóstico del recobro de hidrocarburos es un parámetro incierto, y que este hecho tiene tres causas fundamentales: 1)el modelo, porque es una representación imperfecta de la realidad; 2)los parámetros geológicos, porque son muestras limitadas en el espacio y/o tiempo; 3)errores en las medidas realizadas. De esta forma, un tratamiento estadístico que reconozca la diferencia entre el conocimiento y la incertidumbre es deseable para el desarrollo de yacimientos. Este artículo demuestra, que técnicas como el escalamiento, análisis de primer y segundo orden, superficies de respuesta, y los Diseños de Taguchi y Box Benhken, pueden ser de utilidad y proveer resultados razonablemente exactos de la incertidumbre estimada en el recobro de hidrocarburos. [16]

Friedman et al.,(2001) : Aplicaron técnicas de diseño experimental para analizar yacimientos de arenas canalizadas, por su relevancia en muchos de los yacimientos de aguas profundas. La metodología se aplicó sobre dos modelos de yacimientos sintéticos, uno maduro con suficiente datos de simulación y otro inmaduro con insuficiente información. Los resultados obtenidos determinaron que fue exitoso el uso del diseño experimental, para construir curvas de predicción de producción cuando la información disponible es escasa, y que los modelos polinomiales generados representan una buena analogía a los resultados obtenidos por los simuladores para evaluar la producción en sistemas de arenas canalizadas. [22]

White and Royer, (2003) : Realizaron experimentos utilizando técnicas de diseño experimental, que permitieron concluir que la simulación basada en diseño experimental enumera los factores de influencia, identifica respuestas a las sensibilidades realizadas y contribuye a determinar el rango de incertidumbre de cada parámetro que afecta la respuesta. Por otra parte, también se concluyó que el análisis de varianza, sensibilidad y la metodología de superficie de respuesta puede ser utilizada para diseñar estudios de simulación, además de reconocer que las superficies de respuesta son modelos eficientes para la simulación de yacimientos y que se pueden utilizar para el análisis de incertidumbre, estimación de parámetros, pronósticos y optimización.[50]

Salhi et al., (2005) : Se sometió a estudio uno de los campos de mayor extensión y con mayor tiempo de producción de Oman, luego de 35 años de producción y varios mecanismos de empuje, se determinó que existían reservas que aún podían ser extraídas, con este fin, utilizaron técnicas de diseño experimental y superficie de respuesta para evaluar la incertidumbre asociada a este caso. Se concluyó que el diseño experimental, es una metodología robusta para la exploración del rango de impacto y significancia de las diferentes variables de incertidumbre tanto en la historia del yacimiento como para desarrollos futuros; el diseño experimental combinado con técnicas de cotejo histórico utilizadas en el estudio dieron como resultado un esquema de trabajo poderoso disponible para evaluar el comportamiento del yacimiento maduro en tiempo real. La mayor ventaja del diseño experimental es su capacidad para estudiar como un gran número de variables y sus combinaciones tienen impacto en el comportamiento del yacimiento.[44]

Al-Shamma and Teigland., (2006) : Este artículo provee un estudio de un cotejo histórico en un yacimiento complejo utilizando técnicas de optimización global. Para esto se utilizaron técnicas de diseño experimental, gráficos Tornado y simulación Monte Carlo, herramientas que luego del proceso de optimización fueron utilizadas para el análisis de sensibilidad, creación de modelos y realización del cotejo histórico, este último se realizó con diseños experimentales Factorial 2^k y El Hipercubo Latino. Los resultados obtenidos se consideraron válidos para realizar predicciones y elaborar planes de desarrollo. También se confirmó que utilizar técnicas de diseño experimental reduce el tiempo de estudio al disminuir el número de corridas necesarias y ofrece ventajas para el manejo y estudio de la incertidumbre. [2]

García, Egle.,(2007) : Este estudio implementó técnicas de diseño experimental para la optimización del proceso SAGD en un modelo conceptual de yacimiento. Estas técnicas permitieron discriminar varios parámetros para definir aquellos que tenían mayor influencia en el proceso SAGD y posteriormente ajustarlos a un modelo, de acuerdo con la Metodología de Superficie de Respuesta, los resultados obtenidos se utilizaron para realizar predicciones y optimizaciones del proceso. [24]

Capítulo 2

Metodología

De manera general, el objetivo de la metodología presentada es el de analizar integral y sistemáticamente el impacto de las variables determinísticas y estocásticas sobre el factor de recobro, implementando técnicas de diseño experimental y modelos matemáticos de superficie de respuesta. Para esto se utilizará el nivel de investigación descriptivo, el cual consiste en la caracterización de un hecho, fenómeno o grupo con el fin de establecer una estructura o comportamiento.[4]

En cuanto al diseño de la investigación, se presenta una metodología de investigación experimental, proceso que consiste en someter un objeto o grupo de individuos a determinadas condiciones o estímulos (variables independientes), para observar los efectos que se producen(variables dependientes).[4]

2.1 Definición de Parámetros de Incertidumbre

En esta etapa de la investigación se realizó una revisión de la literatura disponible relacionada con el diseño de experimentos y su inclusión en la simulación numérica de yacimientos. Para esto se consultaron investigaciones precedentes, referentes al problema planteado, considerándose éste como el punto de partida para la definición de los objetivos, alcances, límites y justificación del presente estudio.

Adicionalmente, se realizó una breve reseña del modelaje estocástico de yacimientos y geoestadística, además de hacer una compilación de los conceptos estadísticos necesarios para el desarrollo y entendimiento del análisis a realizar. También se realizó una revisión de trabajos previos que se han desarrollado en esta área.

2.2 Clasificación de las variables de acuerdo con la incertidumbre asociada

Según lo indicado en el Capítulo 1 (Fundamentos Teóricos) las variables presentan 3 tipos de incertidumbre asociada, de acuerdo con su comportamiento estadístico. En esta sección se detalla que tipo de incertidumbre presentan los 14 parámetros sometidos a estudio.

2.3 Rangos de Incertidumbre

De acuerdo a la revisión bibliográfica y a lo establecido en el Apéndice B, los rangos de incertidumbre evaluados, tanto para el cálculo de las realizaciones como para el diseño experimental, son:

Gravedad API : [0.1-0.5 grados API].

Gravedad específica del gas : el rango a utilizar es [0.01% - 1 %] de la lectura realizada. Esto corresponde a [0.0001-0.0123 unidades de gravedad específica].

Relación gas-petróleo en solución : [0.1 %, - 2 %] de la lectura realizada. Lo cual es equivalente a [1.476-29.52 PCN/BN].

Temperatura del yacimiento : [0.04-0.9 °F]. Esto es equivalente en °C, aproximadamente, a [0.02-0.5°C].

2.4 Generación de los modelos sintéticos de yacimiento

Para la evaluación de la metodología desarrollada y el flujo de trabajo planteado se generaron modelos conceptuales de yacimiento fluvial utilizando métodos geostatísticos, desarrollados e implementados con anterioridad en MATLAB®, a través del programa *SReM*. Estos modelos de yacimiento constaron de una realización de facies por 5 realizaciones de propiedades, de esta manera se garantiza la creación de suficientes modelos como para poder clasificarlos de acuerdo con su volumen poroso, en escenarios con volumen poroso alto, medio y bajo. Las características de la malla de alta resolución se presenta a continuación en la tabla 2.1:

Una descripción general del funcionamiento de la metodología utilizada se presenta a continuación.

Número de celdas en x	100
Número de celdas en y	100
Número de celdas en z	10
Tamaño de las celdas en x , [ft]	100
Tamaño de las celdas en y , [ft]	100
Tamaño de las celdas en z , [ft]	10

Tabla 2.1: Características de la malla de alta resolución

2.4.1 Modelado de Facies

Para la reproducción del ambiente de canales fluviales, se utilizó un algoritmo de simulación basada en objetos, el cual permite asignar de manera discreta y no condicional, las facies a objetos generados y distribuidos estocásticamente. Para este caso se simularon dos facies, lutita y arena, en un ambiente fluvial de manera tal de reproducir de la forma más exacta posible las formas observadas en los ambientes naturales. Las proporciones de las facies se presentan a continuación en la tabla 2.2:

Facies	Proporción
Canales de Arena	60%
Llanura de inundación (Lutita)	40%

Tabla 2.2: Proporciones globales de las facies presentes en los modelos de yacimiento

Es importante mencionar que en esta etapa todas las realizaciones producidas reproducen la estadística de los datos, es decir, todas las realizaciones tienen un histograma, media y varianza similar entre ellas que representan las proporciones globales de cada una de las facies presentes.[40]

2.4.2 Modelado de Propiedades

El paso siguiente a realizar el modelado de facies es simular las propiedades petrofísicas independientes para cada una de las facies, para luego hacer el modelo conjunto mediante un proceso de condicionamiento de las propiedades a las facies. Este proceso de condicionamiento funciona asignando a las celdas ceros y unos, de acuerdo con la zona correspondiente (arena o lutita).

En este caso, sólo se simularon la porosidad y la permeabilidad, ésto porque son las propiedades necesarias para, posteriormente, realizar simulación numérica.

Facie/Porosidad[Fracción]	Mínimo	Máximo	Varianza
Arena	0,1	0,3	0,05
Lutita	0,02	0,06	0,02

Tabla 2.3: Parámetros utilizados para la generación de la porosidad

Para generar la porosidad se utilizaron los siguientes parámetros, tabla 2.3:
 La generación de permeabilidades se hizo a partir de regresiones, que relacionan la porosidad con el logaritmo de la permeabilidad, esto por el carácter lognormal que se evidencia en los histogramas de permeabilidad. Las regresiones empleadas se muestran en la tabla 2.4

Facie	Regresión
Arena	$k = 1,000000 * \exp(23,02585 * \text{poro})$
Lutita	$k = 0,630957 * \exp(23,025851 * \text{poro})$

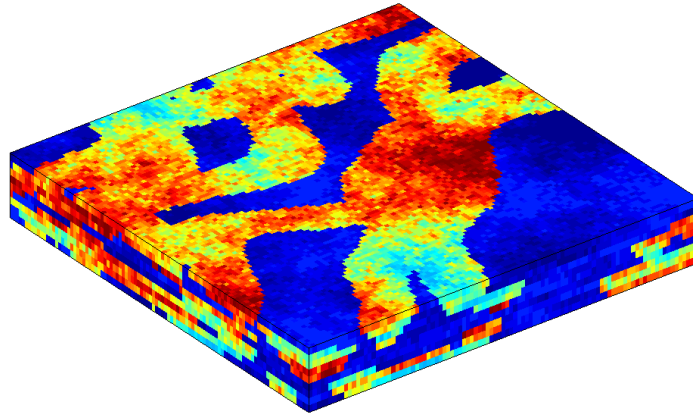
Tabla 2.4: Ecuaciones utilizadas para el cálculo de la permeabilidad

2.4.3 Escalamiento

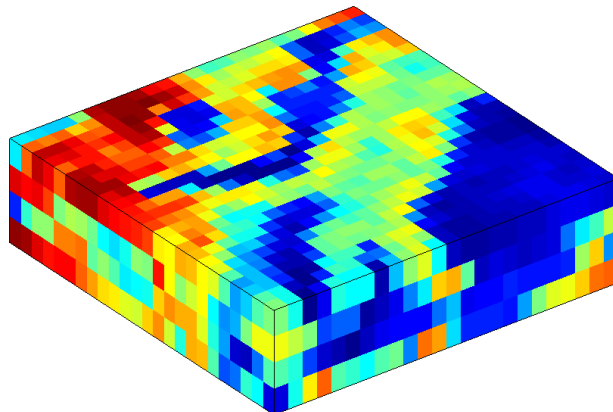
Una vez que se obtiene el modelo condicionado, éste debe ser escalado para poder ser empleado en la simulación.

El escalamiento es un proceso que permite llevar modelos de yacimiento de alta resolución ($\sim 100.000.000$ de celdas) ver Figura 2.1(a), generados mediante técnicas geoestadísticas, a modelos de escala que sean manejables en simulación de yacimientos (~ 100.000 celdas) ver Figura 2.1(b), teniendo en cuenta que debe mantenerse en la medida de lo posible la respuesta de flujo del modelo de alta resolución. Ver Figura 2.1

Para tal fin, el escalamiento de la porosidad se realiza mediante un promedio aritmético, mientras que para la permeabilidad se realiza el escalamiento en cada dirección, resolviendo las ecuaciones de flujo monofásico en estado estable con condiciones de borde de no-flujo. Luego de obtener los modelos escalados, se realiza una corrida de inicialización de ECLIPSE 100. Las características de la malla de los modelos escalados se presentan en la tabla 2.5. [40]



(a) Modelo de alta resolución



(b) Modelo de baja resolución

Figura 2.1: Escalamiento

Número de celdas en x	25
Número de celdas en y	25
Número de celdas en z	5
Tamaño de las celdas en x , [ft]	400
Tamaño de las celdas en y , [ft]	400
Tamaño de las celdas en z , [ft]	20

Tabla 2.5: Características de la malla del modelo escalado

2.4.4 Propiedades de los modelos de yacimiento

De manera general todos los modelos de yacimiento se realizaron de acuerdo a la de la tabla 2.6), el resto de las propiedades se consideraron como factores potenciales de diseño y su valores se encuentran en la tabla 2.8:

Presión Inicial [psia]	7000
Propiedades PVT (Correlación utilizada)	Standing
Saturación de Agua connata (swc)	0.15
Saturación de Petróleo irreducible al Agua (sorw)	0.2
Saturación de Gas Crítica (sgcr)	0.05
Saturación de Petróleo irreducible al Gas (sorg)	0.15
End Point Permeabilidad relativa del Petróleo al Agua (krocw)	1
End Point Permeabilidad relativa del Agua (krwro)	0.3
End Point Permeabilidad relativa del Gas (krgcw)	1
End Point Permeabilidad relativa del Petróleo al Gas (krogro)	1

Tabla 2.6: Propiedades Generales para los modelos de yacimiento

2.5 Esquema de Explotación

Los modelos de yacimiento generados presentan un esquema de explotación constituido por 6 pozos: 2 verticales, 2 desviados y 2 multilaterales de 2 y 3 brazos (Figura 2.2). Todos con diámetro igual a 0.375" y con un factor de daño igual a cero. La descripción detallada de cada pozo se presenta en la tabla 2.7, allí se indica su posición en la malla (i,j,k), CF el cual es el Índice del pozo (Well Index), el producto KH (permeabilidad por espesor) y Ro, el cual es el Radio de Peaceman [40]

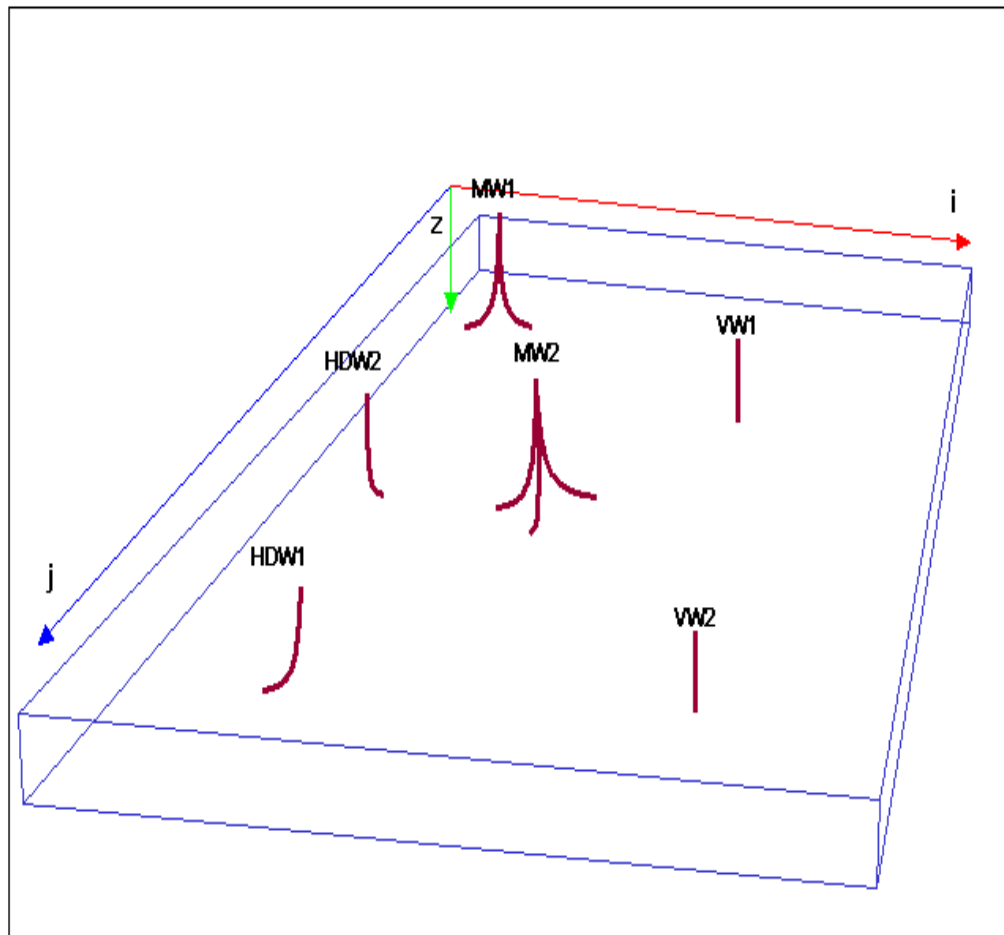


Figura 2.2: Esquema de Explotación Utilizado

Pozo	i	j	k	CF	KH (mD*ft)	Ro
'VW1'	10	10	1	2.326	1.985.533	79.196
'VW1'	10	10	2	1.075	917.842	79.216
'VW1'	10	10	3	2.027	1.730.292	79.235
'VW1'	10	10	4	2.476	2.113.782	79.212
'VW1'	10	10	5	1.267	1.081.822	79.232
'VW2'	18	20	1	0.157	135.795	84.532
'VW2'	18	20	2	0.055	47.014	79.358
'VW2'	18	20	3	0.707	603.440	79.241
'VW2'	18	20	4	6.563	5.602.720	79.196
'VW2'	18	20	5	2.350	2.006.211	79.223
'HDW1'	23	7	1	0.182	129.502	28.806
'HDW1'	23	7	2	12.434	8.684.092	26.365
'HDW1'	22	8	2	1.185	841.201	28.599
'HDW1'	22	8	3	0.334	233.518	26.678
'HDW1'	22	8	4	0.194	137.820	28.890
'HDW1'	22	8	5	1.193	849.203	29.001
'HDW1'	21	9	5	0.019	13.524	28.435
'HDW2'	17	4	1	0.234	166.263	28.814
'HDW2'	17	4	2	0.108	76.849	28.753
'HDW2'	18	4	2	0.236	177.245	38.479
'HDW2'	18	4	3	0.179	127.176	28.806
'HDW2'	18	4	4	0.084	63.644	40.944
'HDW2'	19	4	4	0.148	105.508	28.762
'HDW2'	19	4	5	0.069	49.210	28.753
'MLW1'	5	3	1	0.231	164.102	28.810
'MLW1'	5	3	2	0.523	372.268	28.910
'MLW1'	6	3	2	2.384	1.689.420	28.350
'MLW1'	6	2	2	0.259	174.746	22.137
'MLW1'	6	2	3	0.212	150.346	28.510
'MLW1'	6	2	4	0.800	528.988	20.272
'MLW1'	7	2	4	0.747	493.444	20.173
'MLW1'	7	2	5	1.589	1.125.559	28.321
'MLW1'	7	1	5	1.081	768.014	28.777
'MLW1'	3	4	1	2.083	1.142.678	9.119

'MLW1'	2	4	1	1.391	962.166	25.149
'MLW1'	2	4	2	4.852	3.442.575	28.521
'MLW1'	2	3	2	0.177	126.104	28.887
'MLW1'	2	3	3	0.717	489.551	23.570
'MLW1'	1	3	3	0.076	54.202	28.804
'MLW1'	1	3	4	1.311	861.994	19.751
'MLW1'	1	3	5	5.573	3.907.401	26.884
'MLW1'	4	5	1	0.052	36.763	28.870
'MLW1'	4	6	1	0.042	26.124	15.332
'MLW1'	5	6	1	0.152	108.152	28.686
'MLW1'	5	6	2	5.153	3.652.816	28.392
'MLW1'	5	6	3	0.072	50.655	27.123
'MLW1'	5	7	3	5.742	4.089.747	29.083
'MLW1'	5	7	4	9.813	6.943.880	28.149
'MLW1'	5	8	4	1.344	947.734	27.681
'MLW1'	5	8	5	2.008	1.422.356	28.306
'MLW2'	7	15	1	0.566	450.779	52.819
'MLW2'	7	15	2	1.712	1.220.756	29.231
'MLW2'	8	15	2	1.510	1.070.868	28.502
'MLW2'	8	15	3	15.165	10.647.031	27.064
'MLW2'	8	15	4	0.292	210.167	30.822
'MLW2'	9	15	4	12.550	8.661.499	24.870
'MLW2'	9	15	5	1.139	799.102	27.004
'MLW2'	4	15	1	0.131	92.896	28.759
'MLW2'	4	15	2	0.002	1.286	28.822
'MLW2'	3	15	2	0.119	84.595	28.805
'MLW2'	3	15	3	1.836	1.256.462	23.839
'MLW2'	2	15	3	0.387	256.521	20.461
'MLW2'	2	15	4	19.100	13.260.431	25.605
'MLW2'	2	15	5	0.051	36.171	28.894

Tabla 2.7: Descripción de Pozos

Se utilizó un sistema integrado subsuelo superficie para una instalación Costa Afuera (Ver Figura 2.3) para la simulación, a través de la opción ECLIPSETMNetworks donde se incluyó además de la información correspondiente a los pozos, los datos

del sistema. Para cada pozo el sistema subsuelo-superficie se compone de: el pozo, el cabezal, una línea de flujo hasta la plataforma (riser); una vez que los fluidos provenientes de cada uno de los pozos llegan a la plataforma se unen y pasan a la línea de exportación, la cual llega a un separador colocado a 10 Km.

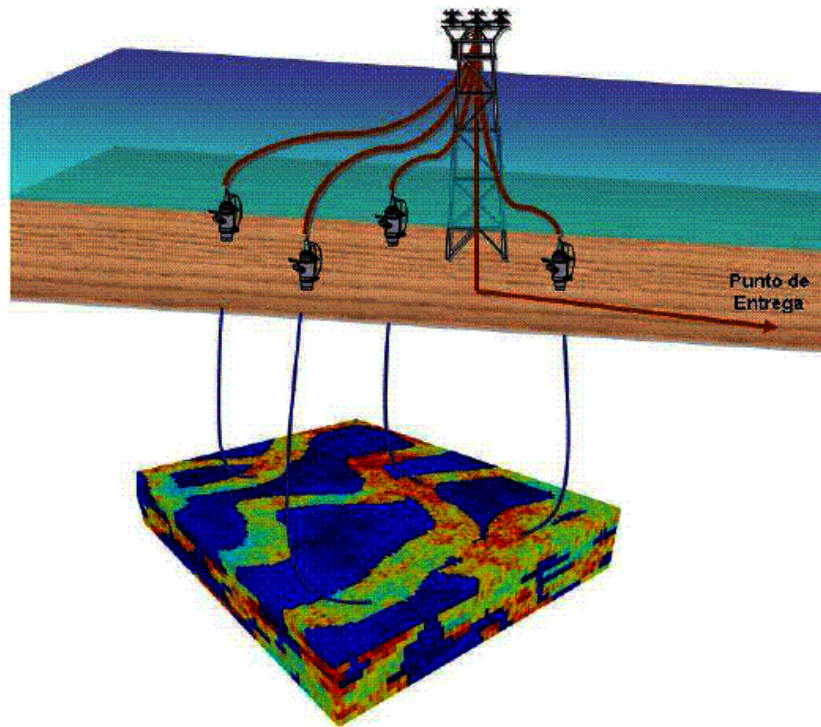


Figura 2.3: Sistema Integrado Subsuelo Superficie para una instalación Costa Afuera

Con el uso de esta opción los cálculos de las caídas de presión y temperatura a lo largo de las líneas de flujo del sistema (vertical, riser, exportación), **ECLIPSETM** hace uso de las tablas VFP's (Vertical Flow Performance) [56], las cuales se generaron previamente, con el uso de las correlaciones de Beggs-Brill para las líneas de producción y Hagerdon Brown para el flujo en las tuberías verticales de los pozos; utilizando el mismo modelo de fluidos del yacimiento.

Controles utilizados:**Condiciones de Borde**

Se utilizó como condición de borde del sistema la presión del separador $P_{sep} = 150\text{ lpc}$, a partir de ésta el simulador realiza cálculos de caída de presión y temperatura en las líneas de flujo, en concordancia con el resto de los controles establecidos y de las curvas VPF.

Como condiciones de borde de los pozos se estableció que su Presión de fondo fluyente debería ser $P_{wf} = 500\text{ lpc}$ y la tasa máxima de cada uno de ellos $q = 5000\text{ BN/d}$

Controles de Producción

Todos los pozos del esquema están controlados por la Presión del Separador, cuyo valor es $P_{sep} = 150\text{ lb}$ y una tasa máxima para el campo de $q = 10000\text{ BN/d}$, para ello ECLIPSETM tiene un algoritmo que calcula la tasa por pozo de manera tal de satisfacer la restricción de la tasa máxima del yacimiento. Para dichos cálculos se utiliza el potencial de producción de cada pozo y las tablas VPF.

Controles Económicos de los pozos

Para todos los pozos se aplicaron como controles un corte de agua de 90% y una tasa $q = 10\text{ BN/d}$, una vez que alguno de los pozos alcanzaran estas condiciones eran cerrados.

2.6 Implementación de la Metodología de Diseño Experimental

Aplicando la metodología descrita en el Capítulo I, (Pautas Generales para realizar Diseños Experimentales), el proceso de implementación constó de las siguientes etapas:

1. Identificación y enunciación del problema: En esta etapa se determinó que el problema que se debía someter a estudio era la influencia de la incertidumbre en las estimaciones de factor de recobro, para ello se estableció que el problema se enunciaría de la siguiente manera: *Evaluación de la Incertidumbre en un Modelo Estocástico de Yacimiento Fluvial Mediante Técnicas de Diseño Experimental*

2. Elección de los factores, niveles y rangos: Inicialmente se consideraron 14 parámetros como factores potenciales de diseño, ellos se agruparon de la siguiente manera (Ver Tabla 2.8):

- Propiedades del Yacimiento: porosidad, relación kv/kh , acuífero y la permeabilidad, para la porosidad se realizó una selección de escenarios con porosidad alta, media o baja, de acuerdo con los resultados estocásticos obtenidos por el programa SReM, el cual será descrito más adelante. La relación kv/kh , el acuífero y la permeabilidad fueron parámetros que variaron en orden de magnitud, para ello se utilizaron multiplicadores, de esta forma se garantizó que la variación fuese igualmente espaciada al momento de considerar los valores altos, medios y bajos de dichos parámetros.
- Propiedades PVT: Rsi, Gravedad API, Gravedad Específica del Gas y la Temperatura del yacimiento. Las propiedades se determinaron de acuerdo con las ecuaciones descritas en el Apéndice A, implementadas en el módulo bopvt2ecl de PRiSMa.
- Propiedades Roca Fluido: Exponente Corey para el agua, para el petróleo respecto al agua, para el gas, para el petróleo respecto al gas, Constante de la ecuación de Presión Capilar, para el sistema petróleo agua y para el sistema gas petróleo. Las ecuaciones correspondientes a estas propiedades se explican en los Apéndices C,D.

Los niveles considerados para que las variables variaran fueron alto, medio y bajo, los cuales se establecieron de acuerdo con los Rangos de Incertidumbre descritos anteriormente.

Posteriormente, se realizó un análisis de sensibilidad para establecer cuales parámetros se debían considerar como factores de diseño, cuales debían permanecer constantes y a cuales se les debía permitir variar. Se estableció que sólo se haría uso de un factor perturbador controlable: las condiciones operacionales, es decir, diámetro interno de la tubería de producción y la presión de entrega del sistema.

3. Selección de la variable de respuesta: Para medir el impacto de la incertidumbre de los parámetros seleccionados se utilizó el Factor de Recobro.
4. Elección del Diseño Experimental: De acuerdo con los objetivos que motivaron

	Parámetro	Bajo	Medio	Alto	
1	poro	-1	0	1	Propiedades de yacimiento
2	multperm (10^x)	-0,5	0	0,5	
3	kv/kh (10^x)	-2	-1	0	
4	Acuifero ($Vr * 10^x$)	0	1	2	
5	Rsi	980	1000	1020	PVT
6	API	21,5	22	22,5	
7	Gas Grav	0,693	0,7	0,707	
8	T	179	180	181	
9	nw	4	5	6	Propiedades roca-fluidos
10	now	2	3	4	
11	pcowconst	0	5	10	
12	nog	3	4	5	
13	ng	1	2	3	
14	pcgoconst	0	0,5	1	

Tabla 2.8: Factores Potenciales de Diseño

a la implementación de esta metodología y a las investigaciones previas se determinó que se implementarían 2 diseños experimentales, se seleccionó el Diseño Plackett Burman para realizar el análisis de sensibilidad, esto por el gran número de parámetros que se consideraron como factores potenciales de diseño; el Diseño Box Benhken para el análisis de incertidumbre de los parámetros de mayor importancia, obtenidos previamente con el análisis de sensibilidad.

- Realización del Experimento: Comprende 2 fases; la primera corresponde al análisis de sensibilidad, para la cual se hicieron 15 matrices de diseño Plackett Burman y se ajustaron el mismo números de ecuaciones, con los parámetros de mayor influencia en el factor de recobro. La segunda fase corresponde al análisis de incertidumbre, para ésta se realizaron 5 matrices de diseño Box Behnken, con la finalidad de ajustar 5 ecuaciones de segundo orden correspondientes a las superficies de respuesta.
- Análisis Estadístico de los resultados: Para la primera fase se utilizaron herramientas como el Análisis de la Varianza, la Prueba de la hipótesis, tanto para el modelo ajustado como para los parámetros utilizados como regresores y Diagramas Pareto. Para la segunda fase se utilizó el Análisis de Varianza para determinar cuán bueno fue el ajuste para el modelo utilizado y Simulación Monte Carlo para estimar las curvas de distribución de probabilidad para el Factor de Recobro.

7. Elaboración de conclusiones y recomendaciones: se plantearon recomendaciones y se hicieron conclusiones de acuerdo con los objetivos propuestos y con los resultados obtenidos.

2.7 Flujo de trabajo propuesto

El flujo de trabajo propuesto se divide en 2 fases, las cuales integran y automatizan todos los pasos previamente descritos; la primera corresponde al análisis de sensibilidad, la segunda al análisis de incertidumbre.

- **Análisis de Sensibilidad:** Esta fase se llevó a cabo con los 14 parámetros considerados como factores potenciales de diseño, con éstos se generaron 15 matrices de diseño de tipo Plackett Burman. La ejecución de las simulaciones se realizó con la disposición de parámetros que indicara cada fila de las matrices de diseño y se realizó en tres partes, primero se generaron 5 modelos de yacimiento (1 realización de facies y 5 de propiedades) y posteriormente generaron las propiedades PVT y de permeabilidades relativas. De los 5 modelos de yacimiento sólo se tomó uno, el que tuviese Volumen Poroso promedio alto o bajo, esto de acuerdo con lo que la matriz de diseño estableciera. Una vez generada toda la información necesaria se integró y se realizaron las simulaciones correspondientes al yacimiento con sus fluidos y al flujo de los mismos. Se consideró como variable para medir el comportamiento de cada uno de los modelos la producción de petróleo acumulada (N_p). Ver Figura 2.4

Con los resultados de las simulaciones se procedió a realizar un análisis estadístico, a través de mínimos cuadrados se ajustaron ecuaciones de primer orden, una por matriz, y con la Prueba de la hipótesis para los regresores, se determinó cuáles de ellos eran significativos o poseían una influencia importante en la variable de respuesta, los que poseían una influencia significativa fueron los seleccionados para la siguiente etapa.

- **Análisis de Incertidumbre:** Una vez seleccionados los parámetros que se considerarían factores de diseño, se procedió a cuantificar su efecto sobre la variable de respuesta, para ello se crearon 5 matrices de diseño de tipo Box Behnken. De acuerdo con lo estipulado en dichas matrices se hicieron variar los valores de los parámetros en sus niveles alto, medio y bajo. Los parámetros que no fueron seleccionados pasaron a formar parte de los parámetros considerados

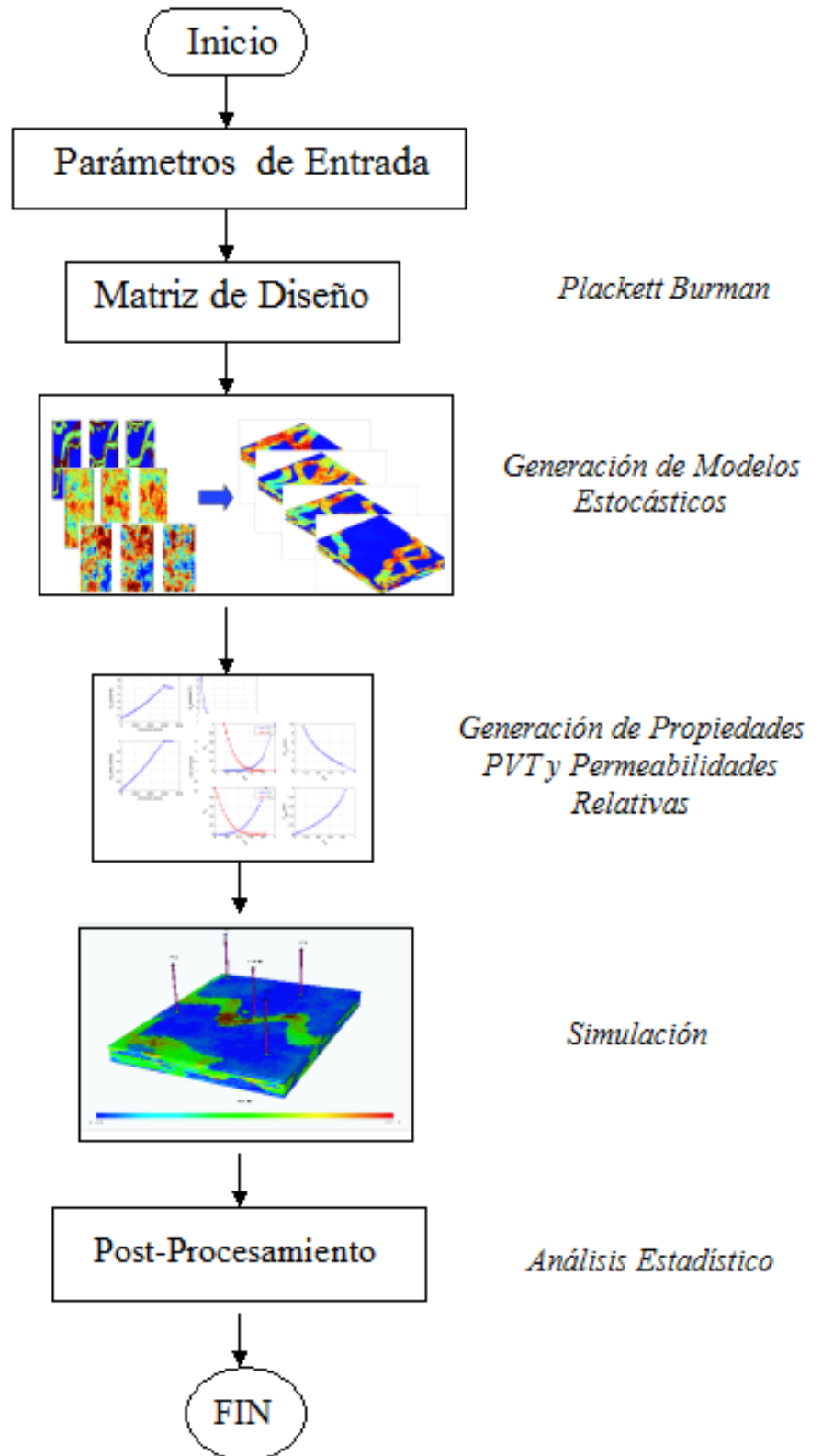


Figura 2.4: Flujo de Trabajo Análisis de Sensibilidad

constantes, y los valores para éstos fueron los correspondientes a su nivel medio.

Al igual que en el análisis de sensibilidad, se consideró como variable para medir la respuesta de las simulaciones la producción de petróleo acumulada; los resultados de las simulaciones fueron sometidos a un análisis estadístico, donde a través de mínimos cuadrados se ajustaron ecuaciones de segundo orden o superficies de respuesta.

Para cada uno de los factores de diseño se seleccionaron funciones de distribución y con la superficie de respuesta obtenida se realizó Simulación Monte Carlo, esto generando números aleatorios para cada parámetro y evaluándolos en la superficie de respuesta. Con esta metodología se caracterizó estadísticamente el comportamiento de la variable de respuesta, lo que permitió crear funciones de distribución y cuantificar la incertidumbre asociada a los factores de diseño. Ver Figura 2.5

2.8 Descripción de las herramientas utilizadas

2.8.1 ECLIPSE™

ECLIPSE se originó a partir de ECL en la década de los 70, ECL se especializaba en adquisición de datos sísmicos y control de calidad, luego se diversificó hacia el modelado de flujo dinámico, lo que se convirtió en una ventaja, y, aunque habían numerosos simuladores de yacimientos para la época ninguno consideraba el uso de modelos de pozos completamente implícitos. Teniendo a ECL como base, los desarrolladores decidieron crear un nuevo producto que tuviese modelos de pozos de manera implícita, siendo este el origen de ECLIPSE, programa que no ha parado su crecimiento, en pro de satisfacer las necesidades de la industria.[52]

La suite de ECLIPSE™ consiste en tres simuladores: ECLIPSE 100 especializado en el modelaje de petróleo negro, ECLIPSE 300, que se utiliza para modelado composicional con opción para modelaje térmico, y FrontSim, utilizado en el modelaje de líneas de flujo. ECLIPSE 100 es un simulador completamente implícito, trifásico, tridimensional, con fines generales, que incluyen controles de pozos versátiles, grupos y controles económicos, funcionalidades de perforación y secuencias de tareas automáticas. Se puede utilizar para crudo muerto y tiene opciones de gas condensado.

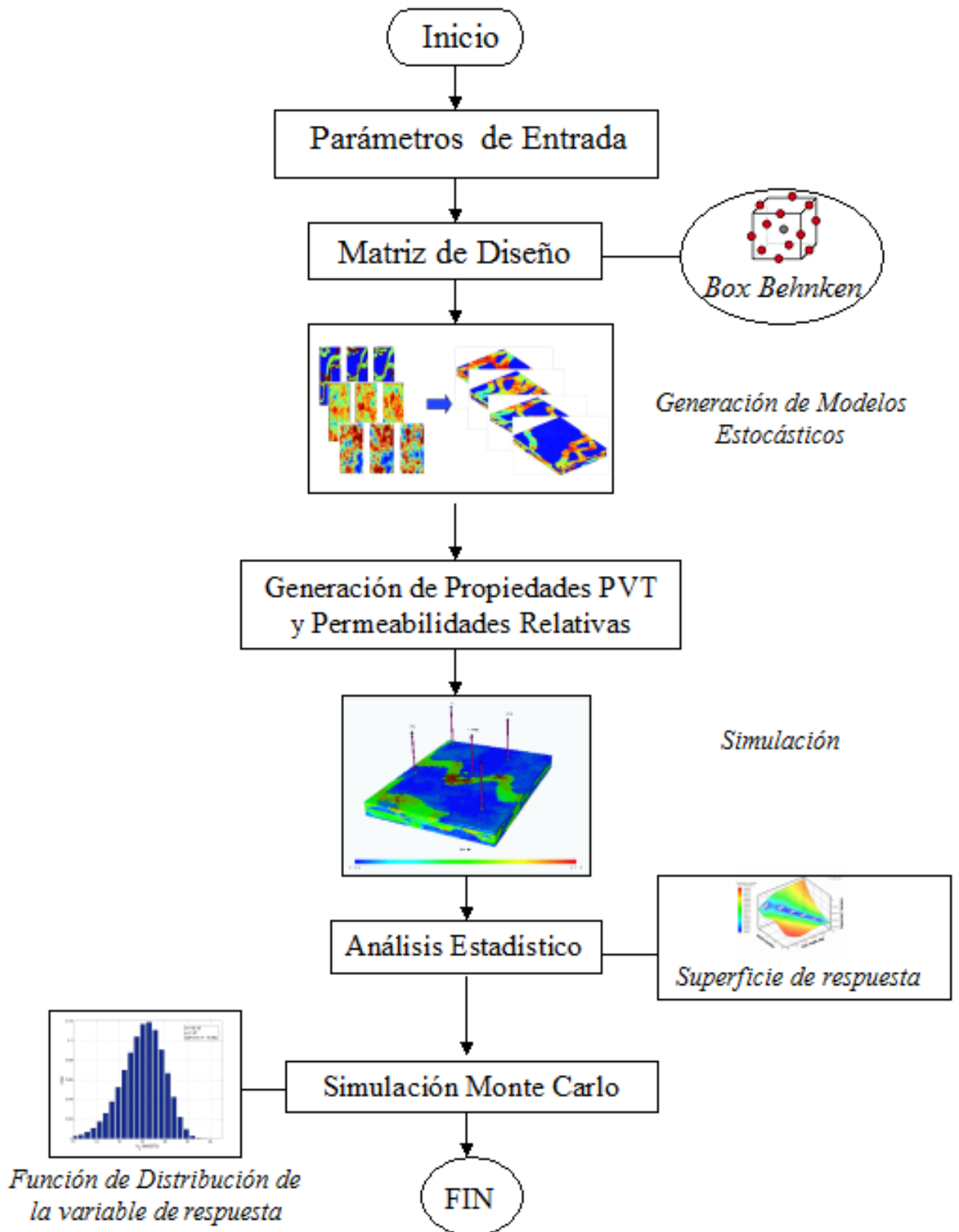


Figura 2.5: Flujo de Trabajo Análisis de Incertidumbre

Para realizar simulaciones en ECLIPSE™ se requiere de un archivo de datos de entrada, el cual se divide en secciones, cada una de las cuales, esta indicada por *keywords*. Ver Figura 2.6

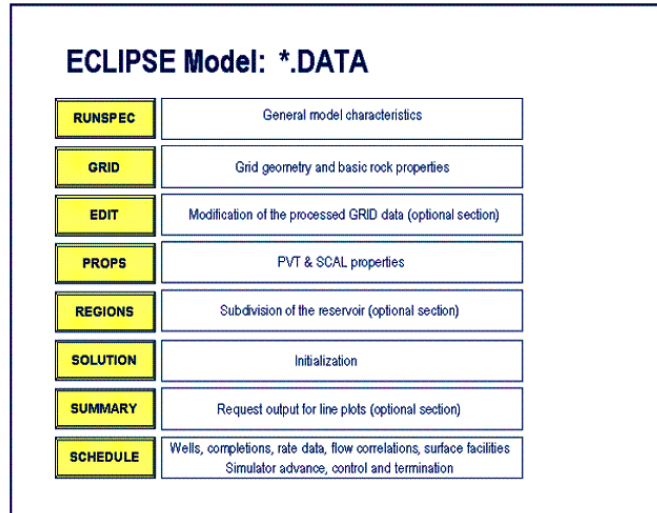


Figura 2.6: Modelo de un archivo de entrada de ECLIPSE™

A continuación se hace una breve descripción de los *keywords* que forman parte del archivo de entrada para hacer simulaciones en ECLIPSE™:

- **RUNSPEC**: En esta sección se especifica el título de la simulación, las dimensiones de la malla, las fases presentes en el yacimiento a simular, los fluidos presentes en el modelo y ciertos comandos que permiten establecer características especiales de la simulación. Además tiene como propósitos principales establecer la localización de varios componentes de la simulación en el área de la memoria principal, y especificar características básicas de la simulación, como por ejemplo, la fecha de inicio de la simulación.
- **GRID**: En esta sección se especifican las propiedades básicas de la roca y la geometría de la malla, datos que conforman el modelo estático, mapas estructurales, porosidades, permeabilidades y espesores netos de las arenas. Estas propiedades se usan para calcular el volumen poroso y la transmisibilidad, y así crear la malla a simular. Los tipos de mallas que se pueden seleccionar se muestran a continuación en la Figura¹ 2.7

¹Imágenes Extraídas de Eclipse Black Oil Reservoir Simulation Training and Exercise Guide Version 1.0

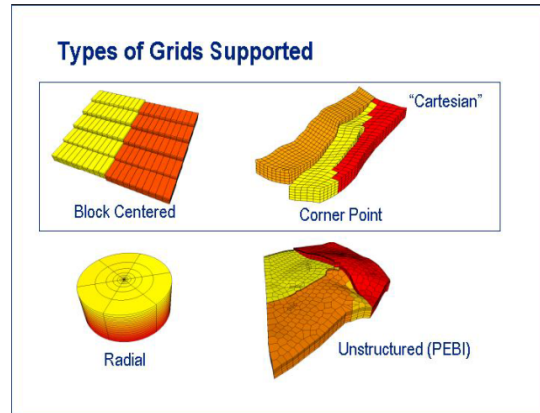


Figura 2.7: Mallas disponibles para modelar yacimientos

- *PROPS*: En esta sección se especifican los parámetros PVT, presiones capilares, permeabilidades relativas, densidades de los fluidos y compresibilidad de la roca, esto con la finalidad de hacer balance de masa, el cual es una parte importante dentro del proceso de simulación.
- *SOLUTION*: En esta sección se definen las condiciones iniciales del yacimiento, las cuales serán utilizadas al comienzo de la simulación, esto para determinar la distribución de los fluidos presentes en el yacimiento.
- *SUMMARY*: Esta sección se utiliza para especificar las variables que serán escritas en los archivos de salida, luego de cada paso de tiempo de la simulación. Los parámetros seleccionados se pueden visualizar en forma de gráficos.
- *SCHEDULE*: Esta sección se utiliza para especificar las condiciones de operación de los pozos a ser simulados, los nodos de presión que se utilizarán y las fechas. Se utiliza en modo Cotejo Histórico o Predicción, donde se especifican los pozos presentes y las facilidades de producción, y los mecanismos de control, pozos nuevos y los límites económicos.

2.8.2 PRiSMa

PRiSMa² es un proyecto de modelaje estocástico, simulación, evaluación de incertidumbre y optimización de yacimientos, está codificado en MATLAB y ha sido desarrollado con el objetivo de implementar técnicas geoestadísticas para modelaje de yacimientos, evaluación de incertidumbre y optimización de localización, tipo y

²<http://villaj.googlepages.com/prisma>

trayectoria de pozos no convencionales.

El flujo de trabajo de PRiSMa consiste en:

- Generación de un modelo estocástico de yacimientos de alta resolución, utilizando técnicas geoestadísticas (fluvsim, sgsim), esto a través de la aplicación SReM.
- Postprocesamiento de las realizaciones del modelo estocástico (postsim).
- Escalamiento de porosidad y permeabilidad de la malla de alta resolución a la malla de simulación, utilizando técnicas de escalamiento estáticas y dinámicas (upscaler, flowsim).
- Determinación de las celdas que conecta un pozo con la malla de simulación y cálculo del well index - WI (wi2ecl.m). Cada pozo puede tener una configuración arbitraria, puede ser vertical, desviado, o multilateral sin limitación con el número de brazos.
- Simulación de diferentes escenarios de explotación con diferentes configuraciones de pozos utilizando ECLIPSE.
- Simulación integrada subsuelo-superficie de una red de superficie de recolección costa afuera usando la opción NETWORKS.
- Postprocesamiento de los resultados de las simulaciones de yacimientos.

SReM

Genera una malla de alta resolución con múltiples realizaciones de sistemas fluviales utilizando técnicas geoestadísticas y realiza el escalamiento de porosidad (analítico-aritmetico) y permeabilidad (numérico-monofásico) utilizando diferentes relaciones de escalamiento. Inicializa cada modelo en ECLIPSE.

PRiSMa

Corre o carga un modelo generado por SReM, calcula las celdas que cada pozo intercepta en cada realización de cada modelo escalado generado por SReM. Cada pozo puede ser vertical, desviado, horizontal o multilateral. Puede manejar múltiples escenarios de explotación (diferentes escenarios con diferentes número y configuración de pozos). Realiza las simulaciones en ECLIPSE. Puede hacer una simulación convencional o una simulación integrada subsuelo-superficie. Genera gráficos de producción total y por pozos.

bopvt2ecl

PRiSMa posee un módulo que permite realizar el cálculo de las propiedades Presión-Volumen-Temperatura, para tal fin, en éste se encuentran implementadas las ecuaciones mostradas en el Apéndice A. A su vez, se generan los archivos necesarios de extensión .INC (INCLUDE) utilizados por Eclipse para realizar las simulaciones.

relperms2ecl

Se encuentra dentro de PRiSMa y su función es determinar las curvas de permeabilidades relativas para los fluidos existentes, los cálculos se realizan a través de las ecuaciones que se encuentran en el Apéndice D. Este módulo también escribe los archivos .INC (INCLUDE) necesarios para realizar las corridas de Eclipse.

2.8.3 Essencial Regression and Experimental Design for Chemists and Engineers

Essential Regression and Experimental Design for Chemists and Engineers fue desarrollado como un software que permite a los no estadísticos realizar y analizar diseños de experimentos y datos cuantitativos usando regresión lineal múltiple y polinomial de una manera sencilla y entendible. Este software se encuentra disponible como una aplicación de MS Excel, es decir, como un macro compilado de Excel, lo que facilita al usuario el manejo de los datos, ya que este no necesita tener conocimientos estadísticos amplios para poder usar el paquete, además reduce el tiempo de trabajo al trivializar todos aquellos pasos dentro de la investigación que consumen tiempo.

El uso de esta herramienta permitió obtener las matrices de diseño, para cada uno de los diseños experimentales utilizados, además de facilitar el postprocesamiento de la información, con técnicas estadísticas como el Análisis de Varianza, Prueba de Significancia de las variables, entre otras.

Essential Regression and Experimental Design for Chemists and Engineers ofrece al usuario, a través de su ventana de Diseño Experimental (Ver Figura 2.8), una variedad de características y diseños que seleccionar de acuerdo con los objetivos del estudio a realizar, como lo son el número de factores a utilizar en el diseño, el número de puntos centrales, en caso que se desee incluirlos, el número de respuestas que se utilizaran como función objetivo y cual diseño utilizar.

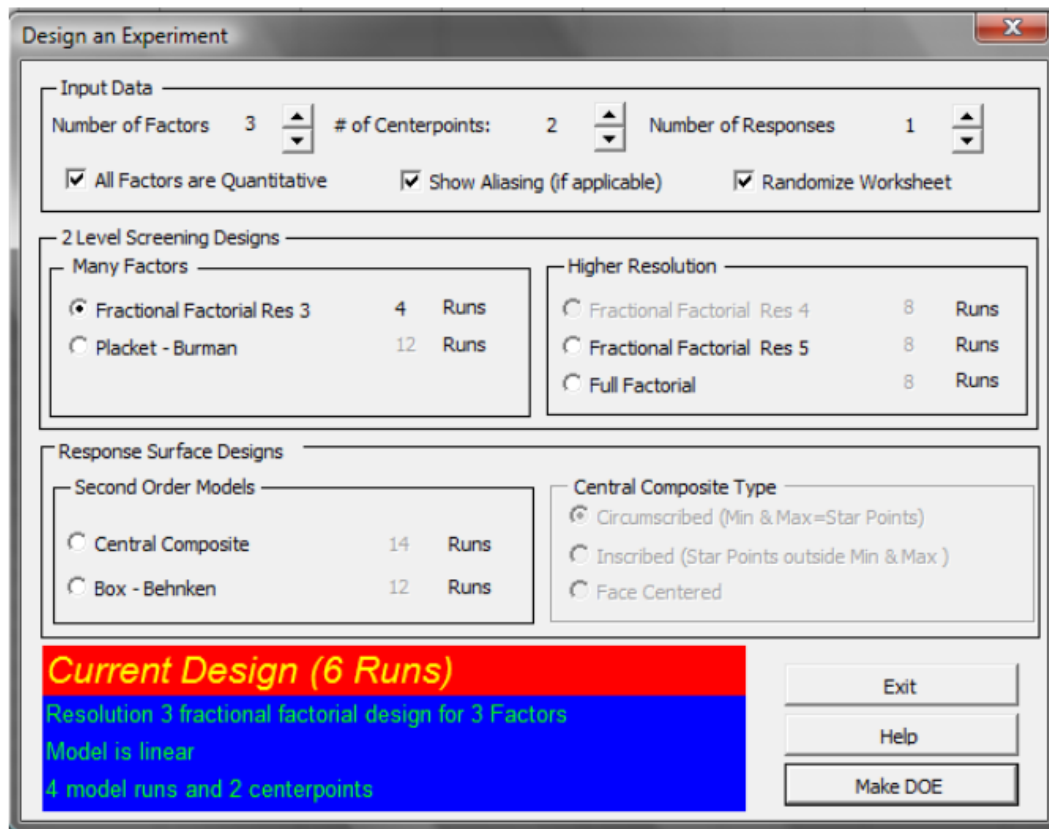


Figura 2.8: Ventana de Diseño de Experimentos

Los diseños se encuentran clasificados en dos grupos:

- Diseños para Análisis de Sensibilidad: diseños de 2 niveles para muchos factores y de alta resolución, como lo son Factorial Completo, Factorial Fraccionado de resolución III, IV y V y Plackett Burman.
- Diseños para ajustar Superficie de Respuesta: Conjunto de Diseños Centrales, compuesto, inscrito e inscrito en las caras y Box Behnken.

Luego de seleccionar todo lo concerniente al diseño, existen opciones para especificar el nombre de las variables de entrada y de las variables para medir la respuesta, las unidades de cada una de ellas y los niveles, alto, medio o bajo, según el diseño.

Para el análisis estadístico, el paquete ofrece regresión lineal múltiple, a través del uso de técnicas de mínimos cuadrados, y para análisis de las variables dentro del estudio y del modelo ajustado se utiliza ANOVA y otros procedimientos descritos en el capítulo 1, todo esto a través de la ventana de regresión lineal (Ver Figura 2.9), en la cual el usuario puede seleccionar cuales factores se deben incluir en la

regresión y se muestran los diferentes elementos que forman el resultado del diseño de experimentos, como lo son la ecuación del modelo, la tabla ANOVA para las variables, y otras pruebas estadísticas. Se puede generar, en caso de así desearlo, un archivo de salida que contendrá toda la información, tanto de entrada como de salida, donde se puede observar los resultados de los cálculos realizados y realizar gráficos de superficies de respuesta.

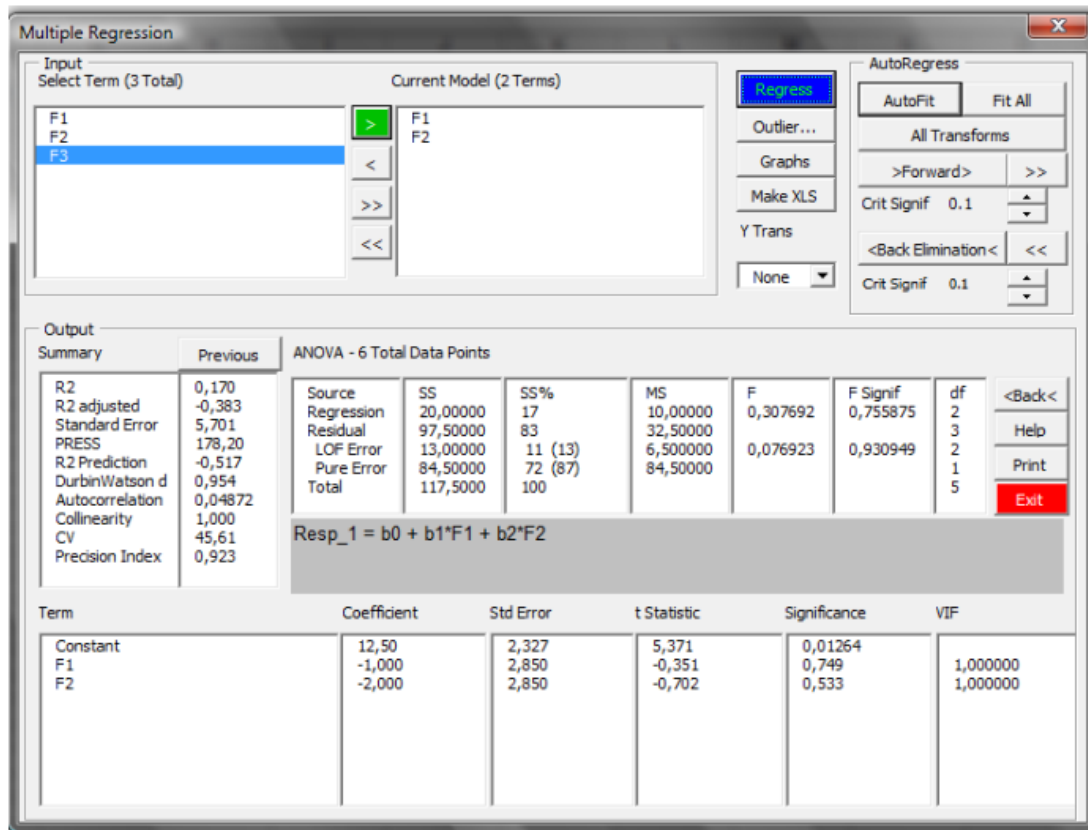


Figura 2.9: Ventana de Regresión Lineal Múltiple

Capítulo 3

Análisis de Resultados

3.1 Generación de los modelos de yacimiento

Los modelos de yacimiento se generaron con técnicas geoestadísticas. Por cada matriz de diseño para análisis de sensibilidad se produjeron 100 modelos de yacimiento, de los cuales se seleccionaron 20, los que tuviesen una porosidad promedio alta o baja, en función de lo que la matriz estableciese. Para el análisis de "incertidumbre" se realizaron 220 modelos de yacimiento y se seleccionaron 44.

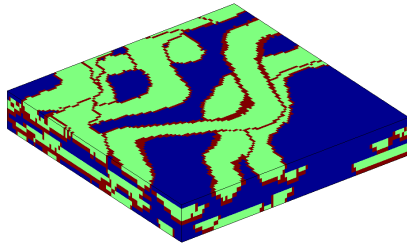
Para generar los modelos de yacimiento se hizo una realización de facies por cinco propiedades. La figura 3.1 muestra los resultados obtenidos para una de las simulaciones.

En la figura 3.2 se muestra la distribución de porosidad y permeabilidad sin condicionamiento a las facies presentes en el modelo, en este caso para la realización 1.

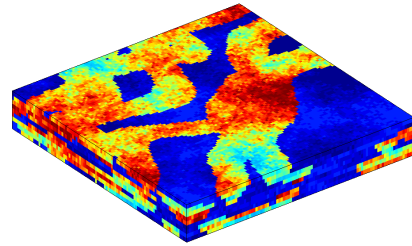
En la figura 3.3 se muestra la distribución de porosidad y permeabilidad condicionada a las facies presentes en el modelo, en este caso para la realización 1. El proceso de condicionamiento es de vital importancia porque permite que el modelo simulado reproduzca con la mayor exactitud el posible el comportamiento del yacimiento real.

3.2 Análisis de sensibilidad

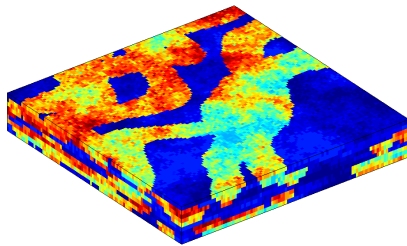
Para realizar el Análisis de Sensibilidad se crearon 15 matrices de diseño Plackett Burman, y se hicieron las corridas de simulación de acuerdo con la disposición



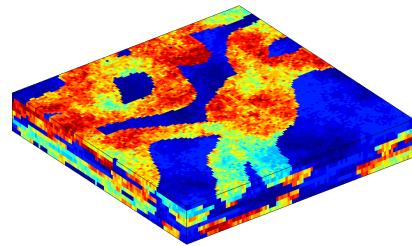
(a) Modelo de facies



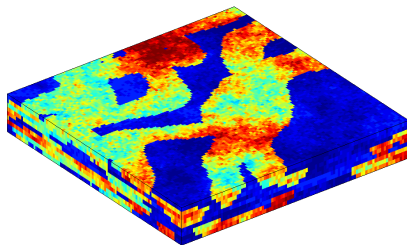
(b) Realización 1



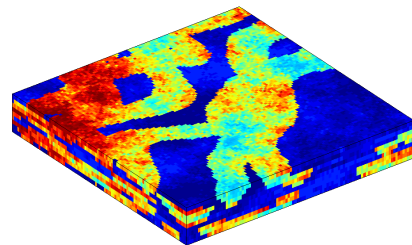
(c) Realización 2



(d) Realización 3



(e) Realización 4



(f) Realización 5

Figura 3.1: Escenario 1, una realización de facies por cinco realizaciones de propiedades

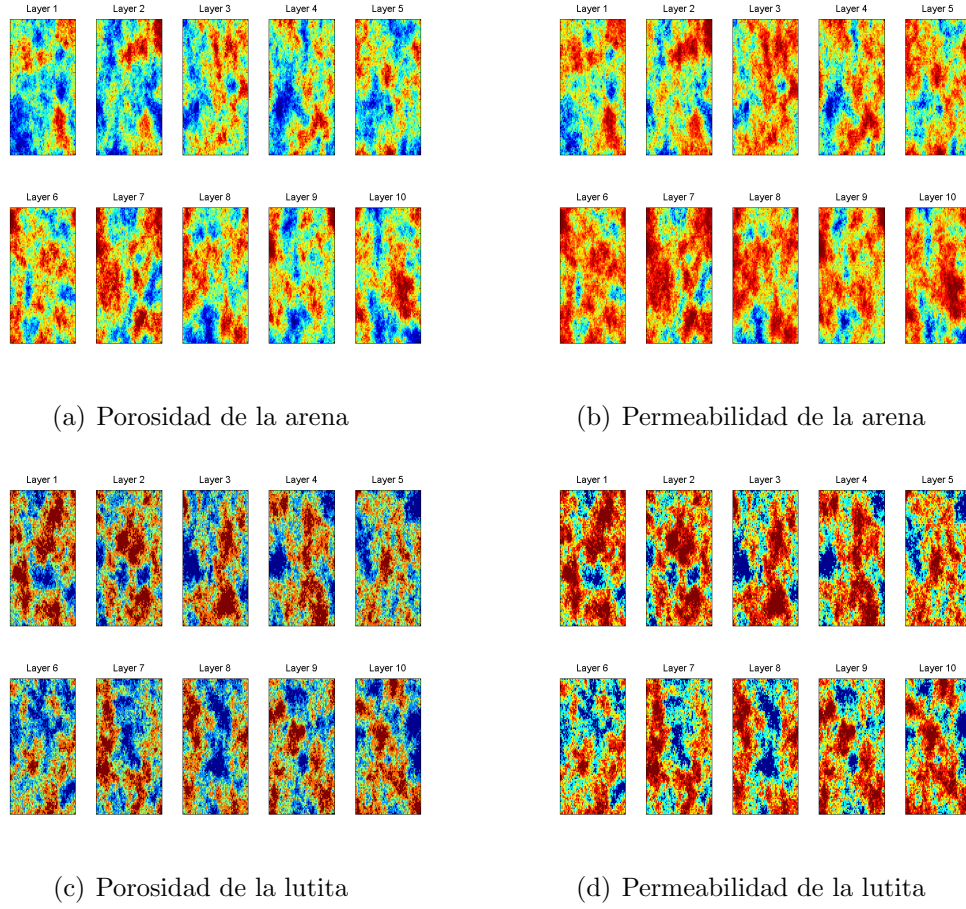


Figura 3.2: Sección horizontal de los modelos de porosidad y permeabilidad sin condicionar. Realización 1

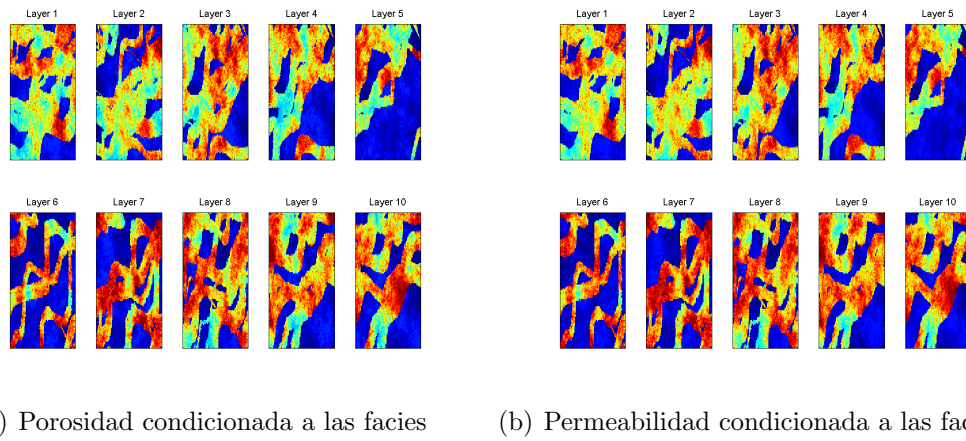


Figura 3.3: Sección horizontal de los modelos de porosidad y permeabilidad condicionados a las facies. Realización 1

de parámetros que la matriz indicara. Para cada parámetro, un valor igual a 1 se traducía en tomar el valor alto para realizar la simulación; un valor igual a -1 hacía referencia al valor bajo de dicho parámetro (los valores altos y bajos de los parámetros se pueden observar en el Capítulo 2, en la tabla 2.8).

Los resultados de esta etapa para cada una de las matrices se muestran en la tabla 3.1, donde con x se indica cuales parámetros registraron una influencia notoria en la variable de respuesta.

Factor \ Matriz	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
T															
GOR															
Gas Grav															
API				x	x			x					x		
poro	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
multperm															
kvkh	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
aquifer															
nwat															x
noilwat	x	x	x	x	x	x	x	x	x	x	x	x		x	x
pcow															
noilgas	x	x	x		x	x			x	x	x	x	x	x	x
ngas	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
pcog															

Tabla 3.1: Parámetros más significativos

La selección de los parámetros señalados en la tabla se realizó de acuerdo con la Prueba de la Hipótesis para los regresores, mencionada en el Capítulo 1, ésto se hizo con el uso de la función de distribución t-Student para contrastar las hipótesis nula y alternativa. La hipótesis nula utilizada establece que los regresores no son significativos, mientras que la hipótesis alternativa enuncia que si son significativos dentro del estudio.

El nivel de significancia considerado para contrastar las hipótesis fue $\alpha = 5\%$, todos los términos cuyo *valor p* fuese menor que el nivel de α establecido se consideraron significativos. En la mayoría de los casos se obtuvieron más de 5 parámetros, pero para todos ellos se seleccionaron sólo 5, los de menor valor p o mayor significancia. Esto se debe a la necesidad de mantener un número fijo de regresores para

poder establecer las comparaciones pertinentes y tener resultados consistentes. Sin embargo, es importante mencionar que, aunque el *valor p* varía de acuerdo con el número de regresores presentes, se verificó que todos los parámetros seleccionados fuesen los de mayor significancia partiendo de una ecuación que involucrara los 14 términos sometidos a este estudio.

Para ilustrar de mejor manera lo anteriormente expuesto se muestra en la tabla 3.2 la Matriz de Diseño 1 y en la tabla 3.3 los resultados de las corridas de simulación hechas con dicha matriz. Estos resultados corresponden a la producción total acumulada (N_p) para el tiempo de estudio de 10 años y a través de los cuales se obtuvo la siguiente ecuación matemática 3.1, cuyos coeficientes se encuentran en la tabla 3.4:

$$\begin{aligned} N_p = & b0 + b1 * \text{poro} + b2 * \text{multperm} + b3 * \text{kvh} + b4 * \text{aquifer} + \\ & + b5 * T + b6 * GOR + b7 * Gas + b8 * API + b9 * nw + \\ & + b10 * now + b11 * pcow + b12 * nog + b13 * ng + b14 * pcog \end{aligned} \quad (3.1)$$

Luego al seleccionar los 5 parámetros de mayor importancia (que son los resaltados en negrillas en la tabla 3.4) se obtuvo la siguiente ecuación:

$$\begin{aligned} N_p = & b0 + b1 * ng + b2 * kvh + \\ & + b3 * nog + b4 * now + b5 * \text{poro} \end{aligned} \quad (3.2)$$

Los coeficientes de dicha ecuación se indican en la tabla 3.5.

Este procedimiento se realizó para las 15 matrices de diseño, y se determinó que para la siguiente fase del flujo de trabajo, es decir, el análisis de riesgo, los parámetros seleccionados serían los que se indican a continuación:

3.2.1 Discusión de los parámetros de mayor significancia

Exponentes Corey

Durante años las permeabilidades relativas han sido objeto de amplio estudio, a nivel de pruebas de laboratorio con la finalidad de entender como ésta influye en la producción. Sin embargo, el uso de métodos empíricos ha cobrado fuerza en los últimos tiempos, sobretodo para el cálculo de esta propiedad dentro del flujo de trabajo de los simuladores. En el presente estudio se utilizaron los modelos

Exp	ϕ	mult perm	kv/ kh	aquifer	T	GOR	Gas Grav	API	nw	now	pcow	nog	ng	pcog
1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1
2	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1
3	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1
4	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
5	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1
6	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1
7	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1
8	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1
9	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1
10	-1	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1
11	1	1	1	-1	-1	1	1	-1	1	1	-1	-1	-1	-1
12	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1
13	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1	1
14	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	-1
15	-1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	1	-1
16	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1	1
17	1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	1
18	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1
19	-1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1
20	-1	1	1	1	1	-1	-1	1	1	-1	1	1	-1	-1

Tabla 3.2: Matriz de Diseño 1

Exp	N_p
1	21565670
2	27577910
3	22555470
4	19526070
5	28933950
6	31301470
7	19923290
8	17261370
9	31412620
10	25132520
11	19984520
12	33171890
13	17012280
14	13459280
15	32189410
16	27278400
17	23209140
18	27100410
19	28799840
20	18461290

Tabla 3.3: Resultados de simulación Matriz 1

		P value	Std Error	t Stat
b0	2,429e07	1,562e-10	147434	164,77
b1	820217	0,00258	147434	5,563
b2	-240834	0,163	147434	-1,634
b3	1,589e06	0,000119	147434	10,78
b4	-149898	0,356	147434	-1,017
b5	-85422,0	0,587	147434	-0,579
b6	-195485	0,242	147434	-1,326
b7	190464	0,253	147434	1,292
b8	197993	0,237	147434	1,343
b9	-797000	0,00293	147434	-5,406
b10	-1,060e06	0,000811	147434	-7,188
b11	44312,0	0,776	147434	0,301
b12	-1,143e06	0,000572	147434	-7,751
b13	4,997e06	4,204e-07	147434	33,89
b14	552890	0,01329	147434	3,750

Tabla 3.4: Resultados Matriz 1. Modelo de 14 parámetros

		P value	Std Error	t Stat
b0	2,429e07	4,018e-20	299166	81,20
b1	4,997e06	1,217e-10	299166	16,70
b2	1,589e06	0,000110	299166	5,311
b3	-1,143e06	0,00188	299166	-3,820
b4	-1,060e06	0,00325	299166	-3,542
b5	820217	0,01590	299166	2,742

Tabla 3.5: Modelo ajustado para 5 parámetros. Matriz 1

empíricos planteados por Brooks y Corey, (ver apéndice D). Dentro de dichos modelos se consideraron como parámetros potenciales de diseño los exponentes, mejor conocidos como Exponentes Corey, los cuales determinan el comportamiento de las curvas de permeabilidades relativas, y su obtención se realizó a partir de pruebas de laboratorio realizadas a núcleos, correspondientes a diferentes tipos de roca yacimiento.

La permeabilidad es una propiedad de la roca, y se define como una medida de la capacidad que tiene una roca para permitir el paso de fluidos a través de los poros interconectados de la misma. Por otra parte, la permeabilidad relativa se produce cuando existe más de un fluido en el medio poroso. Se puede afirmar, desde el punto de vista de ingeniería de yacimientos, es uno de los elementos que tienen mayor impacto en el factor de recobro. Es decir, las curvas de permeabilidades relativas, para los modelos de yacimiento desarrollados, fueron los elementos con mayor incertidumbre asociada.

Dentro del sistema roca-fluido se definieron 4 exponentes: dos para el sistema crudo-agua, uno para el agua (nw) y uno para el petróleo (noilwat); dos para el sistema crudo-gas, uno para el gas (ng) y otro para el petróleo (noilgas). De los 4 exponentes mencionados 3 tienen un impacto considerable en el comportamiento del factor de recobro, estos son: ngas, noilgas y noilwat, es decir, exponente Corey para el gas, para el petróleo en presencia de gas y para el petróleo en presencia de agua, respectivamente. Estos se consideraron como parámetros influyentes dentro del análisis de sensibilidad de acuerdo con los resultados estadísticos, sin embargo, éstos se pueden justificar, desde el punto de vista de ingeniería de yacimientos, de acuerdo con los siguientes aspectos:

- Tienen incertidumbre asociada a la variabilidad, es decir, los exponentes Corey utilizados, se seleccionaron de acuerdo con la literatura consultada y con las

características de los modelos de yacimiento generados. Se asumió que eran representativos de la roca yacimiento, pero es necesario tener en cuenta que para un estudio de yacimiento, por lo general, se cuenta con más de un conjunto de curvas de permeabilidades relativas; para este estudio sólo se utilizó uno por modelo de yacimiento generado.

- Tienen incertidumbre asociada de acuerdo con el grado de precisión con el que se realizaron las pruebas de laboratorio a través de las cuales se obtuvieron los valores seleccionados, por lo general, se toman valores documentados en tablas de experimentos previos.

Relación k_v/k_h

También conocido como el coeficiente de anisotropía, tuvo un efecto positivo dentro del estudio, es decir, al aumentar dicho factor, se registró un aumento en el recobro obtenido, una disminución del mismo, ocasionaba una disminución en el recobro obtenido.

Porosidad

Tiene una gran incertidumbre asociada, por lo que clasifica como un parámetro que influye en la producción de petróleo acumulada, tuvo un efecto positivo, debido a que los escenarios cuyo promedio de porosidad fuese alto registraron un factor de recobro alto, y para los que tenían porosidad promedio baja hubo un factor de recobro bajo.

Otros parámetros

De los parámetros PVT seleccionados para este análisis, de sensibilidad, la gravedad API, la gravedad específica del gas y la temperatura formaron parte de los parámetros de entrada que registraron una significancia importante. Sin embargo, no se tomaron en cuenta para la fase siguiente por su poca ocurrencia, como fue el caso de la Gravedad API; adicionalmente sus respectivos valores p estaban muy próximos al nivel de significancia del estudio. Otro parámetro que se descartó por presentar el mismo comportamiento, pero perteneciente a las propiedades roca-fluido evaluadas, fue el Exponente Corey para el agua (n_{wat}), sólo se presentó como parámetro significativo en una de las 15 matrices analizadas en esta etapa.

3.3 Análisis de Incertidumbre

Para establecer la superficie de respuesta se generaron 5 matrices de diseño experimental de tipo Box Behnken con los factores considerados como de diseño, es decir, los elementos de mayor significancia dentro del estudio, y se ajustaron 5 ecuaciones de segundo orden, una por cada matriz. Al igual que para el análisis de sensibilidad se hicieron simulaciones, las cuales se realizaron de acuerdo con la disposición de parámetros que las matrices indicaran.

Los parámetros que se consideraron constantes fueron todos los concernientes a las propiedades PVT, el acuífero, la permeabilidad, las presiones capilares y el Exponente Corey para el agua. Todos ellos se utilizaron en los cálculos realizados y simulaciones con su nivel medio (Ver Tabla 2.8).

Para el análisis estadístico de los resultados de esta etapa se utilizó la opción AUTOFIT de la herramienta *"Essencial Regression and Experimental Design for Chemists and Engineers"*, de donde se obtuvo una ecuación por matriz; sin importar la cantidad de coeficientes que se ajustaran por ecuación. Las ecuaciones o superficies de respuesta ajustadas a partir de cada matriz se presentan en la sección Superficies de Respuesta.

La metodología de Superficie de Respuesta tiene como principal ventaja el tiempo de computo, es decir, permite obtener la misma cantidad de resultados que se generan con las simulaciones, pero en menor tiempo y con el uso de pocos recursos computacionales. Sin embargo, su poder predictivo depende de la cantidad de simulaciones utilizadas para generarla, por lo que para su creación se debe tener en cuenta la relación tiempo/recursos, para lograr un balance adecuado entre el tiempo disponible para ajustarla y los recursos que se tengan para tal fin.

En el presente estudio, el número de parámetros utilizados fue quien indicó el tipo de diseño a seleccionar, y el número de matrices que se generarían para obtener las superficies de respuesta. Por la cantidad de parámetros las matrices de diseño fueron de 5 columnas por 44 filas, lo que implicó un tiempo de computo considerable, esto a su vez se tradujo en una reducción de la cantidad de matrices que se utilizarían, y en consecuencia se disminuyó el poder predictivo de las superficies ajustadas.

3.3.1 Superficies de Respuesta

Teniendo en cuenta las ventajas y limitaciones anteriormente expuestas se crearon 5 superficies de respuesta. El tipo de diseño Box Behnken permitió ajustar modelos de segundo orden, en los cuales se tomó en cuenta la interacción entre los efectos de las variables, elementos utilizados para ajustar este tipo de ecuaciones.

Matriz de diseño 1

Modelo ajustado para la matriz de diseño 1

$$N_p = b0 + b1 * ngas + b2 * ngas * ngas + b3 * kvkh + \\ + b4 * noilgas + b5 * poro + b6 * kvkh * kvkh + b7 * noilwat \quad (3.3)$$

Cuyos coeficientes se encuentran en la tabla 3.6

	Coeficientes	P value	Std Error
b0	2,778e07	1,634e-53	167490
b1	5,931e06	1,622e-28	179613
b2	-3,253e06	2,429e-16	228834
b3	1,428e06	1,943e-09	179613
b4	-1,160e06	1,702e-07	179613
b5	903182	1,377e-05	179613
b6	-1,048e06	5,399e-05	228834
b7	-431588	0,02155	179613

Tabla 3.6: Coeficientes de la ecuación para la Superficie de Respuesta 1

Al aplicar simulación Monte Carlo a esta ecuación se obtuvo como resultado el histograma mostrado en la siguiente figura 3.4.

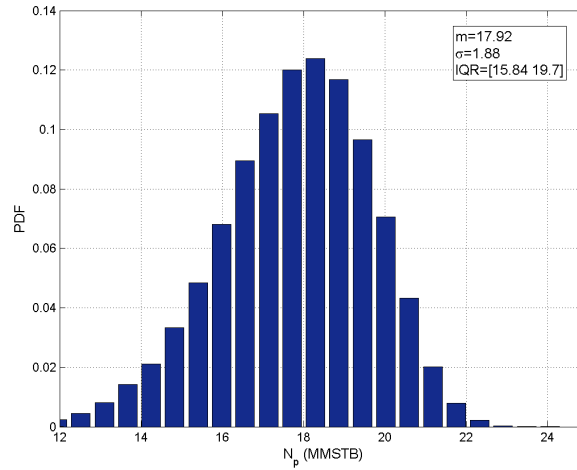


Figura 3.4: Histograma Superficie de Respuesta 1

Matriz 2

Para la Matriz 2 la superficie de respuesta es la siguiente:

$$N_p = b_0 + b_1 * ngas + b_2 * ngas * ngas + b_3 * kvkh + b_4 * noilgas + b_5 * noilwat + b_6 * poro + b_7 * kvkh * kvkh + b_8 * ngas * kvkh \quad (3.4)$$

Donde los coeficientes de dicha ecuación se muestran en la siguiente tabla:

	Coeficientes	P value	Std Error
b0	2,750e07	5,569e-50	192936
b1	5,207e06	5,265e-24	206901
b2	-3,034e06	1,894e-13	263601
b3	1,472e06	2,710e-08	206901
b4	-1,442e06	4,164e-08	206901
b5	-863203	0,000189	206901
b6	844003	0,000248	206901
b7	-854792	0,00260	263601
b8	972620	0,02452	413802

Tabla 3.7: Coeficientes de la ecuación para la Superficie de Respuesta 2

Los resultados de la simulación Monte Carlo se muestran en el siguiente histograma

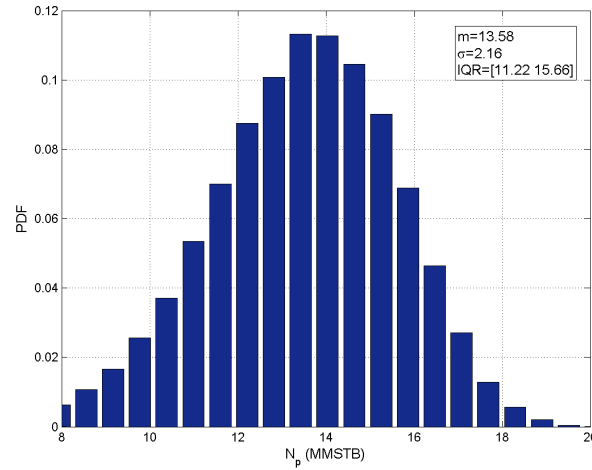


Figura 3.5: Histograma Superficie de Respuesta 2

Matriz 3

Modelo ajustado o superficie de respuesta para la Matriz 3:

$$N_p = b_0 + b_1 * ngas + b_2 * ngas * ngas + b_3 * kvkh + b_4 * noilgas + b_5 * kvkh * kvkh + b_6 * poro + b_7 * noilwat + b_8 * ngas * kvkh \quad (3.5)$$

Coefficientes de la ecuación:

	Coefficientes	P value	Std Error
b0	2,786e07	2,594e-52	167621
b1	5,638e06	3,263e-27	179754
b2	-3,218e06	5,917e-16	229014
b3	1,651e06	7,505e-11	179754
b4	-1,382e06	5,076e-09	179754
b5	-1,458e06	2,537e-07	229014
b6	1,052e06	1,217e-06	179754
b7	-1,007e06	2,577e-06	179754
b8	737550	0,04775	359508

Tabla 3.8: Coeficientes de la ecuación para la Superficie de Respuesta 3

Los resultados de la simulación Monte Carlo se presentan en el siguiente histograma:

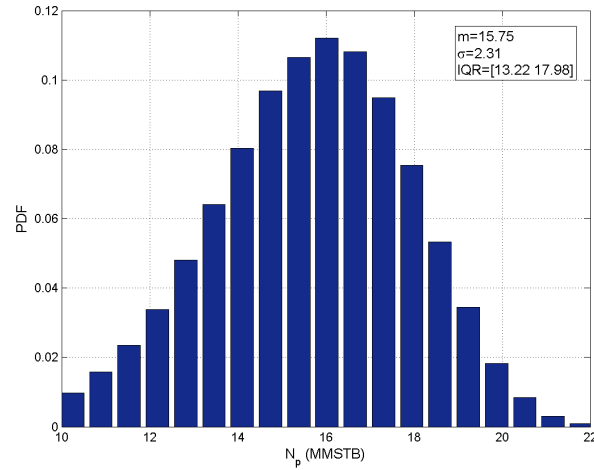


Figura 3.6: Histograma Superficie de Respuesta 3

Matriz 4

Para la Matriz 4 la superficie de respuesta tiene la siguiente ecuación matemática:

$$\begin{aligned}
 N_p = & b_0 + b_1 * ngas + b_2 * ngas * ngas + b_3 * kvkh + \\
 & + b_4 * poro + b_5 * noilgas + b_6 * noilwat + +b_7 * kvkh * kvkh + \\
 & + b_8 * noilwat * noilwat + b_9 * ngas * kvkh
 \end{aligned} \tag{3.6}$$

Coefficientes de la ecuación:

	Coefficientes	P value	Std Error
b0	2,891e07	1,653e-45	253663
b1	5,306e06	5,498e-23	219678
b2	-3,187e06	7,278e-13	286737
b3	1,497e06	7,677e-08	219678
b4	982817	8,171e-05	219678
b5	-832955	0,000586	219678
b6	-808158	0,000804	219678
b7	-921275	0,00287	286737
b8	-611615	0,04023	286737
b9	805678	0,07545	439356

Tabla 3.9: Coeficientes de la ecuación para la Superficie de Respuesta 4

Resultados de simulación Monte Carlo en el siguiente histograma:

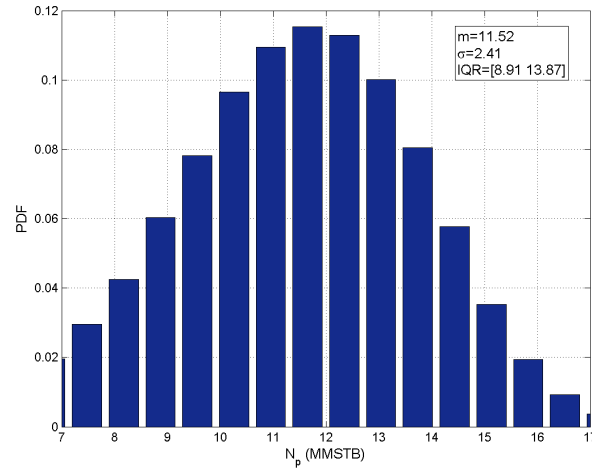


Figura 3.7: Histograma Superficie de Respuesta 4

Matriz 5

Superficie de Respuesta obtenido a partir de la Matriz 5:

$$\begin{aligned}
 N_p = & b0 + b1 * ngas + b2 * ngas * ngas + b3 * kvkh + \\
 & + b4 * poro + b5 * kvkh * kvkh + b6 * noilgas + b7 * noilwat + \\
 & + b8 * ngas * poro + b9 * poro * kvkh
 \end{aligned} \tag{3.7}$$

Donde los coeficientes de la ecuación se muestran en la siguiente tabla:

	Coeficientes	P value	Std Error
b0	2,786e07	4,693e-51	167796
b1	5,726e06	6,904e-27	179941
b2	-3,287e06	5,609e-16	229252
b3	1,656e06	9,301e-11	179941
b4	1,244e06	5,760e-08	179941
b5	-1,412e06	5,364e-07	229252
b6	-979464	4,561e-06	179941
b7	-819951	6,402e-05	179941
b8	-1,197e06	0,00211	359881
b9	790588	0,03495	359881

Tabla 3.10: Coeficientes de la ecuación para la Superficie de Respuesta 5

Los resultados para la simulación Monte Carlo se presentan en el siguiente histograma:

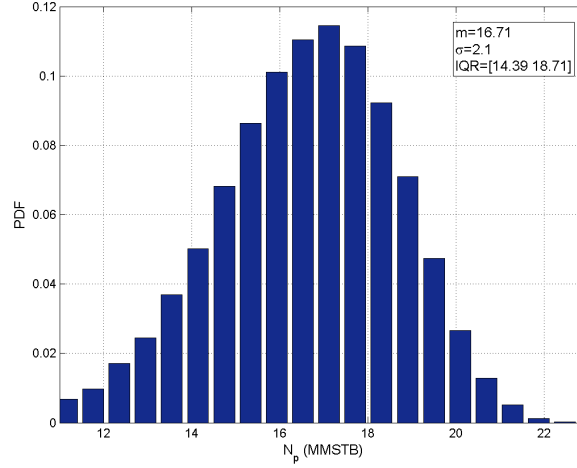


Figura 3.8: Histograma Superficie de Respuesta 5

De acuerdo con los resultados anteriormente mostrados se puede afirmar que para todos los modelos los parámetros de mayor contribución o significancia son: exponente Corey del gas n_g , exponente Corey del gas (término cuadrático) n_g^2 y la relación kv/kh . Por su parte, la porosidad si su cambia su significancia es el parámetro que cambia en en las superficies.

Como primer paso para realizar un análisis comparativo entre los modelos ajustados se hizo un análisis estadístico. Para ello se planteó la hipótesis nula que supone que todos los modelos son equivalentes, y la hipótesis alternativa de que los modelos difieren entre ellos, y por lo tanto no son equivalentes entre si.

Utilizando la *prueba de distribución t para dos muestras de dos colas*, se verificó que los valores t_0 fuesen o mayores que el t_{crit} positivo o menores que $-t_{crit}$. Con los resultados obtenidos (ver tabla 3.11) se determinó que los modelos no eran equivalentes entre ellos, por lo que se procedió a realizar otro tipo de pruebas que permitieran seleccionar una superficie de respuesta.

3.3.2 Análisis de Incertidumbre y Simulación Monte Carlo de la Superficie de Respuesta seleccionada

Dentro del análisis de sensibilidad se realizó Simulación Monte Carlo, con un número de corridas igual a diez mil (10000), para obtener la información necesaria para determinar si las superficies son equivalentes entre ellas, o si por el contrario, es necesario escoger una.

Histogramas	t_0	t_{crit}
1 2	14,33	1,99
1 3	14,13	1,99
1 4	6,97	1,99
1 5	4,02	1,99
2 3	-6,83	-1,99
2 4	6,39	1,99
2 5	-10,12	-1,99
3 4	12,96	1,99
3 5	-3,04	-1,99
4 5	-16,51	-1,99

Tabla 3.11: Resultados Prueba t para dos muestras

Para el análisis de incertidumbre se seleccionó una superficie de respuesta, realizando comparaciones de indicadores estadísticos, tales como: el nivel de significancia, el poder predictivo y el error asociado a cada uno de los modelos. El error se estimó probando cada una de las superficies de respuesta con valores aleatorios y comparando sus resultados con valores de Np obtenidos de simulaciones realizadas con los mismos valores aleatorios.

La superficie de respuesta 1 (RSM1) fue el modelo que reunió todas las condiciones deseadas (ver tabla 3.12), como lo son el Índice de precisión y R^2_{pred} .

	RSM1	RSM2	RSM3	RSM4	RSM5
R2 for Prediction	0,963	0,918	0,960	0,929	0,961
Precision Index	60,008	42,222	63,291	66,685	39,688
LOF Error	1,748e13	2,248e13	1,727e13	2,570e13	1,560e13
Pure Error	1,101e12	1,488e12	8,270e11	5,495e11	2,013e12
F	204,69	113,26	176,23	89,50	158,92
F Signif	5,002e-27	1,064e-22	6,096e-26	4,956e-21	4,108e-25
Mayor significancia	1	4	2	5	3
Mayor prediccion	1	5	3	4	2

Tabla 3.12: Comparación de modelos

Sin embargo, aún cuando la superficie de respuesta 1 fue el modelo con menor error asociado es necesario destacar que el error puro (Pure Error) y el Error de ajuste (Lack of Fit, LOF Error), de todos los modelos, incluyendo el seleccionado, presentan valores de un orden de magnitud elevado, este comportamiento se puede atribuir a:

- Falta de regresores, la inclusión de más parámetros podría disminuir el error de ajuste.
- Falta de más simulaciones, el número de simulaciones utilizados para ajustar las superficies de respuesta pudo haber sido insuficiente, por lo que se ajustaron modelos que aunque tienen alta capacidad predictiva, tienen un elevado error asociado.
- El diseño utilizado, es, dentro del conjunto de diseños para ajustar superficies de respuesta, el que posee menor poder predictivo y baja resolución, sin embargo, es el único que permite manejar una cantidad de parámetros superior a 5.
- Modelo Seleccionado, es decir, para los cálculos realizados se utilizó un modelo de segundo orden, ésto partiendo de la hipótesis que la superficie ajustada tiene ese comportamiento, sin embargo, el error obtenido muestra que el modelo no es adecuado por lo que se sugiere utilizar un modelo de tercer orden, e inclusive de orden superior, para obtener un mejor ajuste.

Posteriormente a la selección de la superficie se realizó simulación Monte Carlo para ésta, con un millón de corridas y se obtuvo como resultado que la variable de respuesta, N_p tiene como valor promedio $17,94MMBN$, varianza o medida de dispersión igual a $\sigma^2 = 1.71$ y una desviación estándar $\sqrt{\sigma^2} = 1.30MMBN$. El Histograma se presenta en la siguiente figura:

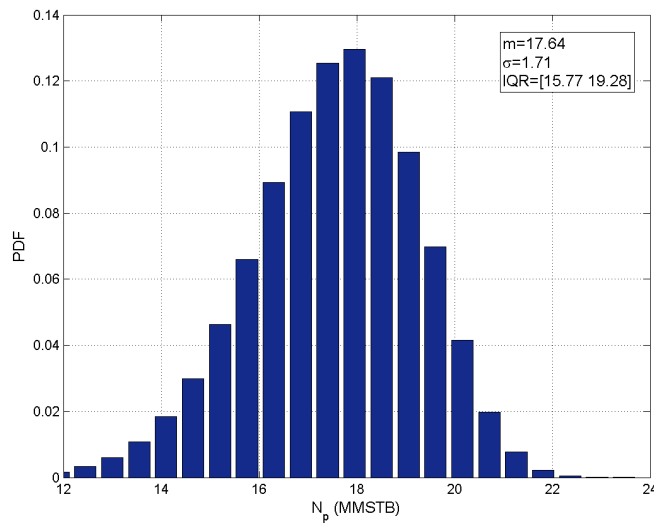


Figura 3.9: Histograma. Superficie de Respuesta seleccionada

Como se puede observar, al incrementar el número de simulaciones Monte Carlo (n), los resultados obtenidos no varían de manera considerable con respecto al primer grupo de simulaciones realizadas (ver tabla 3.13), lo que indica consistencia en la metodología aplicada.

	n=10000	n=1000000
Promedio MMBN	17.92	17.94
Varianza	1.88	1.71
Desviación Estándar MMBN	1.37	1.30

Tabla 3.13: Simulaciones Monte Carlo para distintos n

De acuerdo con los resultados de la tabla 3.13 se puede observar que la producción de petróleo acumulada N_p tiene un valor promedio, para $n = 10000$ de 17.92 MMBN y para $n = 1000000$ de 17.94 , y una varianza de 1.88 y 1.71 , respectivamente, lo que indica que para un incremento del número de simulaciones Monte Carlo, mejor ajuste se tiene. Es decir, el modelo ajustado con un mayor número de simulaciones Monte Carlo presenta menor dispersión, por lo tanto se puede decir que es el que controla de manera más eficiente la incertidumbre asociada a las variables.

Conclusiones

- La incertidumbre es un parámetro de gran importancia que se debe considerar en el estudio de yacimientos, debido a que posee un gran impacto en la elaboración de pronósticos y en la toma de decisiones.
- El Diseño de Experimentos es una técnica de gran utilidad, debido a que permite estudiar procesos que incluyen gran cantidad de variables, reduciendo los costos que esto implica.
- La integración de diversos diseños experimentales con otras técnicas estadísticas, como Análisis Anova y Simulación Monte Carlo, se considera como una herramienta robusta y versátil.
- Para los modelos de yacimiento analizados se verificó la existencia de una relación lineal significativa entre la variable de respuesta, es decir, el factor de recobro y los regresores: Exponentes Corey del gas, petróleo-gas, petróleo-agua, la relación k_v/k_h y la porosidad.
- El gas como fase presente en el yacimiento, es un parámetro de importancia (los exponentes Corey asociados a un sistema con gas libre fueron 2 de los 5 parámetros de mayor impacto en el factor de recobro).
- La relación k_v/k_h y la porosidad tienen una influencia positiva sobre el recobro de los modelos estudiados.
- Los parámetros exponente Corey del gas y k_v/k_h fueron los más significativos dentro del análisis de sensibilidad.
- Es necesario tomar en cuenta dentro de un estudio todos los parámetros que sean significativos en el mismo, esto para evitar, al eliminar alguno de ellos que se disminuya el poder predictivo y se incremente el error de ajuste al momento de elaborar una superficie de respuesta con dichos parámetros.

- Aún cuando el modelo Box-Behnken posee ventajas para el ajuste de superficie de respuesta, cuando se maneja un número de parámetros mayor a 7, el poder predictivo de la superficie de respuesta puede verse afectado por el error asociado al modelo, por lo que puede ser conveniente, si se desea aplicar este tipo de diseño, reducir el número de parámetros para todo el estudio, o dividir el número de parámetros para realizar un diseño de experimentos de tipo secuencial.
- No es conveniente utilizar un gran número de parámetros para aplicar la metodología propuesta, esto debido a que se requiere de un amplio conocimiento estadístico del fenómeno estudiado para discretizar los resultados y poder identificar con mayor precisión las fuentes de incertidumbre dentro de los parámetros analizados.
- La metodología que se aplicó, de acuerdo con los objetivos propuestos, resultó ser eficiente para reducir el tiempo de computo con relación a la cantidad de parámetros utilizados. Sin embargo, se debe tener en cuenta, para otros estudios, el objetivo principal y en función de éste determinar que herramientas se combinaran, es decir, que tipo de diseño seleccionar para cada etapa, cuantas réplicas realizar y que tipo de criterios estadísticos utilizar para analizar los resultados.
- La simulación Monte Carlo, como parte del flujo de trabajo propuesto demostró ser una herramienta muy útil para la evaluación de la incertidumbre en la superficies de respuesta.
- Con el objetivo de disminuir el error asociado al modelo de regresión cuando se maneja un número de parámetros igual o mayor a 5 es conveniente realizar experimentación secuencial y ajustar modelos cuyo orden sea superior al de un ajuste cuadrático.
- Es necesario establecer una relación funcional entre el número de corridas a realizar y el tiempo disponible para realizar los computos, para así seleccionar un diseño que tenga un poder predictivo adecuado a los objetivos del estudio y obtener la mayor cantidad de información posible.

Recomendaciones y Trabajos Futuros

- De acuerdo con los resultados obtenidos, se sugiere aplicar esta misma metodología con 6, 7 y 8 parámetros y observar el comportamiento del error de los modelos ajustados con dichas cantidades de parámetros, con el mismo tipo de diseño para ajuste de superficie de respuesta y con otros diseños.
- Comparar los resultados obtenidos en este estudio con otras metodologías, como por ejemplo Kriging Multivariado con Variograma Lineal.
- Diseñar un módulo adicional en el flujo de trabajo que permita analizar y caracterizar la superficie de respuesta de manera automática, para así determinar si existen puntos de inflexión, y la región donde se ubica el punto estacionario, de forma matemática.
- Realizar las modificaciones necesarias para que la metodología desarrollada pueda ser implementada en otros simuladores, tales como: CMG, VIP, ATHOS, BOAST, SYEP, entre otros.
- Evaluar en forma probabilística el comportamiento de los parámetros operacionales y su optimización.
- Diseñar un módulo adicional que acople el módulo de generación de escenarios óptimos con el módulo de análisis de incertidumbre dentro del flujo de trabajo de PRiSMa.

Bibliografía

- [1] T. Ahmed. *Reservoir Engineering Handbook*. Gulf Professional Publishing, 2001.
- [2] B.R. Al-Shama and R. Teigland. History matching of the valhall field using global optimization method and uncertainty assessment. *Paper SPE 100946 presented at the 2006 Society of Petroleum Engineers Annual Technical Conference and Exhibition held in San Antonio, Texas.*, page 11, 2006.
- [3] A. Al-Shammasi. A review of bubblepoint pressure and oil formation volume factor correlations. *Paper SPE 71302, first presented at the 1999 SPE Middle East Oil Show, Bahrain. Paper peer 15 February pproved 2001.*, page 15, 2001.
- [4] F. G. Arias. *El Proyecto de Investigación. Guía para su elaboración*. 1999.
- [5] E. Barberi and M. Essinfeld. *Yacimientos de Hidrocarburos*. 2001.
- [6] H. Beggs and J. Robinson. Estimating the viscosity of crude oil systems. *Journal of Petroleum Technology - SPE 5434*, page 2, 1975.
- [7] P. Behrenbruch and H. Goda. Two-phase relative permeability prediction: A comparision of the modified brooks-corey methodology with a new carman-kozeny based flow formulation. *Paper SPE 101150 presented at the 2006 Society of Petroleum Engineers Asia Pacific Oil & Gas Conference and Exhibition held in Adelaide, Australia.*, page 18, 2006.
- [8] D. Biswas. Assisted history matching for surface coupled gas reservoir simulation. *Paper SPE 101233 presented at the 2006 Society of Petroleum Engineers Asia Pacific Oil & Gas and Exhibition held in Adelaide, Australia.*, page 10, 2006.
- [9] T. Bu and E. Damsleth. Errors and uncertainties in reservoir performance predictions. *Paper SPE 30604 presented at the 1995 Society of Petroleum Engineers Annual Technical Conference and Exhibition held in Dallas.*, 1995.

- [10] A. Bustamante. Evaluación de riesgo agropecuario. In *UCMA-MAG*.
- [11] C.Li, X. Chen, and Z. Du. A new relationship of rock compressibility with porosity. *Paper Society of Petroleum Engineers 88464 presented at the 2004 Asia Pacific Oil and Gas Conference and Exhibition*, page 5, 2004.
- [12] J. Caers. *Petroleum Geostatistics*. Society of Petroleum Engineers, 2005.
- [13] G. Canavos. *Probabilidad y Estadística. Aplicaciones y Métodos*. Mc Graw Hill, 1988.
- [14] F. Chacín. *Diseño y Análisis de Experimentos*. Ediciones del Vicerrectorado Académico Universidad Central de Venezuela, 2000.
- [15] R. Chambers, J. Yarus, and K. Hird. Petroleum geostatistics for nongeostatisticians part 1- part 2. *The Leading Edge*, 2000.
- [16] J. Chewaroungroaj, O. Valera, and L. Lake. An evaluation of procedures to estimate uncertainty in hydrocarbon recovery predictions. *Paper SPE 59499, presented at the 2000 SPE Asia Pacific Conference on Integrated Modelling for Asset Management held in Yokohama, Japan.*, page 11, 2000.
- [17] A.T. Corey. *Mechanics of Immiscible Fluids in Porous Media*. Water Resorces, 1994.
- [18] A. Cullick, D. Johnson, and G. Shi. Improved and more-rapid history matching with nonlinear proxy and global optimization. *Paper SPE 101933 presented at the 2006 Society of Petroleum Engineers Annual Technical Conference and Exhibition held in San Antonio, Texas.*, page 13, 2006.
- [19] J.P. Dejean and G.Blanc. Managing uncertainties on production predictions using integrated statistical methods. *Paper Society of Petroleum Engineers 56696, 1999 Society of Petroleum Engineers Annual Technical Conference and Exhibition held in Houston, Texas.*, page 15, 1999.
- [20] B. Dindoruk and P. Christman. Pvt properties and viscosity correlations for gulf of mexico oils. *Paper SPE 71633 presented at the 2001 Society of Petroleum Engineers Annual Technical Conference and Exhibition held in New Orleans, Louisiana.*, page 14, 2001.
- [21] R. Ewing. *The Mathematics of Reservoir Simulation*. SIAM, 1983.

- [22] F. Friedmann, A.Chawathe, and D. Laure. Assessing uncertainty in channelized reservoir using experimental desing. *Paper SPE 71622 presented at the 2001 Society of Petroleum Engineers Annual Technical Conference and Exhibition held in New Orleans, Louissiana.*, page 15, 2001.
- [23] C. García. Análisis de errores de presión y pvt sobre las estimaciones de balance de materiales. Technical report, Universidad Central de Venezuela, 2005.
- [24] E. García. Optimización del proceso de sagd en un modelo conceptual de yacimientos utilizando técnicas de diseño experimental. Technical report, Universidad Central de Venezuela, 2007.
- [25] E. Glaso. Generalized pressure-volume-temperature correlations. *Journal of Petroleum Technology, SPE paper 8016*, page 10, 1980.
- [26] G. Guo, M. Diaz, F.Paz, J. Smalley, and E. Waninger. Rock typing as an effective tool for permeability and water saturation modeling: A case study in a clastic reservoir in the oriente basin. *Paper Society of Petroleum Engineers 97033 presented at the 2005 Society of Petroleum Engineers Annual Technical Conference and Exhibition*, page 15, 2005.
- [27] H. Hall. Compressibility of reservoir rocks. *Technical Note*, 198:3, 1953.
- [28] M. Honarpour, L. Koederitz, and H. Harvey. *Relative Permeability of Petroleum Reservoirs*. CRC PRESS, 1986.
- [29] N. Hyne. *Dictionary of Petroleum Exploration, Drilling and Production*. SPE, 1991.
- [30] R. Kuehl. *Diseño de Experimentos*. Thomson, 2001.
- [31] A. Law and D. Kelton. *Simulation Modeling and Analysis*. Mc Graw Hill, 1991.
- [32] J. Lee. *Well Testing*. Society of Petroleum Engineers Textbook Series Vol.1, 1982.
- [33] J. Lee, J. Rollins, and J. Spivey. *Pressure Transient Testing*. 2003.
- [34] F.E. Londono, R.A. Archer, and A. Blasingame. Simplified correlations for hydrocarbon gas viscosity and gas density- validation and correlation of behavior using a large scale database. *Paper Society of Petroleum Engineers 75721 presented at the SPE Gas Technology Symposium hedl in Calgary, Alberta, Canada. 30 Abril - 2 May 2002*, page 16, 2002.

- [35] L. Mendoza. Simulación probabilística de un proceso de reinyección de gas en un yacimiento de gas condensado costa-afuera utilizando diseño experimental. Technical report, UNIVERSIDAD DEL ZULIA. FACULTAD DE INGENIERÍA - ESCUELA DE PETRÓLEO MARACAIBO - EDO. ZULIA, 2005.
- [36] D. Montgomery. *Diseño y Análisis de Experimentos*. Limusa Wiley, 2005.
- [37] K. Narayanan, C. White, L. Lake, and B. Willis. Response surface methods for upscaling heterogeneous geologic models. *Paper Society of Petroleum Engineers 51923 presented at the 1999 Society of Petroleum Engineers Annual Technical Conference and Exhibition*, page 2, 1999.
- [38] G. Newman and J. Martin. Equipment and experimental methods for obtaining laboratory compression characteristics of reservoir rocks under various stress and pressure condition. *52nd Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME*, page 16, 1977.
- [39] G. H. Newman. Pore-volume compressibility of consolidated, friable, and unconsolidated reservoir rocks under hydrostatic loading. *Journal of Petroleum Technology*, page 6, 1973.
- [40] A. Ordoñez. Modelo de pozos en simulación numérica de yacimientos. Technical report, Universidad Central de Venezuela, 2007.
- [41] G.E. Petrosky and F. Farshad. Pressure-volume-temperature correlations for gulf of mexico crude oils. *Paper SPE 51395 first presented at the 1993 SPE Annual Technical Conference and Exhibition. Paper peer approved 28 January 1998.*, 5, 1998.
- [42] F. Rosado-Vazquez, E. Rangel-German, and F. Rodriguez de la Garza. Analysis of capillary, gravity and viscous forces effects in oil/water displacement. *Paper SPE 108722 presented at the 2007 International Oil Conference and Exhibition and Exhibition held in Veracruz, Mexico.*, page 14, 2007.
- [43] C. Sabino. *El Proceso de Investigación*. 1992.
- [44] M.A. Salhi, M. Van Rijen, L. Wei, H Dijk, Z. Alias, and A.Upadhyaya. Structured uncertainty assessment for a mature field through the application of experimental design and response surface methods. *Paper SPE 93529, presented at the 14th SPE Middle East Oil and Gas Conference held in Bahrain.*, page 11, 2005.

- [45] R. Scheaffer and J. McClave. *Probabilidad y Estadística para Ingeniería*. Grupo Editorial Iberoamérica, 1993.
- [46] D. Steppan, J. Werner, and R. Yeater. *Essencial Regression and Experimental Design for Chemists and Engineers*, 1998.
- [47] G. Vergara. Distribución t-student. 2007.
- [48] J. Villa. Geoestadística aplicada. 2007.
- [49] C. White. Effective use of reservoir simulation for sensitivity and optimization studies. 2001.
- [50] C. White and S. Royer. Experimental design as a framework for reservoir studies. *Paper SPE 79676, presented at the SPE Reservoir Simulation Symposium held in Houston, Texas, United States.*, page 7, 2003.
- [51] P. Willhite. *Waterflooding*. SPE, 1986.
- [52] *Eclipse Black Oil.- Schlumberger*.
- [53] *Una Descripción de latex 2e*.
- [54] *The Not So Short Introduction to Latex 2e*.
- [55] *Using Imported Graphics in Latex 2e*.
- [56] Schlumberger. *Technical Description -Eclipse*, 2001.

Apéndice A

Correlaciones Empíricas para el cálculo de Propiedades PVT

Los cálculos de las propiedades PVT se realizaron con correlaciones empíricas, válidas dentro de diversos rangos.

A.1 Presión en el Punto de Burbujeo

- Al Marhoun:

$$p_b = a_1 R_s^{a_2} \gamma_g^{a_3} \gamma_o^{a_4} (T + 460)^{a_5} \quad (\text{A.1})$$

$$a_1 = 5.38088E - 3, a_2 = 0.715082, a_3 = -1.877840, a_4 = 3.143700, a_5 = 1.326570$$

- Dindourck-Christman:

$$p_b = a_8 \left(\frac{R_s^{a_9}}{\gamma_g^{a_{10}} 10^A + a_{11}} \right) \quad (\text{A.2})$$

$$A = \frac{a_1 T^{a_2} + a_3 API^{a_4}}{\left(a_5 + \frac{2R_s^{a_6}}{\gamma_g^{a_7}} \right)^2} \quad (\text{A.3})$$

$$a_1 = 1.4282E - 10, a_2 = 2.84459, a_3 = -6.7489E - 04, a_4 = 1.2252, a_5 = 0.03338, a_6 = -0.027294, a_7 = -0.08422, a_8 = 1.86997, a_9 = 1.22148, a_{10} = 1.37058, a_{11} = 0.011688$$

- Glaso

$$p_b = \text{antilog}(a_1 + a_2 \log(G) - a_3 \log(G)^2) \quad (\text{A.4})$$

$$G = \frac{R_s^{a_4}}{\gamma_g} T^{a_5} \gamma_{API}^{a_6} \quad (\text{A.5})$$

$$a_1 = 1.7669, a_2 = 1.7447, a_3 = 0.30218, a_4 = 0.816, a_5 = 0.3172, a_6 = -0.989$$

- Lasater:

$$p_b = \frac{[p_f(T + 459.67)]}{\gamma_g} \quad (\text{A.6})$$

$$p_f = a_6 - a_7 Y_g + a_8 Y_g^2 \quad (\text{A.7})$$

$$Y_g = \frac{\frac{R_s}{a_1}}{\frac{R_s}{a_1} + \frac{a_2}{M_o}} \quad (\text{A.8})$$

$$M_o = a_3 - a_4 \gamma_{API} + a_5 \gamma_{API}^2 \quad (\text{A.9})$$

$$a_1 = 379.3, a_2 = 350, a_3 = 725.32143, a_4 = 16.03333, a_5 = 0.09524, a_6 = 0.38418, a_7 = 1.20081, a_8 = 9.64868$$

- Petrosky-Farshad:

$$p_b = a_1 \left(\frac{R_s^{a_2}}{\gamma_g^{a_3}} 10^X - a_4 \right) \quad (\text{A.10})$$

$$X = a_5 T^{a_6} - a_7 \gamma_{API}^{a_8} \quad (\text{A.11})$$

$$a_1 = 112.727, a_2 = 0.5774, a_3 = 0.8439, a_4 = 12.340, a_5 = 4.561E - 05, a_6 = 1.3911, a_7 = 7.916E - 04, a_8 = 1.5410$$

- Standing

$$p_b = a_1 \left[\left(\frac{Rs}{\gamma_g} \right)^{a_2} 10^{T a_3 - \gamma_{API} a_4} - a_5 \right] \quad (\text{A.12})$$

$$a_1 = 18.2, a_2 = 0.83, a_3 = 0.00091, a_4 = 0.0125, a_5 = 1.4$$

- Vazquez-Beggs

$$p_b = \left(a_1 \frac{Rs}{\gamma_g} \right) \text{antilog} \left(- \frac{a_3 \gamma_{API}}{460 + T} \right)^{a_2} \quad (\text{A.13})$$

Coeficientes a_i en la Tabla A.1

a_i	$API \leq 30$	$API > 30$
a_1	27,64	0,172
a_2	10,9370	1,0937
a_3	11,172	0,989

Tabla A.1: Coeficientes a_i para p_b de acuerdo con la gravedad API

A.2 Factor de Compresibilidad del Gas (Z)

Correlación de Hall-Yarborough:

$$Z = \left[\frac{0.06125}{Y} p_{pr} t \right] \exp -1.2(1 - t)^2 \quad (\text{A.14})$$

Donde:

p_{pr} : Presión pseudo-reducida

t : Inverso de la temperatura reducida

Y : Densidad Reducida, se puede obtener a partir de la siguiente ecuación:

$$F(Y) = X_1 + \frac{Y + Y^2 + Y^3 + Y^4}{(1-Y)^3} - X_2 Y^2 + X_3 Y^{X_4} = 0 \quad (\text{A.15})$$

Donde:

$$X_1 = -0.06125 p_{pr} t \exp [-1.2(1 - t)^2] \quad (\text{A.16})$$

$$X_2 = (14.76t - 9.76t^2 + 4.58t^3) \quad (\text{A.17})$$

$$X_3 = (90.7t - 242.2t^2 + 42.4t^3) \quad (\text{A.18})$$

$$X_4 = 2.18 + 2.82t \quad (\text{A.19})$$

A.3 Relación Gas-Petróleo

- Al Marhoun: A partir de la ecuación A.1, despejando R_s , se obtiene:

$$R_s = \left(\frac{p}{a_1 \gamma_g^{a_3} \gamma_{API}^{a_4} (T + 460)^{a_5}} \right) \quad (\text{A.20})$$

utilizando los mismos coeficientes a_i que se indican para el calculo de la presión de burbujeo.

- Dindourck-Christman:

$$R_s = \left[\left(\frac{p_b}{a_8} + a_9 \right) \gamma_g^{a_{10}} 10^A \right]^{a_{11}} \quad (\text{A.21})$$

$$A = \frac{a_1 API^{a_2} + a_3 T^{a_4}}{\left(a_5 + \frac{2 API^{a_6}}{p_b^{a_7}} \right)^2} \quad (\text{A.22})$$

$a_1 = 4.8699E - 6$, $a_2 = 5.730982539$, $a_3 = 9.9251E - 03$, $a_4 = 1.776179364$,
 $a_6 = 2.702889206$, $a_7 = 0.744335673$, $a_8 = 3.359754970$, $a_9 = 28.10133245$,
 $a_{10} = 1.579050160$, $a_{11} = 0.928131344$

- Glaso : A partir de la ecuación A.4, calculando las raíces la ecuación, asumiendo que $p_b = p(i)$, y sustituyendo esos valores en la ecuación para R_s obtenida a partir de la ecuación A.5, se determinan los valores de R_s , utilizando los mismos coeficientes que en el cálculo de la presión de burbujeo.

$$R_s = \left(\frac{G}{T^{a_5} \gamma_{API}^{a_6}} \gamma_g \right)^{1/a_4} \quad (\text{A.23})$$

- Lasater:

Partiendo de la ecuación A.6 se obtiene que:

$$p_f = \frac{p \gamma_g}{T + 460} \quad (\text{A.24})$$

ecuación que al igualarse con A.7 permite obtener los valores de las raíces; por otra parte, despejando R_s de la ecuación A.8 se obtiene la forma:

$$R_s = a_1 \left(\frac{\frac{a_2}{M_o} Y_g}{1 - Y_g} \right) \quad (\text{A.25})$$

Utilizando los mismos coeficientes a_i utilizados en el cálculo de la presión de burbujeo, se determinan los valores de R_s para cada punto de presión

- Petrosky-Farshad:

$$R_s = \left(\left(\frac{p}{a_1} + a_2 \right) \bar{\gamma}_g^{a_3} 10^X \right)^{a_4} \quad (\text{A.26})$$

$$X = a_5 \gamma_{API}^{a_6} - a_7 T^{a_8} \quad (\text{A.27})$$

$$a_1 = 112.727, a_2 = 12.340, a_3 = 0.8439, a_4 = 1.7318, a_5 = 7.916E - 04, \\ a_6 = 1.5410, a_7 = 4.561E - 05, a_8 = 1.3911$$

- Standing:

Partiendo de la ecuación A.12 al despejar R_s , se obtiene:

$$R_s = \left[\frac{p + a_1 a_5}{10^{a_3 T - a_1 \gamma_{API}^{a_4}}} \right]^{1/a_2} \gamma_g \quad (\text{A.28})$$

- Vazquez-Beggs:

$$R_s = C_1 \gamma_{gs} p^{C_2} \exp(C_3(\gamma_{API}/(T + 460))) \quad (\text{A.29})$$

Coefficientes en la tabla A.2

C_i	$API \leq 30$	$API > 30$
C_1	0.0362	0.0178
C_2	1.0937	1.187
C_3	25.724	23.93

Tabla A.2: Coeficientes C_i para R_s de acuerdo con la gravedad API

A.4 Factor Volúmetrico de Formación del Petróleo

- Al Marhoun:

$$\beta_o = a_1 + a_2(T - 460) + a_3 M + a_4 M^2 \quad (\text{A.30})$$

$$M = R_s^{a_5} \gamma_g^{a_6} \gamma_{API}^{a_7} \quad (\text{A.31})$$

$$a_1 = 0.4970069, a_2 = 0.862963E - 03, a_3 = 0.182594E - 02, a_4 = 0.318099E - 05, \\ a_5 = 0.74239, a_6 = 0.323294, a_7 = -1.20204$$

- Dindourck-Christman:

$$\beta_o = a_{11} + a_{12} A + a_{13} A^2 + a_{14} (T - 60) \frac{API}{\gamma_g} \quad (\text{A.32})$$

$$A = \frac{\left(R_s^{a_1} \gamma_g^{a_2} / \gamma_{API}^{a_3} + a_4(T - 60)^{a_5} + a_6 R_s\right)^{a_7}}{\left(a_8 + 2R_s^{a_9} / \gamma_g^{a_{10}}(T - 60)\right)^2} \quad (\text{A.33})$$

$$a_1 = 2.51, a_2 = -4.8525, a_3 = 1.1835E + 01, a_4 = 1.3654E + 05, a_5 = 2.2528, \\ a_6 = 1.00E + 01, a_7 = 4.4508E - 01, a_8 = 5.3526, a_9 = -6.309E - 01, \\ a_{10} = 9.0007E - 01, a_{11} = 9.8787E - 01, a_{12} = 7.8651E - 04, a_{13} = 2.6891E - 06, \\ a_{14} = 1.1E - 05$$

- Glaso :

$$\beta_o = 1 + [a_1 + a_2(\log(G)) - a_3(\log(G)^2)] \quad (\text{A.34})$$

$$G = R_s \left(\frac{\gamma_g}{\gamma_{API}} \right)^{a_4} + a_5 T \quad (\text{A.35})$$

$$a_1 = -6.58511, a_2 = 2.91329, a_3 = 0.27683, a_4 = 0.526, a_5 = 0.968$$

- Petrosky-Farshad:

$$\beta_o = a_1 + a_2 [R_s^{a_3} (\gamma_g^{a_4} / \gamma_{API}^{a_5}) + a_6 T^{a_7}]^{a_8} \quad (\text{A.36})$$

$$a_1 = 1.0113, a_2 = 7.2046E - 05, a_3 = 0.378, a_4 = 0.2914, a_5 = 0.6265, \\ a_6 = 0.24626, a_7 = 0.5371, a_8 = 3.0936$$

- Standing:

$$\beta_o = a_1 + a_2 [R_s (\gamma_g / \gamma_{API})^{a_3} + a_6 T^{a_4}]^{a_5} \quad (\text{A.37})$$

$$a_1 = 0.972, a_2 = 1.472E - 04, a_3 = 0.5, a_4 = 1.25, a_5 = 1.175$$

- Vazquez-Beggs:

$$\beta_o = 1 + a_1 R_s + a_2 [(\gamma_g / \gamma_{API})(T - 60)] + a_3 \quad (\text{A.38})$$

Coeficientes a_i en la tabla A.3

a_i	$API \leq 30$	$API > 30$
a_1	4.677E-04	4.677E-04
a_2	1.751E-5	1.1E-05
a_3	-1.8106E-8	1.337E-9

Tabla A.3: Coeficientes a_i para el β_0 de acuerdo con la gravedad API

A.5 Gravedad del Gas corregida para condiciones de separador $p = 100psig$

$$\gamma_{gs} = \gamma_{gp}[1 + 5.912E - 05(\gamma_{API}T \log(p/114.7))] \quad (A.39)$$

γ_{gp} : Gravedad del Gas medida a condiciones de separador p y T.

Compresibilidad del Petróleo

- Dindoruck-Christman:

$$c_o = (a_{11} + a_{12}A + a_{13}A^2)10^{-06} \quad (A.40)$$

$$A = \frac{(R_s^{a_1}(\gamma_g^{a_2}/\gamma_{API}^{a_3}) + a_4(T - 60)^{a_5} + a_6R_s)^{a_7}}{a_8 + (2R_s^{a_9}/\gamma_g^{a_{10}})(T - 60)^2} \quad (A.41)$$

$$a_1 = 0.980922372, a_2 = 0,021003077, a_3 = 0.338486128, a_4 = 20.00006358, \\ a_5 = 0.300001059, a_6 = -0.876813622, a_7 = 1.759732076, a_8 = 2.749114986, \\ a_9 = -1.713572145, a_{10} = 9.999932841, a_{11} = 4.487462368, a_{12} = 0.005197040, \\ a_{13} = 0.000012580$$

- Petrosky-Farshad:

$$c_o = 1.705E - 07R_s^{0.69357}(\bar{\gamma}_g)^{0.1885}\gamma_g^{0.3272}T^{0.6729}p^{-0.5906} \quad (A.42)$$

- Vazquez-Beggs:

$$c_o = \frac{a_1 + a_2R_s + a_3T + a_4\gamma_{gs} + a_5\gamma_{API}}{a_6p} \quad (A.43)$$

$$a_1 = -1433, a_2 = 5, a_3 = 17.2, a_4 = -1180, a_5 = 12.61, a_6 = 10^5$$

γ_{gs} : Gravedad del Gas corregida a condiciones de separador con $p = 100psig$

A.6 Factor Volumétrico de Formación del Gas

En unidades de campo se puede expresar como[1]:

$$\beta_g = 0.0050352 Z \frac{T}{p} \quad (\text{A.44})$$

Z : Factor de Compresibilidad del Gas

T : Temperatura, [°R]

p : Presión, [psia]

β_g : Factor Volumétrico de Formación de Gas en [bbl/SCF]

A.7 Viscosidad

1. Petróleo: Correlación de Beggs-Robinson.[6] Para la viscosidad en el punto del burbujeo

$$\mu_{OD} = 10^X - 1 \quad (\text{A.45})$$

Donde:

$$X = y * T^{-1.163} \quad (\text{A.46})$$

$$y = 10^Z \quad (\text{A.47})$$

$$Z = 3.0324 - 0.02023^\circ API \quad (\text{A.48})$$

La corrección para el petróleo con Gas disuelto:

$$\mu = A * \mu_{OD}^B \quad (\text{A.49})$$

Donde:

$$A = 10.715 * (R_s + 100)^{-0.515} \quad (\text{A.50})$$

$$B = 5.44 * (R_s + 150)^{-0.338} \quad (\text{A.51})$$

R_s : Relación Gas Petróleo [SCF/STB]

T : Temperatura [°F]

μ_{OD} : Viscosidad del petróleo libre de gas a T, [cP]

μ : Viscosidad del petróleo saturado a T, [cP]

2. Gas: Esta propiedad se determina a través de la Correlación de Thodos[34]:

$$[(\mu_g - \mu')\xi + 10^{-4}]^{1/4} = f(\rho_{r,JST}) \quad (A.52)$$

donde, $f(\rho_{r,JST})$ viene dada por:

$$f(\rho_{r,JST}) = a_1 + a_2\rho_{r,JST} + a_3\rho_{r,JST}^2 + a_4\rho_{r,JST}^3 + a_5\rho_{r,JST}^4 \quad (A.53)$$

$$\xi = \frac{T_c(1/6)}{M_w(1/2p_c(2/3))} \quad (A.54)$$

$$a_1 = 0.1023, a_2 = 0.023364, a_3 = 0.05833, a_4 = -0.040758, a_5 = 0.0093324$$

$\rho_{r,JST}$: ρ/ρ_c Densidad Reducida, Adimensional

ρ : Densidad a p y T, [gr/cc]

ρ_c : Densidad en el punto crítico

T_c : Temperatura Critica, [K]

M_w : Peso Molecular, [lb/lb_{mol}]

μ_g : Viscosidad del gas, [cP]

μ : Viscosidad del gas a baja presión, [cP]

A.8 Propiedades del Agua

1. Factor Volumétrico de Formación: Este término se calculó de acuerdo con la siguiente expresión matemática:

$$\beta_w = A_1 + A_2 * p + A_3 * p^2 \quad (A.55)$$

Donde p representa la presión y A_i tiene la siguiente estructura:

$$A_i = a_1 + a_2 * (T - 460) + a_3 * (T - 460)^2 \quad (A.56)$$

Los coeficientes a_i vienen dados para el agua libre de gas y el agua saturada, de acuerdo con los objetivos de este estudio, sólo se hace mención a la tabla para agua libre de gas, los coeficientes correspondientes en la Tabla A.4:

2. Viscosidad: El cálculo de la viscosidad del agua se determina de acuerdo con la

A_i	a_1	a_2	a_3
A_1	0,9947	5,8000E-06	1,02E-06
A_2	-4,2280E-06	1,8376E-08	-6,77E-11
A_3	1,3E-10	-1,3855E-12	4,285E-15

Tabla A.4: Coeficientes a_i para agua libre de gas

siguiente expresión, planteada por Brill y Beggs:

$$\mu_w = \exp 1.003 - 1.479E - 02 * T + 1.982E - 05 * T \quad (\text{A.57})$$

T : Temperatura [°F]

3. Compresibilidad: Este término se determina de acuerdo con:

$$C_w = (C_1 + C_2 * T + C_3 * T^2) * E - 06 \quad (\text{A.58})$$

$$C_1 = 3.8546 - 0.000134 * p \quad (\text{A.59})$$

$$C_2 = -0.01052 + 4.77E - 7 * p \quad (\text{A.60})$$

$$C_3 = 3.9267E - 5 - 8.8E - 10 * p \quad (\text{A.61})$$

T : Temperatura [°F]

p : Presión [psia]

C_w : Compresibilidad del Agua [psi^{-1}]

Apéndice B

Compresibilidad de la roca

La compresibilidad de la roca es un parámetro necesario para la evaluación de yacimientos, debido a que forma parte del sistema de producción natural del mismo por la energía allí acumulada, y la misma cobra mayor importancia cuando se trata de yacimientos subsaturados no sometidos a mantenimiento de presión.

Existen dos tipos de compresibilidad, la Compresibilidad del Volumen Poral o Compresibilidad Bruta C_b , que mide la fracción de cambio del V_p en función del cambio en la presión neta de confinamiento (NOBP, "Net OverBurden Pressure") (Ver ecuación B.1).¹, y la compresibilidad del medio con poros interconectados C_f , la cual refleja los cambios de volumen originados por el aumento de presión en el espacio interconectado de la roca (Ver ecuación B.3)

Las ecuaciones que las define son las siguientes:

$$C_b = -\frac{1}{V_b} \frac{dV}{dP} \quad (\text{B.1})$$

$$C_f = \frac{-\left(\frac{1}{V_p} \frac{dV}{dP}\right)}{\phi} \quad (\text{B.2})$$

es decir:

$$C_f = \frac{C_p}{\phi} \quad (\text{B.3})$$

en la que: ϕ representa la porosidad, P la presión, V_b el volumen bruto de la roca y V_p el volumen de poros interconetados.

Con el objetivo de entender el comportamiento de este factor se han desarrollado

¹<http://inlab.com.ar/Compresib1.htm>

estudios entre los cuales destacan el planteado por Hall [27], fue el primero donde se planteó su importancia, especialmente para yacimientos subsaturados donde omitirla representa errores en la predicción del comportamiento del mismo, debido a que afecta los cálculos del petróleo en sitio de acuerdo con la caída de presión y los estudios del empuje por agua como mecanismo de empuje natural. Hall reconoce que la compresibilidad, total o efectiva, de cualquier roca yacimiento es el resultado de dos elementos independientes, la expansión individual de los granos, que se origina por la disminución de la presión de los fluidos que los rodean, y de la compresión adicional a la que están sometidos los granos como consecuencia de la presión realizada por las rocas suprayacentes.

El estudio fue realizado con núcleos tomados de formaciones de calizas y arenas productoras, los cuales fueron sometidos a condiciones de yacimiento, y presiones por carga litostática con el uso de un equipo constituido por una cámara de presión, manómetros y una bomba hidráulica (Ver Figura B.1), con dicho equipo se realizaron medidas donde se cuantificaba la cantidad de líquido producido por cada paso de presión, los resultados obtenidos fueron: expansión del sistema externo del núcleo, expansión de líquido contenido, y la disminución en el espacio poroso causado por la compresión y expansión de los granos de la roca. La reducción del espacio poroso fue la medida que se utilizó para determinar los valores de compresibilidad para las rocas sometidas a estudio, los mismos se expresan como "cambio en el volumen poroso por unidad de volumen poroso por lpc".

De la observación directa se puede apreciar que, de acuerdo con los resultados obtenidos se sugiere que la compresibilidad de la roca decrece en la medida que la porosidad aumenta. Figura B.2

Posteriormente, en 1973 Newman, hizo un aporte importante al comparar los resultados de laboratorio de su estudio con las correlaciones existentes, con ello verificó que para ese momento sólo habían correlaciones para rocas consolidadas pero que no eran correlaciones adecuadas para rocas no consolidadas. Como conclusiones relevantes se puede mencionar que los valores que él obtuvo en su estudio no describían el comportamiento mostrado por las correlaciones, resultado que era predecible, pues los valores obtenidos eran propios de un yacimiento, los valores de la compresibilidad del volumen poroso dependen del tipo de roca, y que por lo tanto, deberían de existir correlaciones diferentes para rocas consolidadas y rocas no consolidadas.

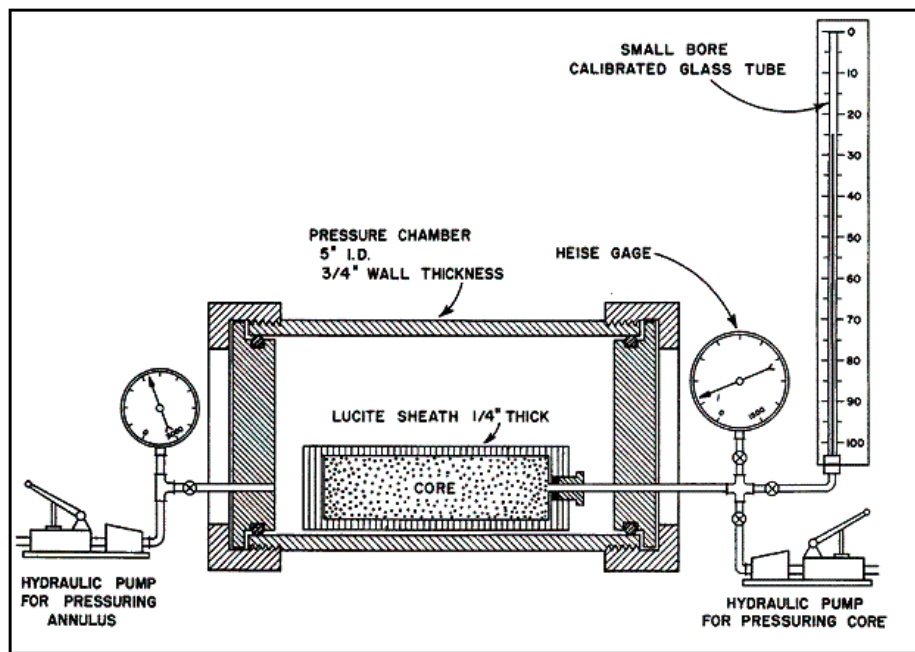


Figura B.1: Diagrama del equipo de laboratorio utilizado par las pruebas de compresibilidad

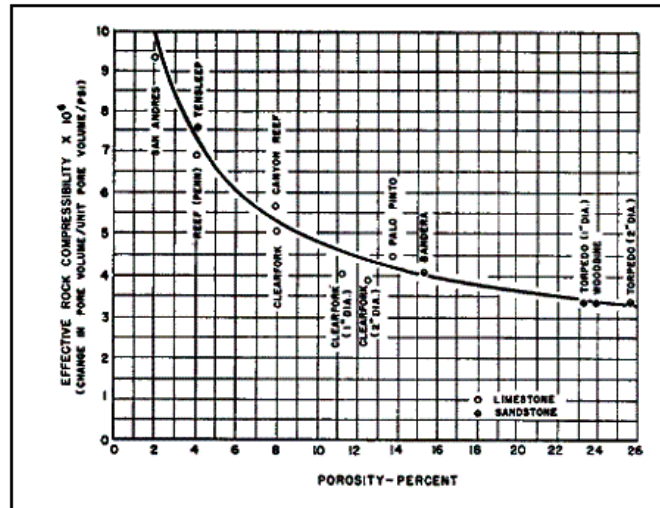


Figura B.2: Compresibilidad efectiva de la roca yacimiento de acuerdo con el estudio de Hall

Aún cuando todos los estudios estuvieron orientados bajo los lineamientos que estableció Hall, en 2004 [11] se presentó una nueva correlación, de acuerdo con la cual, la compresibilidad de la roca se incrementa con la porosidad, es decir, se comporta de manera inversa a lo sugerido por Hall. La nueva correlación toma en cuenta factores como influyentes la relación existente entre los esfuerzos de la roca, el tipo de deformación que se origina producto de los esfuerzos y la compresibilidad del material sólido, es decir, la rigidez del esqueleto, lo que permite obtener la nueva ecuación para describir el comportamiento de la compresibilidad. Ecuación B.4

$$C_p = \left(\frac{\phi}{1-\phi}\right)C_s \quad (\text{B.4})$$

Donde C_s es la parte de la ecuación asociada a las propiedades mecánicas de la roca, como por ejemplo, el Módulo de Poisson y el Módulo de Elasticidad, de manera general tiene como expresión matemática:

$$C_s = -\frac{dV_s}{V_b d\sigma_s} \quad (\text{B.5})$$

y para materiales sólidos:

$$C_s = \frac{3(1-2\nu)}{E_s} \quad (\text{B.6})$$

Aún cuando el uso de estas ecuaciones ofrece un resultado más realista, requiere de datos que no siempre se encuentran disponibles.

Apéndice C

Presión Capilar

La presión capilar, como concepto asociado al medio poroso se desarrolla a partir de la representación de los fenómenos capilares en los tubos capilares, ver Figura C.1. La interface agua/petróleo o aire/agua es plana debido a la acción de las fuerzas de atracción que se generan entre las paredes de los tubos y el fluido que contienen se distribuyen a lo largo de todo el perímetro y no penetran en el interior de las mismas en gran medida, por lo tanto las fuerzas en la interface son iguales.

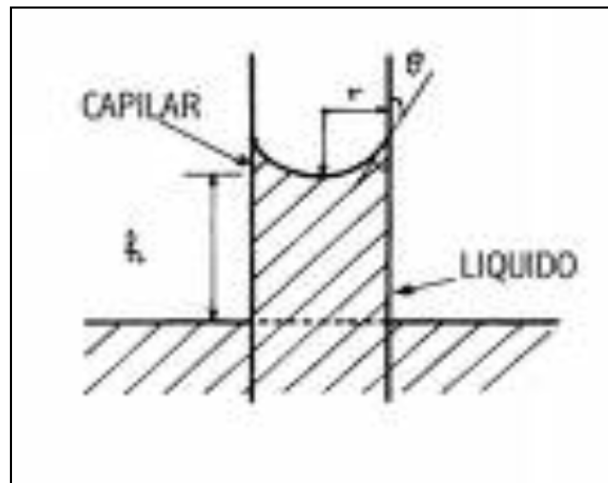


Figura C.1: Presión Capilar

En el yacimiento, los poros de las rocas son análogos a los tubos capilares de diámetro pequeño; cuando el diámetro de los poros es pequeño las fuerzas superficiales, inducidas por la mojabilidad preferencial de la roca por uno de los fluidos, extienden dicho fluido por toda la superficie, ocasionando medidas de presión

diferentes entre los fluidos presentes. Si se considera un sistema agua - petróleo en un tubo capilar de vidrio dispuesto horizontalmente, en equilibrio estático, el agua mojará preferencialmente la superficie del vidrio con un ángulo de contacto cercano a cero, por otra parte, si realizan mediciones en los extremos del tubo capilar se podrá apreciar que la presión de la fase petróleo es un poco mayor que la de la fase agua, independientemente de la longitud del tubo. El agua puede ser desplazada del capilar si se inyecta petróleo y el mismo puede ser desplazado espontáneamente si la presión de la fase es reducida, aún cuando la presión de la fase agua sea menor que la de la fase petróleo.

Si se hace un análisis de fuerza en la línea de contacto entre las fases, se tiene, para condiciones de equilibrio:

$$\sum_{i=1}^n Fx = 0 \quad (C.1)$$

$$P_{nw} - P_w = \frac{2(\sigma_{nw} - \sigma_w)}{r} \quad (C.2)$$

por otra parte, sabiendo que $\sigma_{nw} - \sigma_w = \sigma_{nww} \cos \theta$ y sustituyendo en la ecuación C.2, se obtiene:

$$P_{nw} - P_w = \frac{2\sigma_{nww} \cos \theta}{r} \quad (C.3)$$

La diferencia en la ecuación C.3 es definida como P_c , para los casos donde los sean sistemas agua/petróleo, siendo P_c negativa. Es importante mencionar que esta ecuación fue desarrollada como un caso especial de la desarrollada por LaPlace:

$$P_{nw} - P_w = \sigma_{nw} \left(\frac{1}{r_1} + \frac{1}{r_2} \right) \quad (C.4)$$

donde r_1 y r_2 son los radios de curvatura para la interface. [51]

Dependencia de la Saturación y la Presión Capilar

La presión capilar como una función de la saturación, es, aparentemente la primera relación funcional establecida para que el flujo de fluidos en medio porosos sea medida. Existen 5 métodos para medir dicha relación, entre algunos de ellos se destacan: El Método de la Presión de Vapor, el de la Centrifuga y el de la Presión de Celda, sin embargo, todos tienen implícito el comportamiento descrito a continuación.

Los cambios en el radio de curvatura y en la presión capilar están acompañados por el cambio en la distribución de los fluidos, es decir, la saturación de la fase mojanete es pequeña y la saturación de la fase no mojanete es mayor a grandes elevaciones.

Esto se representa a través de la siguiente ecuación:

$$\sigma \left(\frac{1}{r_1} + \frac{1}{r_2} \right) = \Delta(\rho gh) + P_{co} \quad (C.5)$$

la cual indica que en un sistema estático de dos fases la curvatura media en diferentes puntos de la misma cambia con la elevación, es decir, el radio de curvatura es pequeño a grande elevaciones, es este cambio de curvatura el que permite grandes discontinuidades en la presión a lo largo de grandes elevaciones.

Una relación funcional entre la saturación y la presión capilar puede ser visualizado en un modelo de sección transversal del espacio poroso, Figura C.2:

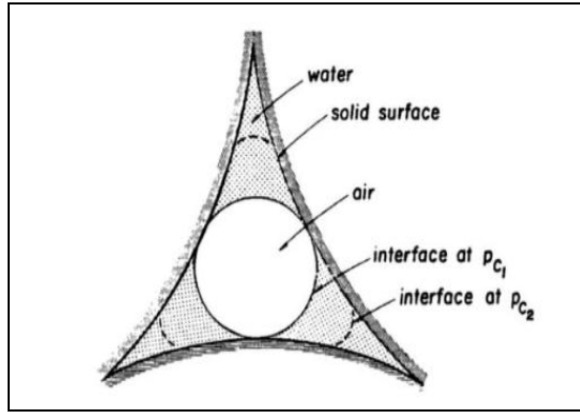


Figura C.2: Modelo de sección transversal de un poro con saturación de agua variable

es decir $S = f(P_c)$.

Para el presente estudio el comportamiento de la Presión Capilar será modelado de acuerdo con la siguiente ecuación[42]:

$$P_c = P'_c \ln(S) \quad (C.6)$$

Donde P'_c es la constante de la Presión Capilar, la cual se determina a través de pruebas de laboratorio.

Apéndice D

Permeabilidades Relativas

La permeabilidad es una propiedad de la roca que se define como una medida de la capacidad que tiene una roca para permitir el paso de fluidos a través de los poros interconectados de la misma. Se pueden distinguir 3 tipos de permeabilidad[5]:

- Permeabilidad Absoluta: es aquella que corresponde a la roca saturada 100% por un fluido que no reacciona con la roca.
- Permeabilidad Efectiva respecto a un fluido: corresponde a aquellos casos donde la roca tiene una saturación de dicho fluido menor a 100%.
- Permeabilidad relativa: se produce cuando existe más de un fluido en el medio poroso, y se expresa como una relación entre la permeabilidad efectiva y absoluta de cada fluido.

Para el estudio de esta propiedad se han desarrollado varios modelos matemáticos que se pueden clasificar en 4 categorías[28]:

1. Modelos Capilares: Asumen que el medio poroso consiste en un conjunto de tubos capilares de diversos diámetros con un fluido que empapa toda la muestra. Este modelo ignora la naturaleza de interconectada del medio poroso y frecuentemente no muestra resultados reales.
2. Modelos Estadísticos: También están basados en que el medio poroso es un conjunto de tubos capilares de diferentes diámetros, distribuidos aleatoriamente, y describen su comportamiento asumiendo que el mismo es un grupo de planos, perpendiculares al eje de los tubos capilares, luego los planos son imaginariamente reorganizados y reensamblados aleatoriamente. Este tipo de modelos tiene la misma deficiencia de los modelos capilares, no toman en cuenta la naturaleza interconectada de los poros.

3. Modelos Empíricos: Tienen como principio relaciones empiricas propuestas a través de la descripción de resultados experimentales, y en general han brindado buenos resultados.
4. Modelos de Redes: Están basados en modelos de flujo de fluidos en medios porosos, recreados a través del uso de resistores eléctricos conectados en red como una computadora analógica. Probablemente son la mejor herramienta para el entendimiento de flujos de fluidos en medios porosos.

La predicción del comportamiento de la permeabilidad relativa es un área de constante investigación, debido a la gran complejidad de la misma, sin embargo, se han presentado grandes aportes a partir de estudios realizados, entre ellos, los planteados por Carman-Kozeny y Brooks-Corey, que serán presentados a continuación[7].

- Modelo Modificado Brooks-Corey: tiene como fundamento los estudios realizados por Corey en 1954, donde a través de estudios de laboratorio se plantearon las siguientes ecuaciones:

$$k_{ro} = (S_{oe})^4 \quad (D.1)$$

$$k_{rg} = (1 - S_{oe})^2(1 - S_{oe}^2) \quad (D.2)$$

Donde:

$$S_{oe} = \frac{S_o - S_{or}}{1 - S_{or}} \quad (D.3)$$

y S_o : Saturación de Petróleo

S_{or} : Saturación Residual de Petróleo

k_{ro} : Permeabilidad Relativa para el Petróleo

k_{rg} : Permeabilidad Relativa para el Gas

Luego Brooks y Corey hicieron una modificación para establecer una relación general que pudiera correlacionar una saturación normalizada de petróleo (ecuación D.3) para la presión capilar y la presión:

$$S_{oe} = (P_d/P_c)^\lambda \quad (D.4)$$

Donde:

P_d : Presión de Entrada

P_c : Presión Capilar

λ : Índice de Distribución del tamaño de los Poros

Utilizando λ y la saturación de petróleo normalizada de la ecuación D.3 ambos autores introdujeron otra formulación:

$$k_{rw} = (S_e)^{\left(\frac{s + 3\lambda}{\lambda}\right)} \quad (D.5)$$

$$k_{rnw} = (1 - S_{oe})^2 (1 - S_e^{\frac{2+\lambda}{\lambda}}) \quad (D.6)$$

Donde:

k_{rw} : Permabilidad Relativa de la Fase Mojante

k_{rnw} : Permeabilidad Relativa de la Fase no Mojante

Sin embargo, las ecuaciones D.5 y D.6 presenta como limitación importante que sólo permiten calcular las permeabilidades relativas para una condición de drenaje, es decir, con ella no se puede definir el proceso de imbibición, lo que conllevó a realizar unos ajustes que se tradujeron en el Modelo Modificado de Brooks Corey, mejor conocido como *Modelo de la Ley de Potencia*, el cual describe el comportamiento de las permeabilidades relativas para sistemas bifásicos a través de las siguientes ecuaciones:

$$k_{ro} = k'_{ro} \left(\frac{1 - S_w - S_{or}}{1 - S_{wir} - S_{or}} \right)^{n_o} \quad (D.7)$$

$$k_{rw} = k'_{rw} \left(\frac{S_w - S_{wir}}{1 - S_{wir} - S_{or}} \right)^{n_w} \quad (D.8)$$

Donde:

k_{ro} : Permeabilidad Relativa del Petróleo normalizada a la permabilidad del aire

k_{rw} : Permeabilidad Relativa del Agua normalizada a la permabilidad del aire

k'_{ro} : Endpoint de la permeabilidad relativa para el petróleo

k'_{rw} : Endpoint final de la permeabilidad relativa para el agua

S_w : Saturación de Agua

S_{wir} : Saturación de Agua Irreducible

S_{or} : Saturación de Petróleo Residual

n_o : Exponente Corey para el Petróleo

n_w : Exponente Corey para el Agua

Algunas de las aplicaciones de estas ecuaciones es la estimación del valor de S_{or} , dar soporte a la regla de pulgar para medir la calidad de la mojabilidad en muestras de roca.

Entre sus limitaciones, destacan: es modelo es función directa de los End points de las curvas de permabilidad relativa, si no se conocen no es posible utilizar sus ecuaciones y no es modelo que sirva para predecir el comportamiento de las curvas sólo sirve para suavizar la tendencia de las ya existentes.

- Modelo Modificado Carman-Kozeny: es el resultado de combinar el trabajo de Kozeny (1927), quien estudió la dependencia que existía entre la porosidad y la permeabilidad planteando el concepto de las propiedades texturales y correlacionando factores como la forma, la tortuosidad, entre otros; y el trabajo de Carman (1937), quien desarrolló la relación directa entre la porosidad y permeabilidad. La ecuación de este modelo se escribe:

$$k = \frac{\phi_e^3}{(1 - \phi_e^3)^2} \left(\frac{1}{F_s \tau S_{gv}^2} \right) \quad (D.9)$$

Donde k es la permeabilidad absoluta en μm^2 , ϕ_e es la porosidad efectiva, F_s es el factor de forma (2 para espacios circulares), τ es la tortuosidad y S_{gv} es la superficie por unidad de volumen de cada grano en μm^{-1} .

Luego, entre 1992 y 1993 se hicieron reajustes de la ecuación, quedando esta de la siguiente forma:

$$0.0314 \sqrt{\frac{k}{\phi_e}} = \frac{\phi_e}{1 - \phi_e} \left(\frac{1}{\sqrt{F_s \tau S_{gv}}} \right) \quad (D.10)$$

Agrupando términos, se pueden definir nuevas propiedades como el Índice de Calidad del Yacimiento, Grupo de Porosidad y el Indicador de Zona de Flujo (RQI, PG, FZI por sus siglas en inglés respectivamente) y se definen como:

$$RQI = 0.0314 \sqrt{\frac{k}{\phi_e}} \quad (D.11)$$

$$PG = \frac{\phi_e}{1 - \phi_e} \quad (D.12)$$

$$FZI = \left(\frac{1}{\sqrt{F_s} \tau S_{gv}} \right) \quad (D.13)$$

Sustituyendo estas ecuaciones en la ecuación D.10 se obtiene una nueva relación, que al aplicarle logaritmo en ambos lados, se obtiene una relación lineal de pendiente 1:

$$\log RQI = \log PG + \log FZI \quad (D.14)$$

Sin embargo, se han hecho más avances y para el año 2006, Behrenbruch estableció un nuevo modelo semi-empírico, para predecir las permeabilidades relativas, basado en las ecuaciones D.11 y D.12, parámetros que fueron redefinidos para modelar adecuadamente el comportamiento de la permeabilidad en presencia de dos fases, por lo que, se tiene que:

Permeabilidad Relativa del Agua

$$k_{rw} = \frac{1014 m_w^2 \phi_e^3 S_w}{k} \left(\frac{S_w}{1 - \phi_e S_w} - \frac{S_{wir}}{1 - \phi_e S_{wir}} \right)^2 \quad (D.15)$$

Permeabilidad Relativa del Petróleo

$$k_{ro} = \frac{1014 m_o^2 \phi_e^3 (1 - S_w)}{k} \left(\frac{1 - S_w}{1 - \phi_e (1 - S_w)} - \frac{S_{or}}{1 - \phi_e S_{or}} \right)^2 \quad (D.16)$$

Apéndice E

Error en la medición de la presión y de los parámetros PVT

Siempre que se mide un fenómeno se hace destacar la importancia de maximizar la exactitud y precisión. La calidad de la medición dependerá de la exactitud y precisión del sistema de adquisición de datos, elementos que vienen dados por la resolución del instrumento y otros factores, como la habilidad del medidor de manipular el instrumento.

Sin embargo, es importante diferenciar entre exactitud y precisión, la exactitud de una medición representa cuán cerca está la medición del valor real de la cantidad que se desea medir, por lo que, ésta representa que tan correcto es el resultado. Por su parte, la precisión de una medición describe cuán exacta fue la determinación del resultado sin hacer referencia a lo que éste significa. Otro término que es la precisión relativa, la cual indica la incertidumbre en una medición como una fracción del resultado.

Para tener una mejor comprensión de los conceptos ya descritos se citará el siguiente ejemplo: Suponga que se está midiendo la altura de una mesa con una cinta métrica y se obtuvo que la longitud es de 1.502 metros. Esto representa que la cinta métrica pudo leer distancias de al menos un milímetro. Para la mayoría de los casos esto se considera bastante preciso, con una precisión relativa alrededor de $1/1500$. Ahora suponga que realizó una nueva medición y el resultado fue 1.510 metros. Después de reflexionar acerca de los resultados obtenidos notó que la técnica de medición para la primera lectura no era la correcta, porque la cinta métrica no estaba completamente vertical. Por tanto la primera medición no fue exacta.

En la industria petrolera es muy importante diferenciar entre estos términos debido a que puede ayudar en la obtención de mediciones de mejor calidad. Por ejemplo, la exactitud de un instrumento se puede mejorar notablemente si se calibra adecuadamente según las indicaciones del fabricante (permitirá que el resultado se acerque más al valor real de la magnitud que se quiere medir). La precisión está relacionada con el instrumento en sí, es decir, los dispositivos que utiliza este para realizar la medición (algo que es más difícil controlar por el ingeniero ya que debería conocer la precisión del dispositivo de medición, o si está a disponibilidad, elegir entre los diferentes tipos de dispositivos de medición con sus respectivas precisiones).

E.1 Error en la medición de presión

Las herramientas actuales que utilizan medidores de cuarzo poseen resoluciones alrededor de [0.1 psi - 0.01 psi], y la precisión de [0.01% - 0.05% del valor de la lectura]. Esto implica que para medir una presión de 10000 psi se tendría un error asociado con la precisión del instrumento de [1 psi - 5 psi]. Sin embargo, hay que recordar que la historia de declinación de presión posee mediciones hechas en el pasado, y a medida que avanza la tecnología los errores asociados con la medición de presión van disminuyendo.

E.2 Error en la medición de la gravedad API del petróleo

La mayoría de los instrumentos disponibles en el laboratorio para medir la gravedad API tienen resoluciones que varían de entre [0.1 °API - 0.2 °API], la precisión también está en ese orden, es decir, [0.1 °API - 0.2 °API]. En el laboratorio es importante conocer la reproducibilidad, es decir, la diferencia entre dos resultados independientes, obtenidas por diferentes operadores trabajando en diferentes laboratorios con el mismo material de medición. La reproducibilidad del método de medición estándar para la gravedad API se encuentra alrededor de [0.5 °API].

E.3 Error en la medición de la gravedad específica del gas

De las herramientas consultadas para la medición de la gravedad específica del gas las resoluciones se encontraron alrededor de [0.02% de la lectura realizada], y la

precisión de los instrumentos en el siguiente rango [0.1% - 1% de la lectura].

E.4 Error en la medición de la temperatura del yacimiento

Las resoluciones están en el orden de [0.0001 °C- 0.001 °C], las precisiones de los equipos varían en el rango [0.02 °C - 0.05 °C]. La exactitud efectiva de los registros de temperatura es de +/- 0.5 °C.

E.5 Error en la medición de la relación gas-petróleo en solución

La medición del gas en solución en el laboratorio se realiza mediante dos procedimientos. Estos son liberación diferencial y la prueba del separador (también llamada liberación flash). Estos dos procesos se llevan a cabo en una celda PVT. Por lo tanto el error en la medición esta asociado con los dispositivos que se utilicen para medir los volúmenes en la celda Bu y Damsleth [9] mencionan que la incertidumbre relativa experimental se encuentra en el rango [1% - 2%] de la lectura realizada.[23]

Nomenclatura

En esta parte del informe se presenta la nomenclatura adoptada para este informe. Los siguientes símbolos son los adoptados por la Sociedad de Ingenieros Petroleros estadounidense (SPE).

Tabla E.1: Símbolos adoptados

Símbolos	Descripción
c	Compresibilidad. lpc^{-1}
c_f	Compresibilidad de la formación. lpc^{-1}
c_o	Compresibilidad del petróleo. lpc^{-1}
c_t	$c_t = c_o S_o + S_w c_w + S_g c_g + c_t$ compresibilidad total. lpc^{-1}
c_w	Compresibilidad del agua. lpc^{-1}
G	Capa inicial de gas. PCN
G_{iny}	Cantidad acumulada de gas inyectado. PCN
G_p	Producción acumulada de gas. PCN
$GOES$	Gas Original en Sitio
h	Espesor de la formación. pie
J	Indice de Productividad. BY/día/lpc
J_p	Indice de Productividad actual. BY/día/lpc
J_f	Indice de Productividad futuro. BY/día/lpc
k	Permeabilidad. mD
k_g	Permeabilidad al gas. mD
k_o	Permeabilidad al petróleo. mD
k_w	Permeabilidad al agua. mD
N	Petróleo original en sitio. BN
N_p	Producción acumulada de petróleo. BN.
m	Relación entre el tamaño de la capa inicial de gas y la cantidad inicial de petróleo. Adimensional
p_i	Presión promedio inicial. lpc
p	Presión promedio actual. lpc
p_b	Presión de Burbujeo. lpc
p_{pc}	Presión pseudocrítica. °F
p_{pr}	Presión pseudoreducida. Adimensional
$POES$	Petróleo Original en Sitio

Tabla E.1: Símbolos adoptados (cont.)

Símbolos	Descripción
q	Tasa de producción. BN/día
q_w	Tasa de producción de gas. MPCN/día
q_o	Tasa de producción de petróleo. BN/día
q_w	Tasa de producción de agua. BN/día
R_p	Relación gas-petróleo acumulada. PCN/BN
RGP	Relación gas-petróleo instantánea. PCN/BN
R_{si}	Solubilidad del gas inicial. PCN/BN
R_s	Solubilidad del gas. PCN/BN
S_w	Fracción del V_p ocupado por agua. Adimensional
S_o	Fracción del V_p ocupado por petróleo. Adimensional
S_g	Fracción del V_p ocupado por gas. Adimensional
T	Temperatura del Yacimiento. °F
T_{pc}	Temperatura pseudocrítica. °F
T_{pr}	Temperatura pseudoreducida. Adimensional
W_{iny}	Cantidad acumulada de agua inyectada. BN
W_e	Influjo de agua. BY
W_p	Producción acumulada de agua. BN
V_p	Volumen poroso. BY
β_o	Factor volumétrico del petróleo. BY/BN
β_g	Factor volumétrico del gas. BY/PCN
β_w	Factor volumétrico del agua. BY/BN
Δp	Variación en la presión. lpc
γ_o	Gravedad específica del petróleo. Adimensional
γ_g	Gravedad específica del gas. Adimensional
μ	Viscosidad, cP
μ_g	Viscosidad del gas, cP
μ_o	Viscosidad del petróleo, cP
μ_w	Viscosidad del gas, cP
ρ_r	Densidad reducida. Adimensional

Tabla E.2: Unidades

Unidad	Significado	Medida
<i>BY</i>	Barriles de Yacimiento	Volumen
<i>BN</i>	Barriles de Normales	Volumen
<i>cP</i>	Centipoise	Viscosidad
°F	Fahrenheit	Temperatura
<i>lpc</i>	Libra por pulgada cuadrada	Presión
<i>mD</i>	Milidarcy	Permeabilidad
<i>PCN</i>	Pies Cúbicos Normales	Volumen
°R	Rankine	Temperatura