

**HERRAMIENTA SOFTWARE PARA EL AJUSTE DE LA ECUACIÓN DE
ESTADO DE PENG-ROBINSON A DATOS EXPERIMENTALES PVT DE
YACIMIENTOS CERCANOS AL PUNTO CRÍTICO UTILIZANDO REDES
NEURONALES**

ING. FERNANDO ENRIQUE CALVETE GONZALEZ

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICO-MECANICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
MAESTRÍA EN INFORMÁTICA
BUCARAMANGA
FEBRERO 2004**

**HERRAMIENTA SOFTWARE PARA EL AJUSTE DE LA ECUACIÓN DE
ESTADO DE PENG-ROBINSON A DATOS EXPERIMENTALES PVT DE
YACIMIENTOS CERCANOS AL PUNTO CRÍTICO UTILIZANDO REDES
NEURONALES**

ING. FERNANDO ENRIQUE CALVETE GONZÁLEZ

Trabajo de investigación para optar el título de Magíster en Informática

Director

ZULY H. CALDERÓN C., PH.D.

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICO-MECANICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
MAESTRÍA EN INFORMÁTICA
BUCARAMANGA
FEBRERO 2004**

DEDICATORIA

“Dedico este trabajo con todo mi corazón a mis padres Andrés y Belén, por su amor, paciencia y esmero”.

AGRADECIMIENTOS

Expreso mis mas sinceros agradecimientos a:

A la Universidad Industrial de Santander, por sus puertas abiertas.

A Maria Eugenia, el motor de la maestría.

A los profesores que quieren plasmar en cada uno la búsqueda de la verdad.

A la doctora Zuly Calderón, por su apoyo incondicional.

A Jairo Enrique Arrieta, un amigo de verdad.

A los compañeros de clase, por aguantarme siempre el mismo cuento.

A los compañeros del grupo GERENE: Mike, Robert y Miguel, creyentes fieles de las redes neuronales.

A Carlos Brewer, Henry Cárdenas y Edgardo Lorduy, amigos, creyentes e incondicionales.

A todos los que se consideran mis amigos, por serlo.

A mis padres, por todo y más.

CONTENIDO

	pág.
1. MEDIDAS PVT DE LABORATORIO.....	6
1.1. ANÁLISIS PRELIMINARES	6
1.1.1. Selección de la muestra.....	7
1.1.2. Análisis composicional de fluidos.	8
1.1.2.1. Precisión de las mediciones composicionales.	9
1.1.3. Validación de las muestras de gas y aceite a condiciones del separador.....	11
1.1.4. Recombinación de muestras.	12
1.2. EXPERIMENTOS PRESIÓN - VOLUMEN.....	13
1.2.1. Expansión a Masa Constante (EMC).....	14
1.2.2. Agotamiento a volumen constante (AVC)	18
1.2.3. Liberación Diferencial (LD)	21
1.3. PRUEBAS DE SEPARACIÓN MULTI-ETAPAS.	21
1.4. PRUEBAS DE INYECCIÓN DE GAS.	22
1.4.1. Prueba de hinchamiento.	23
1.4.2. Prueba de vaporización.	23
1.5. PRUEBAS DE MISCIBILIDAD	24
1.6. DIAGRAMA DE FLUJO DE UNA PRUEBA PARA GAS CONDENSADO.....	25
1.7. SOLICITUD DE UN ANÁLISIS PVT	26
2. MATEMÁTICA DEL COMPORTAMIENTO DE FASES.....	31
2.1. PROCESO DE AGOTAMIENTO EN YACIMIENTOS DE ACEITE VOLÁTIL CERCANO AL PUNTO CRÍTICO.....	31

2.2. PROCESO DE AGOTAMIENTO EN YACIMIENTOS DE GAS CONDENSADO.....	32
2.3. PRUEBAS DE LABORATORIO	33
2.4. CÁLCULOS FLASH.....	35
2.5. MODELADO DEL EQUILIBRIO DE FASES.....	37
2.5.1. Métodos de solución de ecuaciones de equilibrio líquido-vapor.....	38
2.5.1.1. Métodos de Iteración.....	39
2.5.1.2. Métodos que utilizan la Minimización de la Energía Libre de Gibbs.....	43
3. MÉTODOS Y CÁLCULOS UTILIZADOS EN LA SIMULACIÓN PVT	45
3.1. ECUACIONES DE ESTADO	46
3.1.1. Ecuación de Estado de Peng Robinson.	48
3.1.1.1. Formulación de la Ecuación.	49
3.1.1.2. Coeficientes de Interacción Binaria.	54
3.1.1.3. Correlaciones empleadas para determinar Coeficientes de Interacción Binaria.....	55
3.1.1.4. Modificaciones a la Ecuación de Estado de Peng – Robinson.	58
3.2. RELACIONES DE EQUILIBRIO	64
3.2.1. Método de Varotsis para obtener relaciones de equilibrio.	66
3.3. CÁLCULOS FLASH.....	70
3.3.1. Método de Nghiem.....	71
3.3.1.1. Algoritmo del método de Sustituciones Sucesivas modificado de Nghiem.	74
3.3.1.2. Método de Rachford & Rice.	76
3.4. CÁLCULO DE PRESIONES DE SATURACIÓN.....	77
3.4.1. Determinación de la Presión de Rocío.	77
3.4.2. Determinación de la presión de burbuja.....	81
3.5. DETERMINACIÓN DE LAS PROPIEDADES FÍSICAS DE LAS FRACCIONES DE HIDROCARBUROS	84
3.5.1. Método de Katz-Firoozabadi.	85

3.5.2. Método de Riazi & Daubert.....	87
3.5.3. Correlación de Edmister para determinar el Factor Acéntrico (w).	88
3.6. CARACTERIZACIÓN DE LA FRACCIÓN PESADA (C_{7+}).....	88
3.6.1. Método de Ahmed para Caracterización de la fracción pesada.....	89
4. SIMULACIÓN DE PRUEBAS PVT	93
4.1. PROCEDIMIENTO PARA SIMULAR LA EXPANSIÓN A COMPOSICIÓN CONSTANTE	94
4.1.1. Algoritmo para simular la prueba.....	95
5. AJUSTE DE ECUACIONES DE ESTADO.....	98
5.1. ANTECEDENTES.....	99
5.2. METODOLOGÍA DEL AJUSTE	102
5.2.1. Planteamiento del ajuste de los CIB.....	106
5.2.2. Método de regresión múltiple.....	109
6. REDES NEURONALES ARTIFICIALES.....	115
6.1. BREVE HISTORIA	116
6.2. FUNCIONAMIENTO BÁSICO	117
6.2.1. Neuronas y conexiones.....	118
6.2.2. Tipos de aprendizaje básicos.....	121
6.3. MODELO SIMPLE DE RED NEURONAL	122
6.3.1. Regla de aprendizaje de Hebb.	123
6.3.2. Entrenamiento de la red.....	125
6.3.3. Modelo del Perceptrón.	127
6.3.4. Limitaciones del modelo.....	128
6.4. PERCEPTRÓN MULTICAPA	129
6.4.1. Formulación vectorial.	131
6.5. RETROPROPAGACIÓN.....	132
6.5.1. Funcionamiento de una BPN.	133
6.5.2. La regla delta generalizada.	135
6.5.2.1. Actualización de pesos de la capa de salida.....	138
6.5.2.2. Actualizaciones de los pesos de capas ocultas.....	143

6.5.2.3. Resumen de BPN.	145
6.6. CONSIDERACIONES PRÁCTICAS	147
6.6.1. Datos de entrenamiento.	147
6.6.2. Dimensionamiento de la red.....	149
6.6.3. Pesos y parámetros de aprendizaje.....	150
7. METODOLOGÍA DE AJUSTE UTILIZANDO REDES NEURONALES ARTIFICIALES	153
7.1. DISEÑO DEL PROYECTO	154
7.2. METODOLOGÍA DESARROLLADA.....	158
7.2.1. Planteamiento general.....	158
7.2.2. Metodología propuesta.....	160
7.2.3. Datos de entrada.....	162
7.2.4. Entrenamiento y predicción con redes neuronales.....	163
7.2.5. Validación de los parámetros de la predicción.....	165
7.3. SOFTWARE DE REDES NEURONALES	166
7.3.1. Software ANNBack.	166
7.3.1.1. Entrada de datos y entrenamiento.....	167
7.3.1.2. Entrenamiento de la red.	168
7.3.1.3. Evaluación y datos de salida de la red.....	171
7.4. AJUSTE DE LOS PARÁMETROS DE LA ECUACIÓN DE ESTADO UTILIZANDO ALGORITMOS GENÉTICOS	173
7.4.1. Conceptualización.....	174
7.4.2. Metodología de ajuste utilizando algoritmos genéticos.	174
7.4.2.1. Cálculo del volumen relativo.	175
7.4.2.2. Cálculo de la presión de saturación.....	176
8. ANÁLISIS DE RESULTADOS Y ESTUDIO COMPARATIVO.....	178
8.1. CÁLCULO DE LA PRESIÓN DE SATURACIÓN	178
8.2. EXPANSIÓN A COMPOSICIÓN CONSTANTE	181
8.3. ANÁLISIS DE RESULTADOS	183
9. MANUAL DEL USUARIO	186

9.1. INTERFASE GENERAL	187
9.1.1. Barra de menú-Archivo.....	187
9.1.2. Barra de menú-Muestras Experimentales.....	188
9.1.3. Barra de menú-Valores de Prueba.....	189
9.1.4. Barra de menú-Acciones.....	190
9.1.5. Barra de menú-Herramientas.....	191
9.1.3. Barra de menú-Ayuda.	199
9.2. CARPETAS PARA EL TRABAJO EXPERIMENTAL	200
9.2.1. Carpeta de Pruebas.....	200
9.2.2. Carpeta Muestra.....	201
9.2.3. Carpeta Valores de Prueba.....	202
9.2.4. Carpeta Resultados.	205
9.3. EXPERIMENTO DE EXPANSION A MASA CONSTANTE	206
CONCLUSIONES	209
RECOMENDACIONES.....	211
REFERENCIAS BIBLIOGRAFICAS.....	212
ANEXOS	

LISTA DE TABLAS

pág.

Tabla 1. Constantes C_0 a C_7 para las ecuaciones 3.45 y 3.48.....	60
Tabla 2. Valores de a, b, y m para CH_4 y N_2 optimizados.	60
Tabla 3. Valores del parámetro de translación para C_1 - C_6 y CO_2	62
Tabla 4. Valores de d y e para la corrección de Jhaveri y Youngren.	63
Tabla 5. Coeficientes de la Ecuación 3.60.....	68
Tabla 6. Propiedades Físicas Generalizadas (KFW).	86
Tabla 7. Propiedades Físicas Generalizadas (KFW).	87
Tabla 8. Constantes para la Ecuación 3.99.....	87
Tabla 9. Pendiente de la ecuación 3.104 para el cálculo de M_{Cn+}	91
Tabla 10. Coeficientes para la ecuación 3.105.	91
Tabla 11. Casos de entrenamiento en un sistema de tres neuronas.	123
Tabla 12. Casos de entrenamiento en un sistema de tres neuronas.	127
Tabla 13. Modelos neuronales y técnicas estadísticas.	156
Tabla 14. Modelos neuronales y técnicas estadísticas.	162
Tabla 15. Comparación de resultados en el cálculo de la presión de saturación para diferentes sistemas de hidrocarburos por medio de diferentes procedimientos.	179
Tabla 16. Estructuras neuronales utilizadas en el estudio de presión de saturación.	180
Tabla 17. Datos físicos para el entrenamiento de la red neuronal.....	179
Tabla 18. Resultados de la Expansión a Masa Constante para el Gas2*. ..	182
Tabla 19. Parámetros ajustados para la Expansión a Masa Constante del Gas2*.	182

LISTA DE FIGURAS

pág.

Figura 1. Representación esquemática de una Expansión a Masa Constante para un gas condensado.	15
Figura 2. Comportamiento P-V para gas condensado y aceite volátil.	16
Figura 3. Curva de volumen de líquido condensado en función de la Presión.....	17
Figura 4. Representación esquemática del experimento de Agotamiento a Volumen Constante para gases condensados.	19
Figura 5. Curvas de condensación de líquido para un gas condensado obtenidas de las pruebas EMC y AVC.....	20
Figura 6. Esquema de realización de una prueba PVT para gas Condensado.	26
Figura 7. Esquema de la simulación de una prueba AVC.....	34
Figura 8. Esquema de un Flash (P,T).	35
Figura 9. Esquema para el cálculo del comportamiento de fases por medio de una Ecuación de Estado.....	38
Figura 10. Condiciones para la existencia de la solución de $g(\epsilon)$ en $[0,1]$	73
Figura 11. Método de Rachford & Rice para encontrar el valor de β	76
Figura 12. Diagrama de flujo del proceso para encontrar la β en un estado determinado utilizando el método de Rachford & Rice.....	76
Figura 13. Diagrama de flujo del proceso de ajuste de EdE de Peng - Robinson.	104

Figura 14. Diagrama del proceso de búsqueda de los CIB que ajustan la Presión de saturación.	111
Figura 15. Diagrama del proceso de regresión de variables.	114
Figura 16. Estructura de una red neuronal.	118
Figura 17. Pesos sinápticos de una neurona.	119
Figura 18. Neurona, bias y su peso sináptico.	120
Figura 19. Modelo Simple de tres neuronas y una neurona bias.....	122
Figura 20. Intensificación del peso sináptico entre dos neuronas activas..	124
Figura 21. Modelo del Perceptrón.	128
Figura 22. Plano de activación.	129
Figura 23. Arquitectura de una red neuronal multicapa.....	130
Figura 24. Tipos de redes neuronales como clasificadores y según su tipo de aprendizaje.....	130
Figura 25. Arquitectura general de una red de retropropagación.	133
Figura 26. Arquitectura de una BPN de tres capas.	136
Figura 27. Superficie hipotética del espacio de pesos.	139
Figura 28. Función sigmoideal típica de forma exponencial.	142
Figura 29. Sección transversal de una superficie de error hipotética dentro del espacio de pesos.....	152
Figura 30. Ciclo de vida de un proyecto utilizando neurocomputación.....	154
Figura 31. Esquema de solución de las ecuaciones del comportamiento de fases.....	159
Figura 32. Esquema de solución de las ecuaciones del comportamiento de fases utilizando redes neuronales.....	160
Figura 33. Esquema de la metodología propuesta para el ajuste de parámetros.....	161
Figura 34. Estructura de entrenamiento.	164
Figura 35. Estructura de predicción.	165
Figura 36. Archivo de entrada de datos para el software ANNBack.	167
Figura 37. Entrenamiento de redes con el software ANNBack.	169

Figura 38. Entrenamiento de redes con el software ANNBack.	170
Figura 39. Dirección del recorrido para guardar los pesos de cada conexión luego del entrenamiento.....	171
Figura 40. Evaluación de la red entrenada con el software ANNBack.....	172
Figura 41. Archivo de datos para la evaluación de la red entrenada por el software ANNBack.	173
Figura 42. Esquema de solución de las ecuaciones del comportamiento de fases.....	176
Figura 43. Comportamiento de factor acéntrico con respecto al número de carbonos.....	180
Figura 44. Comportamiento del Coeficiente de Interacción Binaria con respecto al número de carbonos.	181
Figura 45. Archivo de entrada de datos para el software ANNBack.	192
Figura 46. Entrenamiento de redes con el software ANNBack.	194
Figura 47. Archivo de la estructura de la red entrenada por el software ANNBack.	195
Figura 48. Dirección del recorrido para guardar los pesos de cada conexión luego del entrenamiento.....	196
Figura 49. Evaluación de la red entrenada con el software ANNBack.....	197
Figura 50. Archivo de datos para la evaluación de la red entrenada por el software ANNBack.	198

LISTA DE ANEXOS

pág.

ANEXO A. ECUACIONES Y PROCEDIMIENTOS NUMÉRICOS	
UTILIZADOS EN EL SOFTWARE PvtLabRNA	218
ANEXO B. AJUSTE DE LA ECUACIÓN DE PENG-ROBINSON A DATOS	
PVT UTILIZANDO ALGORITMOS GENÉTICOS.....	219

TITULO: HERRAMIENTA SOFTWARE PARA EL AJUSTE DE LA ECUACIÓN DE ESTADO DE PENG-ROBINSON A DATOS EXPERIMENTALES PVT DE YACIMIENTOS CERCANOS AL PUNTO CRÍTICO UTILIZANDO REDES NEURONALES.*

AUTOR: FERNANDO ENRIQUE CALVETE GONZÁLEZ**

PALABRAS CLAVES: Ajuste, ecuación de estado, datos PVT, redes neuronales, simulación composicional.

RESUMEN

Los yacimientos de fluidos composicionales además de ser muy valiosos presentan una gran complejidad e inestabilidad. En ingeniería de petróleos, se han utilizado ecuaciones de estado para predecir y estimar el comportamiento de las fases presentes en ellos, constituyéndose en la base fundamental de la simulación composicional. Sin embargo, su utilización es limitada porque pierden validez en diferentes rangos de temperatura y presión.

Por otro lado, se han utilizado algoritmos de análisis estadístico, basados en el método de mínimos cuadrados, con el fin de ajustar los datos predictivos de las ecuaciones de estado a los datos experimentales PVT y de esta manera obtener los parámetros que permitan utilizarla a diferentes condiciones de temperatura y presión. Estas metodologías, tiene gran aplicación y una alta precisión, pero ofrecen una gran desventaja: el ajuste estadístico fuerza a cambiar los parámetros de la ecuación de estado sin tener en cuenta que estos reflejan características físico-químicas de los compuestos que hacen parte de la mezcla de hidrocarburos.

El presente trabajo tiene como fin, presentar una metodología y una herramienta software, basada en redes neuronales, que permite generar los parámetros de la ecuación de estado que reproducen con exactitud los datos PVT de laboratorio; permitiendo posteriormente predecir el comportamiento de fases de los fluidos a condiciones de yacimiento.

Esta nueva metodología ajusta los Coeficientes de Interacción Binaria – CIB —, los cuales son parte fundamental de la ecuación de estado y reflejan las interrelaciones moleculares de los compuestos químicos de los fluidos. Las redes neuronales que se utilizan no permiten que los cambios de los parámetros sean forzados sino que lleguen a valores tales que continúen representando las características físicas de los compuestos y permitan predicciones del orden del 99% de precisión como se muestra en los respectivos análisis de resultados.

* Trabajo de grado.

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería de sistemas. Zuly Himelda Calderón Carrillo.

TITULO: SOFTWARE TOOL FOR TUNNING THE PENG ROBINSON EQUATION OF STATE TO EXPERIMENTAL PVT DATA FOR NEAR CRITICAL POINT RESERVOIRS, USING ARTIFICIAL NEURAL NETWORKS.

AUTOR: FERNANDO ENRIQUE CALVETE GONZÁLEZ**

PALABRAS CLAVES: Tuning, Equation of state, PVT Data, Neural Networks, Compositional simulation.

ABSTRACT

Added to their extreme value, the reservoirs of compositional fluids show a very high complexity and instability. In petroleum engineering, equations of state have been used to predict and determine the phase behaviour in such systems, becoming easily the fundamental tool for compositional simulation. However, their use is limited because they loss accuracy in some pressure and temperature ranges.

On the other hand, algorithms provided by statistical analysis have been used, based on the minimum square method, trying to tune the predictive data of the equations of state to experimental data PVT by achieving the necessary parameters than allow to implement them at the different pressure and temperature conditions. These methodologies have an extensive application and a very high precision, but they offer a main disadvantage: the statistical tuning in certain way forces to change the equation's parameters, with out taking to account that these reflect the physical-chemical characteristics of the compounds constituents of the hydrocarbon mixture.

The present work has the main purpose, show up a methodology and a software tool, based in artificial neural networks, that allow to generate the parameters of the equation of state that reproduce exactly the laboratory data PVT; in this way the phase behaviour of reservoir fluids can be predicted more accurately.

This new methodology tunes the Binary interaction coefficients --CIB--, which are the fundamental part of the equation of state, so that they reflect the molecular interrelationships between the chemical compounds of the fluids. Artificial neural networks do not allow the forced changes appear, they simply let them achieve to a certain values that still represent the physical characteristics of the compounds making accurately of the 99% as shown in the respective results analysis.

* Trabajo de grado.

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería de sistemas. Zuly Himelda Calderón Carrillo.

INTRODUCCION

La observación de los fenómenos del mundo que nos rodea es el primero de los pasos que se efectúan para así aspirar a comprender tales eventos de una manera íntegra. De lo anterior se desprende que la información experimental obtenida de fuentes confiables, controlando las variables que se deseen medir, es el punto de partida para la elaboración de modelos que reproduzcan esas situaciones y puedan extenderse a otras diferentes a las que comúnmente se puedan observar.

Una expresión matemática que pretenda reproducir situaciones relativamente complejas se sustenta en una teoría referente al fenómeno estudiado, pero es común que aún así no se alcancen los niveles de precisión requeridos para fines prácticos, haciéndose necesaria la inclusión de términos de naturaleza empírica -susceptibles de ser ajustados- con los que se logre la exactitud deseada. La inclusión de tales términos puede ocasionar errores significativos al tratar de reproducir situaciones diferentes para las cuales se adaptó el modelo.

En yacimientos de hidrocarburos donde la presión y la temperatura se encuentran en una región lo suficientemente cercana al punto crítico de la mezcla, pequeñas variaciones en la presión o en la composición -variables termodinámicas- repercuten en cambios significativos en el estado del sistema. En consecuencia las cantidades de los productos en superficie y las propiedades que determinan su calidad y precio – gravedad API, GOR – también se ven considerablemente afectadas.

La simulación de este tipo de yacimientos debe ser composicional, esto es, auxiliada por una expresión que determine las condiciones de equilibrio termodinámico entre las fases, como lo es una Ecuación de Estado –EdE-. No obstante, la expresión original de cualquier EdE es insuficiente para describir acertadamente el comportamiento de una mezcla de sustancias, por lo que se hace necesaria la modificación de ciertos parámetros de dicha ecuación o del fluido con el fin de representar con mayor exactitud los datos experimentales obtenidos de una muestra representativa del yacimiento.

Se requiere de una EdE para la estimación inicial de reservas así como la predicción de las cantidades de los fluidos recuperables bajo un esquema dado de producción, buscando determinar la estrategia más apropiada para la óptima explotación del yacimiento. Si la EdE no ha sido convenientemente ajustada –respetando los criterios termodinámicos en los cuales se sustenta– las estrategias seleccionadas no serán las adecuadas para la economía del proyecto. Por tal razón el uso de una EDE ajustada otorgará mayor confiabilidad en las predicciones y certeza en la toma de decisiones.

El objetivo principal de este proyecto de investigación fue desarrollar una herramienta software que permita llevar acabo el proceso de ajuste de los parámetros de la ecuación de estado de Peng-Robinson con respecto a la información PVT de laboratorio, haciéndolo de forma automática y termodinámicamente consistente aplicando la tecnología de las Redes Neuronales Artificiales -RNA-, con el fin de ofrecer una nueva forma de efectuar este tipo de cálculo solventando los problemas derivados de la complejidad y la heurística propia de los modelos de ajuste previos, basados en la termodinámica de los fluidos petrolíferos.

El presente trabajo resume los aspectos principales que se tuvieron en cuenta para llegar a la construcción de la metodología y la herramienta

software producto de ella. En los primeros cuatro capítulos se describe el comportamiento de fases de los fluidos de yacimientos petrolíferos, la matemática que rige las relaciones físicas de las muestras y las herramientas conceptuales de ingeniería que permiten simular el comportamiento de los fluidos en el laboratorio PVT. En el quinto capítulo se describe con detalle las metodologías tradicionales de ajuste utilizadas en la industria del petróleo con el fin de esclarecer el problema numérico que trabajan y de esta manera seleccionar las técnicas de inteligencia artificial que permita de una manera meta-heurística resolver el problema. En el sexto capítulo se muestran los fundamentos de las redes neuronales utilizadas para realizar la metodología de trabajo. El séptimo capítulo esboza la metodología conjunta de simulación de pruebas y redes neuronales que se implementó para resolver el problema de ajuste, y que se aplica en la herramienta desarrollada. En el capítulo octavo se presenta el análisis de resultados y el estudio comparativo donde se muestra numéricamente las diferencias con respecto a los cálculos tradicionales, así como las ventajas y desventajas de la metodología propuesta. Por último se muestra el manual del usuario donde se han plasmado las opciones más relevantes del software desarrollado y la forma como se puede utilizar en la simulación de las pruebas PVT.

1. MEDIDAS PVT DE LABORATORIO

Con el fin de estudiar las propiedades de los fluidos de yacimiento en el laboratorio, es necesario recombinar las mezclas de líquido y gas del separador, para ello se mezclan las dos muestras proporcionalmente a su relación gas líquido de producción (GOR).

Existen diferentes análisis de laboratorio que se pueden realizar a una muestra de fluido de yacimiento para determinar datos importantes que serán utilizados en una gran variedad de cálculos de ingeniería.

En este capítulo se describirán las pruebas que se realizan para simular el comportamiento de los gases condensados y los aceites volátiles, las cuales proporcionarán información requerida durante el desarrollo del yacimiento.

1.1. ANÁLISIS PRELIMINARES

El fluido que llega del pozo está separado en líquido y gas en el separador. Después del control de calidad, la muestra de gas está sujeta a análisis composicional en el cromatógrafo, posteriormente se calienta y se deja en equilibrio. De este análisis es posible determinar si la muestra está contaminada, con aire u otro compuesto.

La muestra de líquido se debe estabilizar a condiciones atmosféricas por un proceso instantáneo simple. Una pequeña porción de la muestra a alta

presión, se conecta a un separador instantáneo, convirtiéndola en un fluido en dos fases. Se determina el volumen de gas y se espera hasta que el líquido se deposite en fondo de la cámara flash para determinar su peso y densidad.

La relación gas - líquido de la muestra del separador a condiciones estándar, GOR_0 , puede ser calculada de la siguiente manera:

$$GOR_0 = \text{Volumen de gas} / \text{Volumen de aceite} \quad (1.1)$$

A volúmenes medidos a condiciones estándar.

La composición del líquido de separador se obtiene recombinando matemáticamente los análisis cromatográficos de gas de laboratorio y líquido del tanque.

Para realizar los análisis previamente mencionados, es necesario haber calculado la relación gas- aceite (GOR) con precisión.

1.1.1. Selección de la muestra. Este es el primer procedimiento realizado en el laboratorio al recibir la muestra presurizada. Antes de comenzar un análisis, es importante evaluar las muestras disponibles y escoger la más representativa de los fluidos de yacimiento. Se escogen las muestras tomadas durante las condiciones más estables.

La presión de burbuja del líquido del separador se determina en el laboratorio a la temperatura de separación. La muestra se considera representativa si la presión de burbuja determinada en el laboratorio es igual a la presión del separador de prueba en el momento del muestreo.

Los recipientes con las muestras se conectan a una bomba de mercurio y a un medidor de presión y se registra la presión de apertura. En el caso de muestras de líquido, la bomba se utiliza para elevar la presión en la botella con el fin de realizar la gráfica de presión contra volumen inyectado a temperatura ambiente. Un cambio repentino en la pendiente de la gráfica indica el punto de burbuja de la muestra de líquido del separador.

Como regla del dedo gordo, el punto de burbuja de una muestra de líquido debe disminuir entre 3 y 4 psi por grado Fahrenheit de diferencia entre las condiciones de separador y las condiciones ambiente. Esta regla se aplica también para muestras de fondo.

La presión de apertura en la fase gas correspondiente, puede revisarse y debe ser igual a la presión del separador en el momento del muestreo. Si estas dos presiones concuerdan el muestro puede realizarse.

1.1.2. Análisis composicional de fluidos. La técnica estándar usada para determinar composiciones de mezclas de hidrocarburos involucra cromatografía y destilación. Las muestras de gas se analizan únicamente utilizando cromatografía. El equipo separa los componentes hidrocarburos llevando una pequeña cantidad de muestra a una columna absorbente a diferentes tiempos, dependiendo de su afinidad con el material de la columna. El detector responde a la masa de los componentes presentes y se calibra usando cantidades de mezclas de composición conocida. Con el fin de corroborar toda la información se requiere realizar corridas a través de diferentes medios y con varias muestras.

Los líquidos también se analizan usando las técnicas de cromatografía gaseosa, pero son separados en fracciones de destilación. Para líquidos recolectados a una presión del separador alta, se realiza una separación

flash (bajo condiciones controladas) y el gas liberado se analiza por separado. En estas circunstancias se requieren dos recombinaciones para calcular la composición de los fluidos de yacimiento. Aunque algunas veces un análisis hasta el C_{11} es suficiente, en la mayoría de los casos, un análisis extendido es muy útil para caracterizar la fracción pesada para propósitos de agrupamiento usado en ecuaciones de estado. El análisis extendido se realiza usando la técnica de destilación y midiendo la cantidad de destilado recolectado entre los rangos de punto de destilación (True Boiling Point o técnica TBP) de hidrocarburos parafínicos normales.

Otra técnica comúnmente usada es la destilación simulada por cromatografía de gas, ya que es considerablemente más económica y consume menos tiempo. Sin embargo la cantidad y calidad de datos medidos se reduce porque las fracciones nunca están disponibles para estudiarlas por separado y entonces la mayoría de propiedades asignadas a las fracciones se deben calcular.

El método de espectrometría de masas puede ser combinado con la destilación simulada para examinar la distribución de parafinas, naftenos y aromáticos en las fracciones de puntos de ebullición diferentes. Esta es una técnica especializada que no se aplica frecuentemente.

1.1.2.1. Precisión de las mediciones composicionales. Este aspecto tiene gran importancia en la estimación de la validez de datos, ya que las diferencias reportadas en la composición de un grupo de muestras no necesariamente significan que estas no sean válidas. En general, la precisión en la medición de la concentración de un componente en particular es función de la concentración y naturaleza del mismo.

Eyton¹ desarrolló una relación para estimar repetibilidad de las medidas composicionales, la cual se expresa según la ecuación 1.2:

$$R_{xi} = 0.07 \times i^{0.43} \quad (1.2)$$

Donde R es la repetibilidad de la medición generada por diferentes operadores utilizando el mismo equipo.

$$R = 2.8 \mathbf{s} \quad (1.3)$$

Siendo **s** el estimativo de imparcialidad de la desviación estándar de la población. Esto significa que R no representa un error sistemático ni una parcialidad en la medición sino la influencia del azar en la misma.

Por lo tanto existe un nivel de confiabilidad dado por:

$$\text{Límite inferior} = (\text{Menor resultado}) * \frac{R}{1.4} \quad (1.4)$$

$$\text{Límite superior} = (\text{Mayor resultado}) * \frac{R}{1.4} \quad (1.5)$$

Como regla general, si los resultados de dos laboratorios difieren más de dos veces la repetibilidad de los estimativos en un laboratorio, es recomendable que el experimento se repita.

Para la repetibilidad de un experimento de medición de composiciones (R_{xi}) se encontró un límite de repetibilidad cercano al 0.5% molar cuando el porcentaje de un componente alcanza el 100% y de 0.01 % molar cuando la concentración del componente es 0.01% molar.

1.1.3. Validación de las muestras de gas y aceite a condiciones del separador. Habiendo obtenido el análisis composicional de las muestras del separador, es posible revisar si los valores de K calculados son compatibles a las condiciones de separador reportadas. La técnica de Hoffman, Crump y Hocott² es la más efectiva en este aspecto.

Este es un método gráfico para revisar los datos composicionales del equilibrio de fases. Realmente se debió a Buckley y Lighfoot¹ pero no se usó ampliamente hasta la publicación de Hoffman et al².

El método es muy preciso en el rango de presiones y temperaturas encontradas en el separador y por lo tanto es ideal para revisar la compatibilidad de las composiciones reportadas de las muestras.

De las composiciones de gas y líquido se pueden calcular los valores de K para cada componente el cual es una relación entre la fracción molar de un componente en la fase vapor y su correspondiente fracción en la fase líquida. En este procedimiento los datos se suavizan realizando una gráfica para cada componente del logaritmo del producto KP (la constante de equilibrio por la presión absoluta) contra una función $b(1/T_b - 1/T)$, donde b es una constante característica de cada hidrocarburo, T_b es su punto de ebullición en oR y T es la temperatura del separador en oR. Se ha encontrado empíricamente que para el equilibrio entre un gas y un líquido a cualquier presión, esta gráfica será una línea recta.

La gráfica resultante debe mostrar todos los puntos que representan las composiciones individuales que están fuera de la línea recta.

Los valores de b requeridos pueden ser calculados a partir de la presión de vapor usando la ecuación 1.6.

$$\log \frac{P_2}{P_1} = b \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \quad (1.6)$$

La experiencia ha indicado que los valores de b calculados a partir de presiones y temperaturas críticas y de los puntos de ebullición normal son satisfactorios para propósitos de correlación para la mayoría de yacimientos de gas condensado. De esta forma, b puede ser calculada usando la ecuación 1.7.

$$b = \frac{(\log P_c - \log 14.7)}{\left(\frac{1}{T_b} - \frac{1}{T} \right)} \quad (1.7)$$

Los valores para altos pesos moleculares fueron determinados de una curva suavizada de b contra el número de carbonos de la molécula. Los valores de puntos de ebullición y propiedades críticas se pueden encontrar en la referencia mencionada.

Esta no es una relación arbitraria, puede deducirse de la termodinámica de equilibrio usando la ecuación de Clausius-Clapeyron y suposiciones de comportamiento de gases perfectos para la fase vapor. Sin embargo, dependiendo de estas suposiciones para la validez de esta relación, esta comienza a fallar a altas presiones. Sin embargo, desviaciones significativas de la línea recta indican problemas de equilibrio de las fases o imprecisiones en el análisis. Desviaciones menores pueden corregirse, moviendo los puntos a la línea recta.

1.1.4. Recombinación de muestras. Habiendo obtenido y validado el análisis composicional del gas y líquido del separador, la composición del

fluido de yacimiento puede ser calculada conociendo la relación gas - aceite en producción. Pero antes de esto, las mediciones de tasas de flujo reportadas necesitan ser corregidas ya que ellas se basan en cálculos que necesitan valores para Z y gravedad del gas. La gravedad se mide en campo, pero sin la clase de precisión que puede obtenerse recalculándola a partir de la composición conocida.

Por otra parte, el factor Z se basa en correlaciones que están en función de la gravedad del gas. La corrección, por lo tanto, recalcula la tasa de flujo basada en las composiciones de la gravedad del gas y del factor Z para obtener de nuevo la relación gas - aceite en el separador.

La recombinación consiste en volver a determinar el número de moles de cada componente en cada fase para el GOR establecido y agregarlas a cada composición de fase calculada previamente. Finalmente renormalizar y así obtener un 100%.

1.2. EXPERIMENTOS PRESIÓN - VOLUMEN

Los dos experimentos fundamentales realizados en sistemas de gas condensado y aceite volátil son el Agotamiento a Volumen Constante (AVC) y la Expansión a Masa - Composición - Constante (EMC). En los dos se emplea una celda PVT, la cual es un recipiente cuyo volumen interno es conocido para un amplio rango de presiones y temperaturas y que además puede mantenerse a temperatura constante. Esta celda está equipada con una ventana resistente a altas presiones a través de la cual puede verse y medirse cualquier cambio de fase en la muestra analizada. Los cambios de presión y volumen se realizan introduciendo o retirando mercurio

directamente a la celda o usando pistones flotantes. Inicialmente la celda se carga con los volúmenes de gas y líquido necesarios para formar de nuevo una muestra representativa. Posteriormente se deja en equilibrio a presión y temperatura de yacimiento, algunas de ellas tienen mecanismos de agitación para lograr alcanzarlo más rápidamente.

La prueba de Liberación Diferencial también es parte de la experimentación P-V pero no es tan convencional como las pruebas AVC y EMC y se efectúa a crudos livianos que no son tan volátiles.

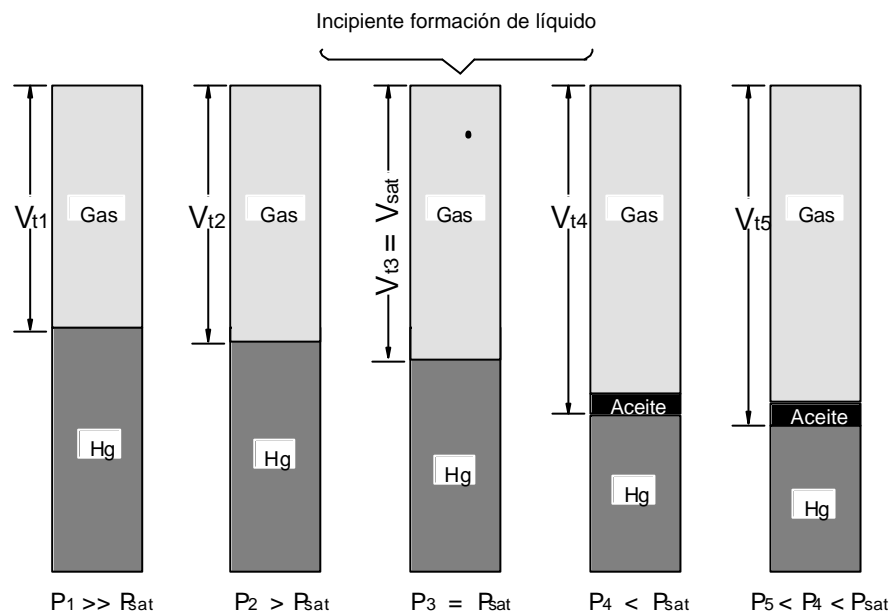
1.2.1. Expansión a Masa Constante (EMC)⁵. Esta prueba se conoce también como Expansión a Composición Constante o vaporización flash y es la que inicia el proceso de pruebas PVT para un fluido de yacimiento.

Esta prueba se realiza para cualquier tipo de fluido y se utiliza para medir el volumen total de fluido y su compresibilidad en un amplio rango de presiones desde la presión inicial de yacimiento hasta valores por debajo de la del separador. Para aceites negros y volátiles también se usa en la determinación de la presión de saturación a condiciones de yacimiento.

A fin de alcanzar un fluido de yacimiento representativo, la muestra se transfiere a la celda en la misma relación gas - líquido medida durante el muestreo. Una vez hecho esto, el sistema en la celda PVT, se lleva a la temperatura de yacimiento y se comprime hasta alcanzar una presión superior a la de yacimiento medida. En este punto, la celda se deja (usualmente durante la noche) para lograr equilibrio, posteriormente el sistema se revisa para ver si está todo en una fase y se registra el volumen total. En este momento puede comenzar el proceso de expansión a masa constante que se presenta en la figura 1.

La presión se reduce paso a paso. Se observa en la celda visual el sistema y se registra el volumen en cada cambio de presión, según las lecturas de la bomba manual.

Figura 1. Representación esquemática de una Expansión a Masa Constante para un gas condensado.

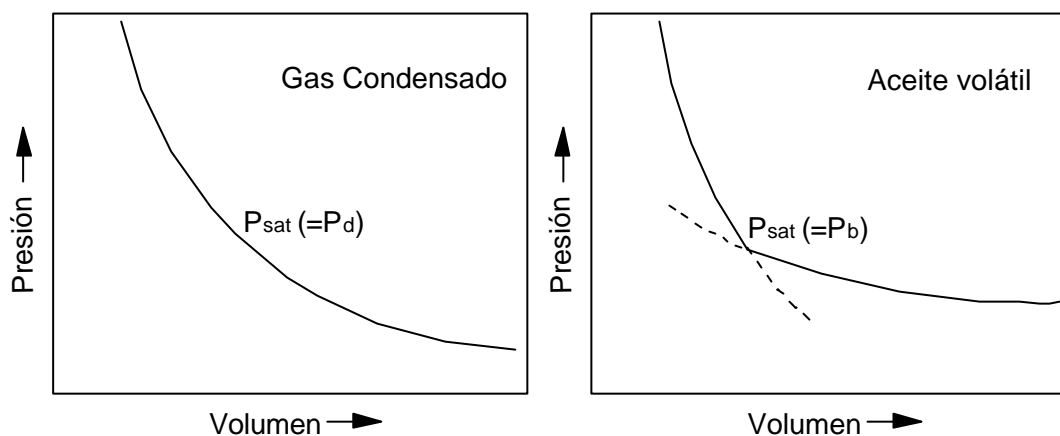


Fuente: AHMED, T. H. "Hydrocarbons Phase Behavior".⁵

En el caso de fluidos volátiles cercanos al punto crítico, puede ser difícil determinar si la presión de saturación es un punto de rocío o de burbuja, porque las dos fases son cercanamente idénticas. Los cambios de fases son rápidos y exhiben grandes variaciones en su comportamiento para pequeños cambios de presión y temperatura. Una vez la presión de saturación del sistema se ha encontrado, y el estudio a masa constante se ha realizado, el programa de trabajo continúa de la misma manera que para un gas condensado.

Para aceite volátil lejano al punto crítico la transición desde el sistema en una fase al sistema en dos fases es fácil de reconocer en la gráfica, pero en condensados la separación de la fase líquida no es muy notoria, por lo tanto debe hacerse visualmente. La figura 2, muestra las curvas para condensados y volátiles.

Figura 2. Comportamiento P-V para gas condensado y aceite volátil.



Fuente: AHMED, T. H. "Hydrocarbons Phase Behavior".⁵

Debido a que la presión del punto de rocío puede ser un poco difícil de reconocer en gases condensados, el agotamiento continúa por debajo de esta presión. En este momento se debe determinar el volumen de líquido condensado en cada paso, con lo cual obtenemos la curva de condensación de líquido que se observa en la figura 3 sobre la cual se determina el punto de rocío extrapolando para un volumen de líquido de cero.

En la región de una fase las curvas P-V también pueden ser usadas para calcular³:

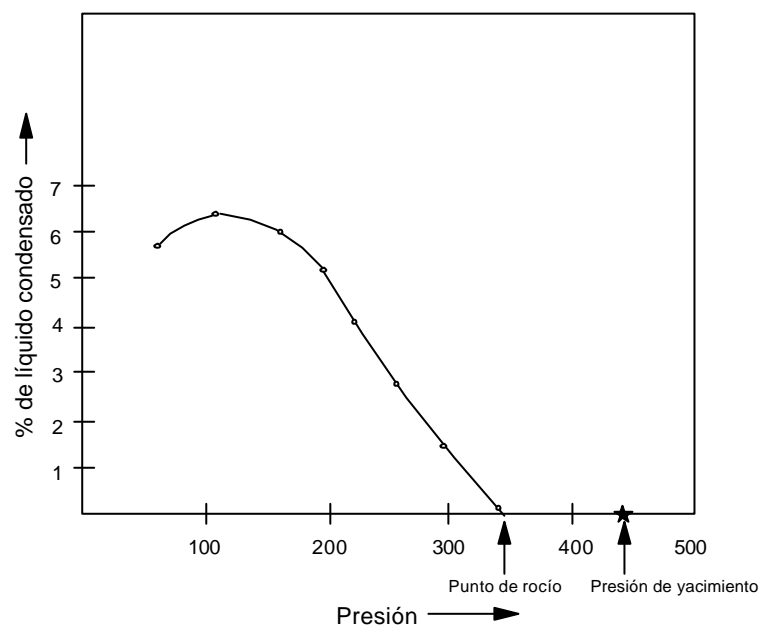
- La variación de la densidad con la presión.

- La compresibilidad isotérmica.
- La expansión térmica si el volumen de muestra se registra a más temperaturas.

El número de reducciones y el rango de presión que se va a cubrir, generalmente se decide consultando con el ingeniero que solicita el estudio.

La expansión a masa constante puede realizarse además a otras temperaturas para ayudar a modelar las condiciones de producción para procesos futuros.

Figura 3. Curva de volumen de líquido condensado en función de la Presión.



Fuente: CORE LABORATORIES. "Basic Phase Behavior Course".³

Para muestras de aceite, la expansión volumétrica por debajo del punto de burbuja es mayor que por encima de este. Esto permite determinar

fácilmente dicho punto en la gráfica, el cual será aquel en donde la pendiente de la curva cambia. El gran cambio en expansión se debe a la presencia de vapor, el cual es mucho más compresible que la fase líquida.

Es importante anotar que el volumen que se registra en las gráficas es un volumen relativo al volumen medido en el punto de saturación, aunque algunos laboratorios lo definen como la relación del volumen de líquido condensado sobre el volumen de fluido total³.

1.2.2. Agotamiento a volumen constante (AVC)⁵. Este experimento se realiza generalmente para aceites volátiles y gases condensados para simular el comportamiento del fluido en el yacimiento durante la producción. Se diferencia de la liberación diferencial (prueba que se hace a aceites negros) en que no todo el gas se remueve en cada etapa de presión, de esta forma, se puede modelar una formación de volumen constante que está produciendo gas rico, pero reteniendo la fase líquida en equilibrio.

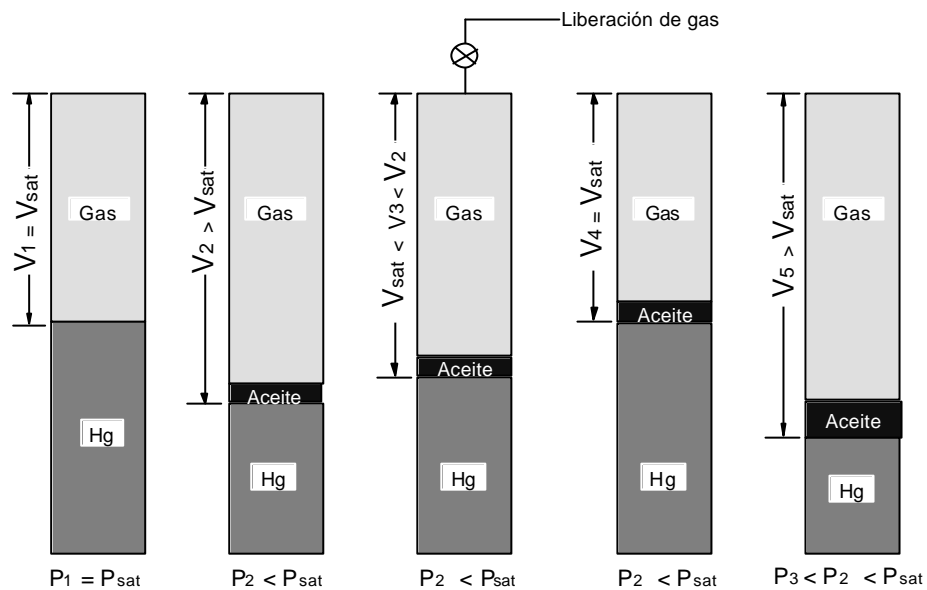
Esta prueba se realiza para aceites volátiles debido a los grandes volúmenes de gas liberados a condiciones de yacimiento cuando la presión cae por debajo del punto de burbuja y al alto encogimiento de la fase líquida cuando esto ocurre.

La figura 4 muestra un diagrama de este procedimiento.

De nuevo el experimento comienza con una muestra de fluido en una fase, de volumen conocido, a presión de saturación y temperatura de yacimiento. De la misma manera que el experimento anterior, la presión se reduce paso a paso, resultando una expansión del volumen en cada uno, luego de alcanzar el equilibrio, se retira fluido de la celda hasta alcanzar el volumen

inicial. En algún momento durante la reducción, el fluido pasa por el punto de burbuja (volátiles) o por el punto de rocío (condensados).

Figura 4. Representación esquemática del experimento de Agotamiento a Volumen Constante para gases condensados.



Fuente: AHMED, T. H. "Hydrocarbons Phase Behavior".⁵

El fluido removido en cada etapa, se lleva a condiciones estándar y se analiza para determinar la composición y el peso molecular de la fracción pesada - C_{7+} -. Generalmente se realizan entre cinco y diez reducciones hasta alcanzar la presión atmosférica; después de la etapa final, el líquido se remueve de la celda para analizarlo.

Los datos que son determinados en la prueba son los volúmenes de gas y líquido a presión y temperatura, la cantidad de gas producida acumulada, el factor Z del gas y la composición total de la corriente del pozo. Todos los

datos volumétricos se reportan como volúmenes relativos al punto de rocío o de burbuja según sea el caso.

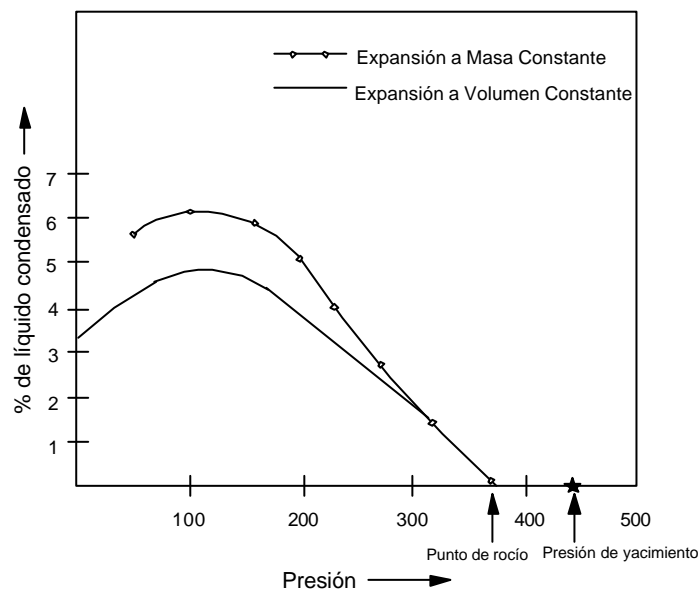
Al final del análisis se realiza un balance molar para saber si se ha perdido fluido durante el experimento. El balance puede escribirse como:

$$\# \text{ moles cargadas a la celda} = \# \text{ moles producidas de la celda} \quad (1.8)$$

Se admite un error máximo de 1% en la medición molar con el fin de mantener la consistencia de los datos.

La figura 5 muestra la diferencia entre las curvas de condensación de líquido para una expansión a masa constante y un agotamiento a volumen constante.

Figura 5. Curvas de condensación de líquido para un gas condensado obtenidas de las pruebas EMC y AVC.



Fuente: AHMED, T. H. "Hydrocarbons Phase Behavior".⁵

1.2.3. Liberación Diferencial (LD)⁵. La prueba de liberación diferencial se lleva a cabo cargando una muestra de aceite del yacimiento en una celda PVT. La muestra de líquido se encuentra a temperatura del yacimiento y presión de punto de burbuja. Con esta prueba se puede obtener valores tales como:

1. Factor volumétrico de formación del aceite, B_{od}
2. Relación gas-aceite en solución, R_{sd}
3. Factor de merma del aceite, S_{od}
4. La densidad del fluido de yacimiento a diferentes presiones
5. Factor de compresibilidad del gas a presión y temperatura dadas.

El subíndice d se refiere a la prueba diferencial.

En esta prueba la presión se reduce en etapas que usualmente están entre diez y quince niveles; en cada nivel una vez alcanzado el equilibrio, se determina el volumen de aceite y se remueve el gas de la celda a presión constante y su volumen se mide a condiciones estándar. Este procedimiento continúa hasta alcanzar la presión atmosférica donde se mide y convierte el volumen de aceite residual a un volumen a 60°F, V_{cs} .

1.3. PRUEBAS DE SEPARACIÓN MULTI-ETAPAS.⁵

Las pruebas de separador o pruebas de separación flash o multi-etapas se realizan para determinar los cambios en el comportamiento volumétrico del fluido de yacimiento a medida que este pasa a través del separador o separadores hasta el tanque de superficie, y tiene como objetivo suministrar información esencial en la determinación de las condiciones de separación

óptimas de superficie y de esta manera maximizar la producción de aceite en el tanque. El comportamiento volumétrico de este proceso depende en mayor grado de las condiciones de operación (presión y temperatura) de las instalaciones de separación en superficie.

Se debe realizar un determinado número de pruebas que permitan encontrar la presión óptima a la que se debe trabajar el separador la cual debe ser aquella a la que se obtenga un máximo valor de gravedad API en el tanque, un mínimo factor volumétrico de formación del aceite y una menor cantidad de gas recuperado.

La información obtenida con estas pruebas es la siguiente:

1. Presión y temperatura óptimas de separación.
2. Composición de las fases líquida y gas separadas.
3. Factor volumétrico de formación del aceite.
4. Relación gas-aceite de producción.
5. Gravedad API del aceite en el tanque de superficie.

Las pruebas de separador rara vez se realizan en sistemas de gases condensados, debido a la dificultad para obtener una cantidad suficiente de fluido de yacimiento y obtener un volumen apreciable de líquido experimental.

1.4. PRUEBAS DE INYECCIÓN DE GAS.⁵

Los experimentos de inyección de gas son usualmente llevados a cabo bajo dos tipos de condiciones. En primer lugar, un volumen variable de gas de

inyección se mezcla con un fluido saturado y luego por cambio de presión esta mezcla se lleva a la condición de saturación (una fase). Alternativamente, se efectúa un experimento de volumen constante en el cual un volumen de gas se adiciona al fluido que se encuentra por debajo de su condición de saturación, se sigue con una remoción de gas enriquecido hasta que se logre el volumen original de fluido. Estos dos experimentos son conocidos como pruebas de hinchamiento y vaporización, respectivamente.

1.4.1. Prueba de hinchamiento.⁵ En esta prueba el gas inyectado en un fluido bajo-saturado puede entrar en solución al “hinchar” la fase aceite. También puede suceder que la inyección de compuestos livianos enriquezca los componentes más pesados incrementando de esta manera la presión de saturación.

Esta prueba se puede efectuar también a gases condensados pero la caída de la presión de saturación se ve retardada debido a la presencia de compuestos livianos; la simulación de esta prueba consiste en encontrar la presión de saturación y definir un volumen de control (V_{sat}), posteriormente, se adiciona un gas de composición definida. A esta nueva mezcla se le encuentra su presión y volumen de saturación para definir la relación existente entre estos volúmenes, que se denomina factor de hinchamiento.

1.4.2. Prueba de vaporización.⁵ Este experimento es similar a la prueba de hinchamiento excepto que se efectúa a volumen y presión constante. Un volumen de fluido a una presión y temperatura usualmente por debajo de la presión de saturación y en dos fases es contactado por medio de la inyección de un gas. En cada contacto un volumen de gas o aceite enriquecido es removido de la celda y se devuelve el sistema a su volumen original, la composición de la muestra removida es medida en cada remoción.

Este experimento muestra el efecto de la vaporización de los componentes intermedios y livianos de la fase líquida del fluido de yacimiento a la fase vapor por gas inyectado.

1.5. PRUEBAS DE MISCIBILIDAD⁵

Esta prueba permite encontrar las relaciones de composición y presiones a las cuales debe ser sometido el fluido del yacimiento para que pueda ser miscible con un fluido de inyección. Generalmente se pueden efectuar tres tipos de pruebas de miscibilidad:

- Prueba de Mínima Presión de miscibilidad. En esta prueba se busca encontrar la presión mínima que se debe tener en el yacimiento para lograr una miscibilidad completa.
- Prueba de Primera Presión de Miscibilidad. Se busca encontrar la presión para que se logre miscibilidad al primer contacto.
- Prueba de Miscibilidad de Múltiples Contactos. En esta prueba se verifica el comportamiento del fluido de yacimiento al entrar en contacto con el fluido de inyección. El desplazamiento va más allá de la zona de contacto y se verifica como cambia la composición total de fluido a medida que se va barriendo el espacio poroso ocupado por el fluido original. En la prueba de miscibilidad de múltiples contactos se pueden ver dos fenómenos bien definidos:
 - Vaporización. Este fenómeno ocurre cuando el gas contacta el aceite original de yacimiento, esto permite el enriquecimiento de los

componentes intermedios del gas. Aquí puede ocurrir una conversión de los estados termodinámicos y crearse una sola fase y se alcanza la miscibilidad completa.

- Condensación. Se considera en este tipo de comportamiento que el fluido original es contactado por gas fresco de inyección. En este proceso los componentes intermedios se transfieren del gas al aceite, resultando una combinación completa, formándose una sola fase y se alcanza miscibilidad completa.

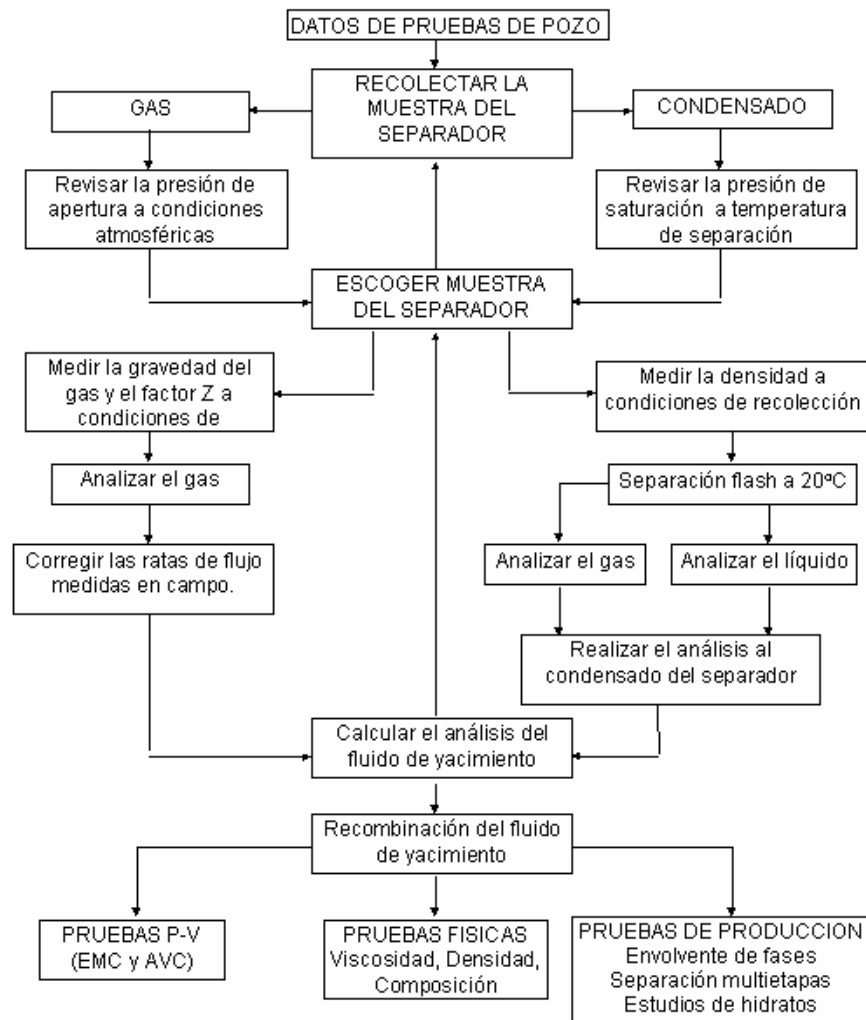
1.6. DIAGRAMA DE FLUJO DE UNA PRUEBA PARA GAS CONDENSADO

Conociendo los procedimientos a que debe ser sometida la muestra, estos se deben efectuar siguiendo un camino típico en el laboratorio.

El punto inicial es siempre recolectar la mayor cantidad de datos del pozo a muestrear y del yacimiento. Posteriormente se puede proceder a la recolección de la muestra según el procedimiento adecuado. Esta debe validarse antes de iniciar las pruebas. Después de ejecutar el estudio de laboratorio el paso final será realizar el análisis solicitado y presentar un reporte.

Este procedimiento puede observarse claramente en la figura 6.

Figura 6. Esquema de realización de una prueba PVT para gas Condensado.



Fuente: SUNBURY B. P. RESEARCH CENTER. Conferencias sobre Comportamiento de Fases.⁴

1.7. SOLICITUD DE UN ANÁLISIS PVT

La empresa que requiere un estudio PVT debe incluir en su solicitud un resumen de la historia de producción del pozo, de los registros que se han

corrido y de las pruebas de presión y producción realizadas. Además se debe especificar que análisis se requieren para las muestras del estudio y cuantas.

Finalmente se debe anexar la documentación de las pruebas realizadas entre las cuales se encuentran:

- Temperatura y presión de yacimiento.
- Los resultados de las pruebas preliminares.
- Condiciones de todos los periodos de flujo preliminares al muestreo.
- Descripción del periodo de flujo durante el cual se tomó la muestra.
- Hojas de muestreo, en donde se presenta la lista de las muestras tomadas.
- Mapas del proyecto.

A continuación se presentan los requerimientos técnicos para un estudio de fluidos, desde la validación hasta los análisis mismos de laboratorio:

1. Validación de muestras.
 - Determinar una muestra válida de aceite, midiendo la presión de saturación de las muestras a temperatura del separador.
 - Chequear la presión de apertura de las muestras de gas del separador a temperatura del separador.
 - Medir el punto de burbuja de las muestras de cabeza de pozo
 - Reportar los datos por télex y esperar notificación de cómo proceder.
2. Análisis del fluido de yacimiento.
 - Revisar la compresibilidad del gas a presión y temperatura de la muestra.

- Realizar el estudio hasta C_{10} incluyendo CO_2 , N_2 , Benceno y Tolueno.
- Medir la presión de saturación de las muestras del separador válidas a temperaturas de $20^\circ F$ por encima y por debajo de la temperatura del separador.
- Realizar una prueba flash al líquido de muestra desde condiciones del separador hasta condiciones estándar:
 - a. Reportar la cantidad de gas separado y el análisis de líquido residual hasta C_{20} . Reportar el peso molecular de la fracción de C_{20+} .
 - b. Dar la curva de punto de ebullición (TBP) para el aceite residual.
 - c. Reportar la densidad del líquido a condiciones del tanque y la del gas a condiciones estándar.
 - d. Calcular el factor de merma (Factor volumétrico en el separador)
 - e. Calcular la composición del fluido del yacimiento.
 - f. Calcular y reportar el factor volumétrico de formación y el GOR.
- Obtener el análisis del fluido de yacimiento matemáticamente, recombinando las muestras de gas y aceite. Corregir la relación gas-aceite, usando el factor Z medido previamente.
- Realizar una prueba de separación en una etapa para una muestra de cabeza de pozo desde condiciones de yacimiento:
 - a. Reportar los análisis de gas separado y líquido residual. Reportar peso molecular del C_{20+} .
 - b. Reportar la densidad del líquido en el stock tank y de gas a condiciones estándar.
 - c. Calcular la composición del fluido de yacimiento.
 - d. Realizar la curva de punto de ebullición (TBP) para aceites residuales.
 - e. Calcular y reportar el factor volumétrico de formación y el GOR.

3. Relaciones presión – volumen.
 - Se realizará a temperatura de yacimiento y a presiones ligeramente mayores que la presión de yacimiento hasta la presión del separador.
 - Obtener la relación de presión contra volumen relativo al volumen en el punto de burbuja.
 - Reportar la densidad del fluido en el punto de burbuja y en la región de una sola fase hasta presión de yacimiento.
 - Reportar la compresibilidad a temperatura de yacimiento y a un rango de presiones cercanas a la presión de yacimiento.
 - Reportar los datos de expansión térmica.
4. Viscosidad del aceite.
 - Medir la viscosidad para las temperaturas especificadas y sobre un rango de presiones desde la de yacimiento hasta la de burbuja.
5. Vaporización diferencial.
 - Se realizará a temperatura de yacimiento.
 - Reportar la presión, densidades de aceite, merma y factor volumétrico de formación.
 - Reportar las densidades de gas y los GOR, y la densidad del gas total liberado.
 - Analizar el gas liberado.
 - Analizar el gas condensado si se forman suficientes cantidades.
 - Medir la viscosidad del aceite.
6. Separación multietapas.
 - Realizarla a las siguientes condiciones:
Etapas 1: Presión de yacimiento y Temperatura de separador 1
Etapas 2: Presión < Presión de yacimiento y Temperatura de separador 2.

Etapas 3: Presión < Presión de yacimiento y 60 °F

Temperatura de separador 2 < Temperatura de separador 1

- En cada etapa medir densidad de aceite, densidad de gas, GOR, Factor volumétrico de formación y gas condensado formado. Reportar el FVF y la densidad de aceite en el tanque.
- Analizar el gas en cada etapa.
- Reportar el análisis de aceite residual.
- Analizar el gas condensado si se ha formado una cantidad suficiente.
- Medir la viscosidad del aceite en cada etapa.

2. MATEMÁTICA DEL COMPORTAMIENTO DE FASES

En este capítulo se hará una descripción matemática de la simulación de las pruebas PVT que se efectúan típicamente para yacimientos de gas condensado y aceite volátil, haciendo énfasis en los cálculos de equilibrio líquido-vapor que son el corazón del simulador. Este capítulo recopila la formulación numérica basada en la termodinámica del comportamiento de fases y que permite al simulador predecir las condiciones presentes en cada etapa de la prueba.

2.1. PROCESO DE AGOTAMIENTO EN YACIMIENTOS DE ACEITE VOLÁTIL CERCANO AL PUNTO CRÍTICO

Estos fluidos se conocen generalmente como aceites volátiles y se caracterizan por encontrarse en el yacimiento a una temperatura cercana a su valor crítico presentando así una merma que puede llegar a alcanzar en algunos casos el 45% del espacio poroso de hidrocarburos con una reducción de sólo 10 psi por debajo de su punto de burbuja. Este efecto crea una alta saturación de gas en el espacio poroso, el cual alcanza una alta movilidad inmediatamente después de alcanzar la presión de saturación. Un estudio de fluidos de yacimiento adecuado, debe habilitar al ingeniero de yacimientos a realizar un balance de material composicional, teniendo en cuenta tanto la producción de condensado retrógrado como de líquido del yacimiento.

2.2. PROCESO DE AGOTAMIENTO EN YACIMIENTOS DE GAS CONDENSADO

Un estudio de yacimiento completo para un gas condensado debe incluir composición de gas y líquido del separador y del fluido recombinado. Además debe comprender la medición del punto de rocío retrógrado, la compresibilidad del fluido por encima de este y los volúmenes de gas y líquido durante una expansión a masa constante.

Finalmente el estudio debe incluir una depleción simulada, la cual consta generalmente de una serie de expansiones y desplazamientos a presión constante de tal forma que el volumen de la celda es el mismo al final de cada desplazamiento; este procedimiento se menciona como agotamiento a volumen constante. El gas producido durante cada etapa se carga a un equipo analítico para determinar su volumen y composición.

Dentro de la roca que contiene gas, una caída de presión puede causar condensación del aceite originalmente vaporizado. Mientras la saturación del aceite en el yacimiento este por debajo de la saturación crítica, el aceite quedara inmóvil. En este punto en la vida del yacimiento, un proceso de agotamiento a volumen constante puede representar el comportamiento de depleción del yacimiento y la variación composicional. Mientras la Expansión a Masa Constante representará el comportamiento volumétrico del fluido antes de caer por debajo de su punto de saturación.

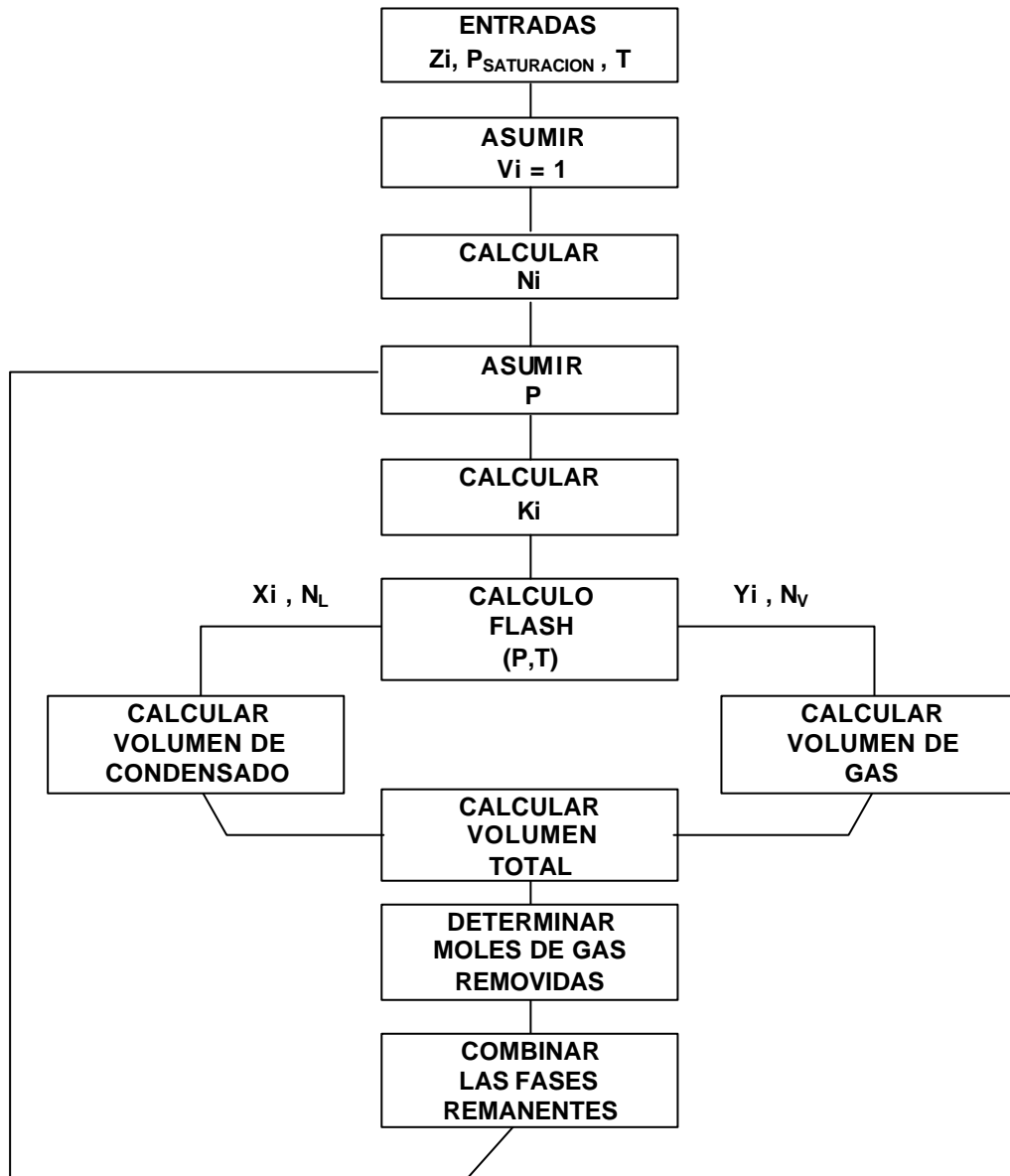
2.3. PRUEBAS DE LABORATORIO

La primera prueba de reducción de presión que se realiza a los fluidos de yacimiento es la Expansión a Masa Constante (E.M.C.). Con esta prueba se quiere encontrar el comportamiento del fluido del yacimiento al caer la presión en los poros de la roca, sin que ocurra el proceso de producción. Esta prueba sirve para conocer el grado y el punto de separación de las fases líquida y gaseosa.

Posteriormente se realiza un agotamiento a volumen constante (A.V.C.), el cual representará el comportamiento volumétrico del fluido dentro de un volumen poroso constante. El proceso de disminución de presión se simula por medio de una serie de expansiones isotérmicas; después de cada expansión el vapor se extrae de la celda de tal manera que se alcanza el volumen original y luego la celda se agita para alcanzar el equilibrio termodinámico antes de la siguiente expansión.

Una simulación de esta prueba utilizando una Ecuación de Estado como la de Peng-Robinson (EDE-PR) puede efectuarse siguiendo el diagrama mostrado en la figura 7.

Figura 7. Esquema de la simulación de una prueba AVC.



Fuente: AHMED, T. H. "Hydrocarbons Phase Behavior".⁵

2.4. CÁLCULOS FLASH

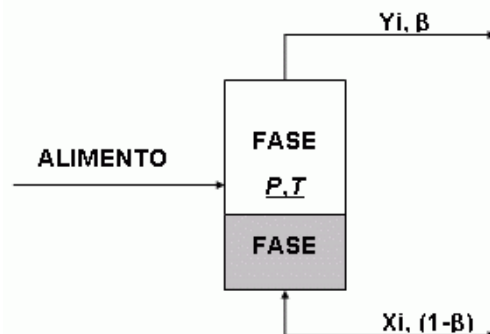
Los procesos FLASH son variaciones de las condiciones fisicoquímicas de una mezcla de fluido que se efectúan de una manera instantánea en los cuales se mantienen constantes algunas propiedades termodinámicas, para poder predecir el comportamiento del fluido como una función de las variables que afectan la mezcla.

Existen muchos tipos de Flash, entre los más importantes tenemos:

- A Presión y Temperatura Constantes. (P,T)
- A Presión y Entropía Constantes. (P,S)
- A Presión y Entalpía Constantes. (P,H)
- A Presión y Composición Constantes. (P,β)
- A Temperatura y Entropía Constantes. (T,S)
- A Temperatura y Entalpía Constantes. (T,H)
- A Presión y Composición Constantes. (T,β)

El flash que ocurre en los experimentos PVT es de tipo (P,T) y es una expansión isotérmica del fluido, como se presenta en la Figura 8:

Figura 8. Esquema de un Flash (P,T) .



Fuente: PEDERSEN, K. S., FREDENSLUND, A. and THOMASSEN, P., "Properties of Oil and Natural Gases".⁶

Los resultados de un Flash P-T son:

- Número de fases presentes en el equilibrio.
- Cantidades molares de cada fase.
- Composición molar de cada fase.
- Densidades de fases.
- Cantidades molares de entropía, entalpía y capacidad calorífica de cada fase.

La formulación matemática de un flash P-T es la siguiente:

$$1. \text{ Balance total de materia: } F = L + V \quad (2.1)$$

$$2. \text{ Balance de materia por componente: } Z_i F = X_i (1-b) + Y_i b \quad (2.2)$$

$$b = \frac{V}{F} \quad (2.3)$$

$$3. \text{ Equilibrio termodinámico: } f_i^L = f_i^V \quad (2.4)$$

$$K_i = \frac{Y_i}{X_i} = \frac{f_i^L}{f_i^V}$$

Si asumimos $F=1$ y si reemplazamos X_i por (2.5) en (2.2) obtendremos:

$$Y_i = \frac{K_i \times Z_i}{1 + b \times (K_i - 1)} \quad X_i = \frac{Z_i}{1 + b \times (K_i - 1)}$$

Aplicando el balance de materia en cada fase:

$$\sum_{i=1}^{N_C} Z_i = \sum_{i=1}^{N_C} X_i = \sum_{i=1}^{N_C} Y_i = 1$$

De tal manera tendremos que:

$$\sum_{i=1}^{N_c} X_i = \sum_{i=1}^{N_c} \frac{Z_i}{1 + b \times (K_i - 1)} = 1$$

Y por lo mismo:

$$\sum_{i=1}^{N_c} Y_i = \sum_{i=1}^{N_c} \frac{K_i \times Z_i}{1 + b \times (K_i - 1)} = 1$$

En resumen, se requiere resolver el siguiente sistema de ecuaciones simultáneas no lineales:

- Balance de materia:

$$Z_i = b \times Y_i + (1 - b) \times X_i$$

- Restricción de las fracciones:

$$\sum_{i=1}^{N_c} (Y_i - X_i) = \sum_{i=1}^{N_c} \frac{Z_i \times (K_i - 1)}{1 + b \times (K_i - 1)} = 0$$

$$f_i^v = f_i^L \quad \Phi_i^v = \frac{f_i^v}{y_i p} \quad \Phi_i^L = \frac{f_i^L}{x_i p}$$

Por lo tanto:

$$Y_i \times f_i^v = X_i \times f_i^L$$

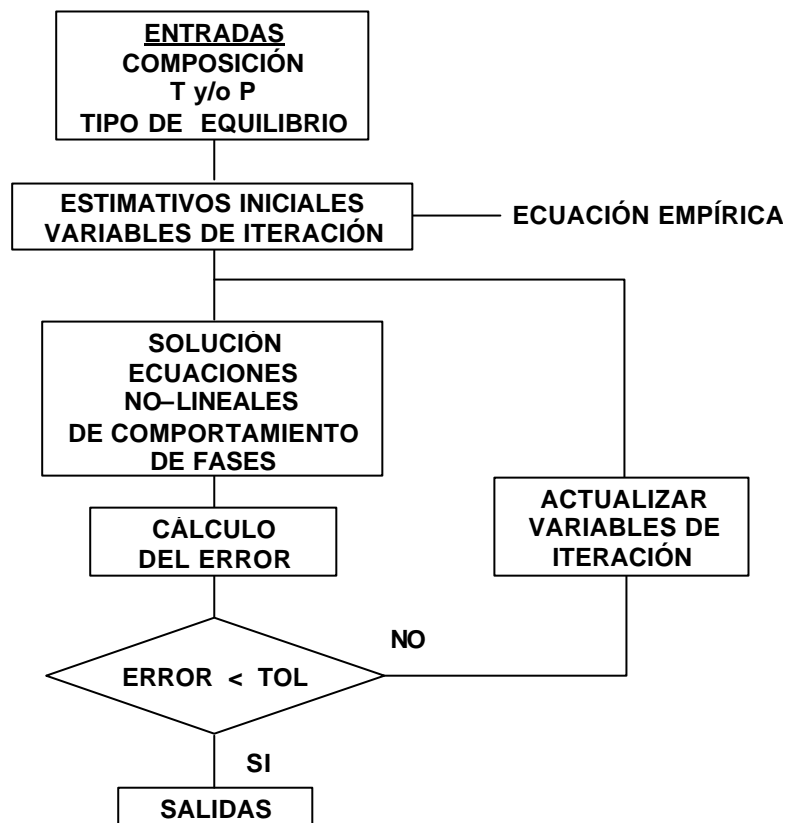
2NC+12NC+1

2.5. MODELADO DEL EQUILIBRIO DE FASES

El cálculo del comportamiento de fases.

La Ecuación de Estado. La figura 9 representa la metodología de trabajo en la solución del problema del equilibrio de fases de compuestos petrolíferos.

Figura 9. Esquema para el cálculo del comportamiento de fases por medio de una Ecuación de Estado.



Fuente: VAROTSIS, N. "A Robust Prediction Method for Rapid Phase Behavior Calculations".⁷

2.5.1. Métodos de solución de ecuaciones de equilibrio líquido-vapor.

Existen métodos que solucionan un conjunto de ecuaciones simultáneas de balance de materia junto con relaciones termodinámicas, los cuales son la aproximación clásica de la solución de problemas de equilibrio (Ver ecuaciones 2.11 a 2.14).

Debido a que estos métodos solo utilizan balances de material y el criterio de igualdad en los potenciales químicos para encontrar el equilibrio de fases,

sufren una mayor desviación y en algunos casos pueden fallar en la predicción de el numero de fases o conducir a soluciones triviales. Estos resultados erróneos son causados porque en dichos estados de equilibrio no se ha minimizado la energía de Gibbs aun cuando ambos criterios de equilibrio se hayan satisfecho.

También estos métodos pueden arrojar soluciones triviales si los valores iniciales de equilibrio no son apropiadamente estimados. Esto se presenta aun para dos fases, especialmente en cercanías al punto crítico.

Otro tipo de método, utiliza como herramienta cálculo, la energía libre de Gibbs buscando minimizarla con respecto a los números molares (o las composiciones) de los diferentes componentes de la mezcla.

Es importante resaltar que los métodos iterativos pueden llegar a encontrar mínimos locales de la energía de Gibbs y no un mínimo global.

2.5.1.1. Métodos de Iteración. Todos estos métodos utilizan técnicas de optimización y se pueden dividir en:

a. Métodos de Iteración por Sustituciones Sucesivas (ISS).

Se basan en las ecuaciones de balance de material (global y por componentes) y en la igualdad de fugacidades de las fases en el equilibrio.

Su esquema iterativo es muy sencillo:

1. Hacer un estimativo inicial del valor de la constante de equilibrio (K_i) por medio de una correlación empírica.
2. Calcular composiciones de prueba de las fases presentes en equilibrio.

3. Calcular las fugacidades de cada componente en cada fase por medio de una Ecuación de Estado.
4. Verificar la convergencia por medio de la igualdad de las fugacidades en equilibrio: $f_i^L - f_i^V < ERROR$. Si este criterio se cumple entonces se puede encontrar las composiciones de las fases en equilibrio, de lo contrario hay que volver al paso 2.

Los principales problemas de los métodos ISS son su baja convergencia y la falta de estabilidad en puntos cercanos a la saturación y en la región crítica.

La fortaleza del método radica en que no requieren estimativos iniciales cercanos a la solución.

Se han efectuado grandes esfuerzos para acelerar y estabilizar la convergencia de estos métodos:

Risnes et al. (1981)⁸ presenta un algoritmo acelerado basado en la suposición de una solución aproximada y un procedimiento de "Bring-Back" para parar la convergencia de soluciones triviales.

Mehra (1983)⁸ ha investigado tres propuestas basadas sobre el método de extrapolación del gradiente y ha dado mejoras significativas en la rata de convergencia en todos los casos aunque esta es errática.

b. Métodos tipo Newton o de primer orden.

El objetivo de estos métodos es minimizar la diferencia entre las fugacidades del líquido y del vapor del sistema, las cuales son calculadas por medio de una Ecuación de Estado. El método de iteración se consigue al predecir los siguientes valores de prueba por medio de un esquema de Newton-Raphson:

1. Calcular K_i por medio de una correlación empírica.
2. Asumir un valor de β y calcular:
$$f(\mathbf{b}) = \sum_{i=1}^{N_c} \frac{Z_i \times (K_i - 1)}{1 + \mathbf{b} \times (K_i - 1)}$$
3. Si la ecuación 2.15 es menor que una tolerancia dada entonces se pueden calcular las composiciones de las fases por medio de las ecuaciones de equilibrio.

Si esto no se cumple entonces se efectúa una iteración de Newton-Raphson:

$$f'(\mathbf{b}) = \sum_{i=1}^{N_c} \frac{Z_i \times (K_i - 1)^2}{[1 + \mathbf{b} \times (K_i - 1)]^2}$$

$$\mathbf{b}_{NUEVO} = \mathbf{b} + \frac{f(\mathbf{b})}{f'(\mathbf{b})}$$

Con este nuevo valor de β ir al paso 2. Hasta que la respuesta se encuentre dentro de la tolerancia. Luego hay que iterar sobre K_i para lograr que las fugacidades sean independientes del valor K_i asumido inicialmente hasta que se obtenga igualdad de fugacidades de las fases presentes para cada componente.

La ventaja que presentan frente a un esquema de sustituciones sucesivas radica en que pueden obtener soluciones cuando las condiciones del cálculo de fases están próximas al punto crítico o a las presiones de saturación. La convergencia de estos métodos depende fuertemente del estimativo inicial, por lo tanto tienen un menor radio de convergencia, razón por la cual si las suposiciones iniciales difieren mucho de los valores finales, se puede dar el caso de la no-convergencia o la convergencia hacia soluciones triviales, que no serían representativas del comportamiento de los fluidos.

Fussell y Yanosik⁸ han utilizado un procedimiento de mínima variable de Newton-Raphson (MVNR) con derivadas analíticas. Este método ha probado tener una mejor tasa de convergencia que los ISS y puede ser utilizado en cálculos en las cercanías del punto crítico.

La dificultad del procedimiento MVNR se basa en la necesidad de computar una complicada matriz Jacobiana y su inversa en cada iteración, además como lo ha demostrado Baker y Luks (1980) su convergencia depende altamente del estimativo inicial. Asselineau (1979) lo ha utilizado para calcular equilibrio liquido-vapor. Boston y Britt (1978) han usado un método Quasi-Newton en el cual la matriz identidad se utiliza como el estimativo inicial de la inversa del Jacobiano, y esta a su vez es actualizada por el Método de Broyden.

c. Métodos Híbridos.

Estos métodos intentan lograr esquemas de solución que aprovechen la ventaja del amplio radio de convergencia de los métodos de 1° orden y la rapidez y el amplio rango en las condiciones de aplicación que poseen los esquemas de 2° orden.

En un esquema híbrido, en las primeras etapas de cómputo cuando los valores calculados están alejados de la solución, se utiliza un método de 1° orden. Luego cuando se obtienen buenos valores iniciales, se cambia a un esquema de 2° orden para acelerar la convergencia.

Estos métodos además no necesitan realizar inicialmente cálculos de presión de saturación para saber en que zona del diagrama de fases está el sistema a las condiciones de cálculo (P,T).

Nghiem⁹ combina el método ISS con un método híbrido de Powell (combinación de un método Newton-Raphson cuadráticamente convergente con un método de Mayor-Descenso linealmente convergente).

Mehra (1982) combina su método acelerado de ISS con el método de Powell en el cual se utilizan derivadas analíticas.

Sin embargo, Mott (1983) ha reportado recientemente que el uso del método de Newton-Raphson con derivadas analíticas como el método de segundo orden es más poderoso que el uso del método de Powell.

Michelsen ha acelerado el esquema de sustituciones sucesivas básico utilizando el método del Valor-propio Dominante General de Crowe y Nishio (1975) pero se debe efectuar un cambio al esquema de Newton-Raphson si la convergencia no se obtiene.

2.5.1.2. Métodos que utilizan la Minimización de la Energía Libre de Gibbs. Este tipo de métodos son de reciente formulación para resolver problemas de equilibrio de fases con fórmulas termodinámicas básicas: Heidemann (1974); Gautam y Seider (1979); Trangenstein (1985).

Michelsen, ha desarrollado un algoritmo importante, basado en el concepto del plano tangente para probar la estabilidad de fases. El algoritmo minimiza la distancia vertical desde la línea tangente a la curva de Gibbs. Eubank y Barrufet (1988) han desarrollado un algoritmo para predecir las composiciones en equilibrio de mezclas binarias. Este algoritmo depende de la presencia de puntos de inflexión y de la continuidad uniforme de la función de Gibbs.

Muchos de estos algoritmos traen dos características comunes:

1. Se enfatiza mucho más en la velocidad computacional y la eficiencia que en la confiabilidad y efectividad. Esto era de esperarse porque las técnicas de optimización derivativa no son realmente técnicas de “Minimización” como lo demostró Trangenstein (1985).

Aclarando, estos métodos (derivativos) dan soluciones a los sistemas de ecuaciones no lineales buscando puntos estacionarios los cuales deberían ser mínimos locales, máximos locales, o puntos de concavidad; y en dichos estados la energía de Gibbs no es necesariamente minimizada.

2. Proveen estimativos iniciales para comenzar con el proceso computacional.

Swank y Mullins (1986) anotan que todos estos algoritmos pueden fallar en la detección de la inestabilidad de las fases y por ende en el equilibrio de fases si los estimativos iniciales están muy alejados de las composiciones en equilibrio.

3. MÉTODOS Y CÁLCULOS UTILIZADOS EN LA SIMULACIÓN PVT

Este capítulo describe los procedimientos, métodos, ecuaciones y las correlaciones utilizadas en el desarrollo de la herramienta software predictiva que permite simular las pruebas PVT composicionales.

El corazón del simulador de pruebas es la ecuación de estado de Peng-Robinson, la cual produce cálculos volumétricos y de fugacidades basándose en parámetros físicos de los componentes de la mezcla. En este capítulo también se incluye una descripción de las modificaciones más importantes y necesarias para mejorar la predicción volumétrica y la sensibilidad de la ecuación al componente pesado.

Se presenta a su vez en este capítulo, el método de Nikos Varotsis¹⁷ para efectuar el cálculo inicial de la relación de equilibrio líquido - vapor, el cual ofrece una mejor predicción que los métodos tradicionales. Además se incluye una descripción del algoritmo flash utilizado en las pruebas basados en un criterio que permite efectuar una minimización de la energía libre de Gibbs y que por ello asegura la convergencia del sistema de ecuaciones formado en el cálculo de equilibrio de fases.

Buscando optimizar los resultados del simulador, se incluyó el procedimiento para la corrección de los valores $a(T)$, b y α de la ecuación de estado de Peng-Robinson propuesta por Tarek Ahmed¹⁵.

3.1. ECUACIONES DE ESTADO

Las ecuaciones de estado son herramientas versátiles para aplicaciones de ingeniería, ya que permiten relacionar presión, volumen y temperatura para componentes puros y para mezclas además relacionan la composición.

El desarrollo de ecuaciones de estado para representar sistemas reales comenzó en 1873 con la publicación de la ecuación de Van der Waals – VdW -. Desde entonces, cientos de ecuaciones se han propuesto.

Los cálculos de Presión - Volumen - Temperatura y el equilibrio de fases para gas y aceite están basados en una ecuación de estado.

Leland (1980) distinguió cuatro familias de ecuaciones de estado:

1. Familia de Van der Waals
2. Familia de Benedict - Webb - Rubin
3. Ecuaciones de fluido de referencia
4. Ecuaciones de cuerpo rígido aumentadas.

La familia de Van der Waals abarca las ecuaciones de estado más sencillas, la mayoría de ellas son cúbicas. Su principal característica es la separación entre el efecto de repulsión y de atracción. Pero a pesar de su simplicidad, estas ecuaciones muestran un comportamiento cuantitativamente correcto, siendo capaces de predecir equilibrio multifásico, puntos tricríticos, y otros fenómenos complicados.

La familia de Benedict - Webb - Rubin incluye ecuaciones de estado complicadas y extensiones empíricas de la ecuación virial. Además de la

ecuación del mismo nombre, son miembros importantes de esta familia la de Starling (1973) y la de Kesler & Lee (1975).

Las ecuaciones de fluido de referencia pretender representar precisamente una gran cantidad de datos PVT para sustancias puras, para lo cual utilizan muchos parámetros. Una aplicación importante de este tipo de ecuaciones es su uso como fluidos de referencia en la teoría de los estados correspondientes.

La familia de las ecuaciones de cuerpo rígido aumentado combina la descripción de fuerzas de repulsión entre moléculas duras de formas diferentes con expresiones para la atracción molecular. La base teórica del término repulsivo es la característica más importante y promisorio de este tipo de ecuaciones. La expresión de Carnahan & Starling (1972) para la repulsión entre esferas duras ha sido usada frecuentemente.

Actualmente, las cuatro familias reciben considerable atención, sin embargo, la literatura reciente contiene desarrollos principalmente para las ecuaciones de la familia de Van der Waals debido a su simplicidad y para la familia de cuerpo rígido debido a su fundamento teórico y a su potencial para realizar predicciones de gran precisión.

Dentro de las ecuaciones cúbicas más conocidas se puede mencionar la de Redlich-Kwong y la de Soave-Redlich-Kwong, las cuales, a pesar de su buen comportamiento en la determinación de la densidad del vapor, tienen algunas deficiencias en común, entre las cuales la más notable es la incapacidad para generar valores satisfactorios en el cálculo de la densidad del líquido.

La necesidad de mayor exactitud en el comportamiento volumétrico de las fases coexistentes en los cálculos de equilibrio, condujo a la creación de una

nueva ecuación de estado de dos constantes desarrollada por Peng Ding Yu y Robinson Donald B. (1975), la cual combina simplicidad y exactitud y muestra grandes ventajas en la predicción de las densidades de la fase líquida. Por esta razón, los cálculos PVT y de equilibrio de fases en este proyecto estarán basados en dicha ecuación.

Aunque las ecuaciones de estado complicadas puedan dar mejores predicciones en cuanto a las densidades del líquido que una ecuación cúbica, un tercer parámetro que provee una corrección del volumen permite subsanar esta desventaja, sin perder la sencillez y la honradez que caracterizan a las ecuaciones cúbicas.

3.1.1. Ecuación de Estado de Peng Robinson. Desde la aparición de la ecuación de Van der Waals en 1873, muchos autores han propuesto variaciones en la relación semiempírica. Una de las más exitosas modificaciones fueron realizadas por Redlich-Kwong (1949). Desde entonces se han propuesto numerosas modificaciones a esta ecuación (Redlich-Dunlop, 1963; Chueh-Prausnitz, 1967; Wilson, 1969; Zudkevitch-Joffe, 1970; y otros). Algunos han introducido funciones de desviación para ajustar datos PVT de sustancias puras, mientras que otros han mejorado la capacidad de la ecuación para predecir el equilibrio vapor - líquido - EVL -. Una de las más recientes modificaciones a la ecuación RK fue propuesta por Soave (1972). La ecuación de Soave-Redlich-Kwong (SRK) ha tenido gran acogida en la industria del procesamiento de hidrocarburos, debido a su relativa simplicidad y su capacidad para generar con razonable exactitud relaciones de equilibrio en cálculos de EVL.

La ecuación original RK y la SRK, a pesar de su buen comportamiento en la determinación de la densidad del vapor, tienen algunas deficiencias en

común, entre las cuales la más notable es la incapacidad para generar valores satisfactorios en el cálculo de la densidad del líquido.

La necesidad de mayor exactitud en la predicción del comportamiento volumétrico de las fases coexistentes en cálculos EVL, condujo a la creación de una nueva ecuación de estado de dos constantes desarrollada por Peng Ding-Yu y Robinson Donald B. (1975)¹⁰; la cual combina simplicidad y exactitud y muestra grandes ventajas en la predicción de las densidades de la fase líquida.

3.1.1.1. Formulación de la Ecuación. Las ecuaciones de estado semiempíricas generalmente expresan la presión como la suma de dos términos, una presión de repulsión P_R y una presión de atracción P_A donde:

$$P = P_R + P_A \quad (3.1)$$

La presión de repulsión se expresa mediante la ecuación de esfera dura de Van der Waals,

$$P_R = \frac{RT}{v - b} \quad (3.2)$$

La presión de atracción puede ser expresada como,

$$P_A = -\frac{a}{g(v)} \quad (3.3)$$

Donde $g(v)$ es una función del volumen molar v y la constante b se relaciona con el tamaño de las esferas duras. El parámetro a puede ser interpretado como una medida de las fuerzas de atracción intermoleculares. Aplicando la ecuación 3.1 en el punto crítico donde la primera y segunda derivada de la presión con respecto al volumen son cero, se pueden obtener expresiones

para a y b en función de las propiedades críticas. Mientras que b generalmente se usa como un parámetro independiente de la temperatura, a es constante solamente en la ecuación de Van der Waals. En las ecuaciones RK y SRK se usan factores de escala adimensionales para describir la dependencia de temperatura del parámetro de energía.

Un estudio de las ecuaciones semiempíricas que tienen la forma de la ecuación 3.1 indica que mediante la selección de una función apropiada para $g(v)$, el factor de compresibilidad crítico predicho puede acercarse a un valor más real. La aplicabilidad de la ecuación a muy altas presiones es afectada por la magnitud de b/v_c , donde v_c es el volumen crítico determinado.

Además, comparando la ecuación original RK y la ecuación SRK, es evidente que utilizando el factor de escala adimensional para el parámetro de energía como una función del factor acéntrico y la temperatura reducida, ha mejorado significativamente la predicción de presiones de vapor para sustancias puras y consecuentemente las constantes de equilibrio para mezclas.

Peng y Robinson propusieron la siguiente ecuación,

$$P = \frac{RT}{v - b} - \frac{a(T)}{v(v + b) + b(v - b)} \quad (3.4)$$

Esta ecuación expresada en términos de factor de compresibilidad,

$$Z^3 - (1 - B)Z^2 + (A - 3B^2 - 2B)Z - (AB - B^2 - B^3) = 0 \quad (3.5)$$

donde,

$$A = \frac{aP}{R^2T^2} \quad (3.6)$$

$$B = \frac{bP}{RT} \quad (3.7)$$

$$Z = \frac{P_v}{RT} \quad (3.8)$$

La ecuación 3.5 produce una o tres raíces dependiendo del número de fases existentes en el sistema. En la región de dos fases, la raíz mayor pertenece al factor de compresibilidad de la fase vapor, mientras que la raíz positiva más pequeña corresponde al de la fase líquida.

Aplicando la ecuación 3.4 en el punto crítico se obtiene,

$$a(T_c) = 0.45724 \frac{R^2 T_c^2}{P_c} \quad (3.9)$$

$$b(T_c) = 0.07780 \frac{RT_c}{P_c} \quad (3.10)$$

$$Z_c = 0.307 \quad (3.11)$$

A temperaturas diferentes de la crítica se tiene:

$$a(T) = a(T_c) \mathbf{a}(T_r, \mathbf{w}) \quad (3.12)$$

$$b(T) = b(T_c) \quad (3.13)$$

Donde $\mathbf{a}(T_r, \mathbf{w})$ es una función adimensional de la temperatura reducida y el factor acéntrico, e igual a la unidad a la temperatura crítica.

$$\mathbf{a} = (1 + m\alpha(1 - T_r^{0.5}))^2 \quad (3.14)$$

m es una constante característica de cada sustancia y ha sido correlacionada con el factor acéntrico de la siguiente manera :

$$m = 0.37464 + 1.54226 \mathbf{w} - 0.26992 \mathbf{w}^2 \quad (3.15)$$

Mas adelante (1978), esta ecuación fue extendida y expresada de la siguiente forma:

$$m = 0.379642 + 1.48503 \mathbf{w} - 0.1644 \mathbf{w}^2 + 0.016667 \mathbf{w}^3 \quad (3.16)$$

Es importante tener en cuenta que la ecuación 3.15 es similar a la obtenida por Soave (1972) para la ecuación SRK. Esta ecuación se obtiene para cada sustancia usando datos de presión de vapor desde el punto de ebullición normal hasta el punto crítico; mientras que Soave utilizó únicamente el punto crítico y la presión de vapor calculada a $T_r=0.7$ basado en el valor del factor acéntrico.

Aplicando la relación termodinámica

$$\ln \frac{f}{p} = \int_0^p \left(\frac{v}{RT} - \frac{1}{P} \right) dP \quad (3.17)$$

A la ecuación 3.4, se puede derivar la siguiente expresión para la fugacidad de compuestos puros,

$$\ln\left(\frac{f}{p}\right) = (Z-1) - \ln(Z-B) - \frac{A}{2\sqrt{2}B} \ln\left(\frac{Z+2.414B}{Z-0.414B}\right) \quad (3.18)$$

La fugacidad es una medida de la energía molar de Gibbs de un gas real. La fugacidad puede ser interpretada como una presión de vapor modificada para representar correctamente la tendencia de escape de las moléculas desde una fase a otra.

Se introduce la fugacidad del componente en cada fase para desarrollar un criterio de equilibrio termodinámico. Físicamente la fugacidad de un componente en cada fase con respecto a la fugacidad del componente en una segunda fase es una medida del potencial de transferencia del componente entre las fases. La fase con la fugacidad del componente más baja, acepta el componente proveniente de la segunda fase. Fugacidades iguales de un componente en las dos fases da como resultado una transferencia neta de cero. Una transferencia de cero para todos los componentes, implica que un sistema de hidrocarburos se encuentra en

equilibrio termodinámico. Por lo tanto la condición de equilibrio termodinámico se puede expresar matemáticamente como:

$$f_i^V = f_i^L \quad (3.19)$$

El coeficiente de fugacidad del componente i en una mezcla de hidrocarburos sea líquida o gaseosa, es una función de la presión del sistema, la fracción molar y la fugacidad del componente:

$$\Phi_i^V = \frac{f_i^V}{y_i p} \quad (3.20)$$

$$\Phi_i^L = \frac{f_i^L}{x_i p} \quad (3.21)$$

El coeficiente de fugacidad del componente i en una mezcla se puede calcular mediante la siguiente ecuación:

$$\ln \Phi_i^L = \frac{b_i}{b} (Z^L - 1) - \ln(Z^L - B) - \frac{A}{2\sqrt{2}B} \left[\frac{2y_i}{y} - \frac{b_i}{b} \right] \ln \left[\frac{Z^L + 2.414B}{Z^L - 0.414B} \right] \quad (3.22)$$

donde:

$$y_i = \sum_j \left[x_j (a_i a_j \mathbf{a}_i \mathbf{a}_j)^{0.5} (1 - d_{ij}) \right] \quad (3.23)$$

$$y = \sum_i \sum_j \left[x_i x_j (a_i a_j \mathbf{a}_i \mathbf{a}_j)^{0.5} (1 - d_{ij}) \right] \quad (3.24)$$

La ecuación 3.22 también se puede usar para determinar Φ_i^V utilizando la composición de la fase vapor y_i en el cálculo de A, B, Z^V , y los otros términos dependientes de la composición.

Los parámetros de mezcla utilizados en las ecuaciones 3.23 y 3.24 se definen mediante las siguientes reglas de mezcla,

$$a = \sum_i \sum_j x_i x_j a_{ij} \quad (3.25)$$

$$b = \sum_i x_i b_i \quad (3.26)$$

donde:

$$a_{ij} = (1 - d_{ij}) a_i^{1/2} a_j^{1/2} \quad (3.27)$$

En la ecuación 3.27 d_{ij} es un coeficiente de interacción binaria determinado empíricamente que caracteriza el binario formado por el componente i y el componente j.

3.1.1.2. Coeficientes de Interacción Binaria. Los coeficientes de interacción binaria se introducen para tener en cuenta el efecto de las fuerzas polares en las interacciones entre dos componentes. Usualmente, d_{ij} para interacciones hidrocarburo-hidrocarburo se consideran iguales, o al menos cercanos a cero.

Las ecuaciones cúbicas de estado inicialmente no eran destinadas para usarse con componentes que se desvían de la esfericidad, pero dicha desviación ocurre con los compuestos pesados de un fluido de yacimiento de petróleo. Una posible manera de compensar la no esfericidad de los hidrocarburos pesados podría ser introducir coeficientes de interacción binaria diferentes de cero para las interacciones que incluyan hidrocarburos pesados.

Peng y Robinson encontraron que los coeficientes de interacción binaria dependen de la diferencia en tamaño molecular de los componentes en un sistema binario.

Las propiedades que caracterizan los coeficientes de interacción binaria son las siguientes¹¹:

- La interacción entre componentes hidrocarburos incrementa a medida que la diferencia relativa entre sus pesos moleculares aumenta:

$$d_{i, j+1} > d_{i, j}$$

$$d_{i+1, j} < d_{i, j}$$

- Componentes hidrocarburos con el mismo peso molecular tienen un coeficiente de interacción binaria de cero:

$$d_{i, i} = 0$$

- La matriz de coeficientes de interacción binaria es simétrica:

$$d_{i, j} = d_{j, i}$$

3.1.1.3. Correlaciones empleadas para determinar Coeficientes de Interacción Binaria. Los coeficientes de interacción binaria entre los diferentes binarios se pueden determinar utilizando múltiples correlaciones, que tienen como objetivo, generalizar una propiedad física de uno de los componentes del binario que represente fidedignamente la susodicha desviación de la esfericidad y por consiguiente la aparición y el subsecuente efecto de las fuerzas polares en el cálculo de las propiedades del binario.

Existen múltiples correlaciones de este tipo. Katz y Firoozabadi (1988) determinaron una correlación gráfica para encontrar los coeficientes de interacción binaria entre el metano y los grupos de Número de Carbono Simple, representada por la siguiente ecuación:

$$d_{C1-Ci} = 0.14 g_i - 0.0668 \quad (3.28)$$

Una de las correlaciones más utilizadas y que posee una base física y estadística muy fuerte es la propuesta por Nikos Varotsis y colaboradores en 1986¹².

En esta se evalúan los coeficientes de interacción binaria para la EDE-PR como una función de la presión, temperatura y el factor acéntrico del hidrocarburo. La forma generalizada de la correlación se basa en un polinomio de Chevishev de cuarto grado, de la forma:

$$d_{i,j} = k_2 T_{rj}^2 + k_1 T_{rj} + k_0 \quad (3.29)$$

Donde el subíndice i se refiere a los componentes principales (N₂, CO₂, CH₄) y j se refiere a los otros componentes hidrocarburos del binario. Los coeficientes k₀, k₁ y k₂ dependen del factor acéntrico y se determinan para cada grupo binario aplicando las siguientes expresiones:

Para Nitrógeno - Hidrocarburos:

$$k_0 = 0.1751787 - 0.7043 \log(w_j) - 0.862066 [\log(w_j)]^2 \quad (3.30)$$

$$k_1 = -0.584474 + 1.328 \log(w_j) + 2.035767 [\log(w_j)]^2 \quad (3.31)$$

$$k_2 = 2.257079 + 7.869765 \log(w_j) + 13.50466 [\log(w_j)]^2 + 8.3864 [\log(w_j)]^3 \quad (3.32)$$

Efectuando corrección por presión, se tiene:

$$d_{ij}' = d_{ij} (1.04 - 4.2 \times 10^{-5} p) \quad (3.33)$$

donde p es la presión en psia.

Para CO₂ - Hidrocarburos:

$$k_o = 0.4025635 + 0.1748927 \log(w_j) \quad (3.34)$$

$$k_1 = -0.94812 - 0.6009864 \log(w_j) \quad (3.35)$$

$$k_2 = 0.741843368 + 0.441775 \log(w_j) \quad (3.36)$$

Realizando corrección por presión tenemos:

$$d_{ij}' = d_{ij} (1.044269 - 4.375 \times 10^{-5} p) \quad (3.37)$$

Para CH₄ - Hidrocarburos:

$$k_o = -0.01664 - 0.37283 \log(w_j) + 1.31757 [\log(w_j)]^2 \quad (3.38)$$

$$k_1 = 0.48147 + 3.3542 \log(w_j) - 1.0783 [\log(w_j)]^2 \quad (3.39)$$

$$k_2 = -0.4114 - 3.5072 \log(w_j) - 0.78798 [\log(w_j)]^2 \quad (3.40)$$

La aplicabilidad de las anteriores ecuaciones (3.30 a la 3.40) debe ser extendida con precaución para casos donde T_j exceda en mucho el valor de 1.

Para los sistemas que contienen H₂S se utilizó la correlación de Carroll y Mather¹³ que depende de la temperatura de ebullición normal del hidrocarburo parafínico y se expresa como:

$$d_{ij} = 0.084 + 5.99 \times 10^{-5} T_b - 4.23 \times 10^{-7} T_b^2 \quad (3.41)$$

donde T_b está en grados Kelvin.

Firoozabadi y Arbabi¹⁴ encontraron que en fluidos cercanos al punto crítico la predicción de los coeficientes de interacción binaria debe ser muy tenida en cuenta ya que influyen en gran medida en la capacidad predictiva de la ecuación de estado. Por ello propusieron la siguiente relación:

$$d_{i,j} = d_{ij}^o \left[1 + \frac{g_{ij}}{b+z} \right] \quad (3.42)$$

donde:

$$z = -\frac{n^2}{RT} \left[\frac{p}{n} \right] T \quad (3.43)$$

Además:

γ_{ij} , β y δ_{ij}^o son constantes para cada sistema binario. Las cuales se pueden encontrar por medio de una minimización de datos estadísticos.

Para mezclas de fluidos hidrocarburos propusieron:

$$d_{c1,cn} = 0.0314 + 2.179 \times 10^{-4} (MW)_j \quad (3.44)$$

Siendo $(MW)_j$ el peso molecular del hidrocarburo.

La cual es válida para C_n : $C_3 - C_{20}$ y en un rango de temperatura de 160-176°F.

3.1.1.4. Modificaciones a la Ecuación de Estado de Peng – Robinson.

1. Ahmed (1988)¹⁵ plantea una modificación importante de la EDE PR en cuanto a la obtención de los parámetros de la ecuación. La modificación propuesta se basa en que la EDE PR, al igual que la mayoría de las ecuaciones de estado cúbicas, se apoya en dos suposiciones que difícilmente se cumplen en el caso de mezclas de hidrocarburos; tales suposiciones son:

i) Las expresiones para los parámetros a y b se obtienen partiendo del supuesto que $\partial P / \partial V$ y $\partial^2 P / \partial V^2$ son cero en el punto crítico. Esto es

cierto solo para compuestos puros; pero en el caso de mezclas de hidrocarburos, donde la fracción pesada agrupa gran cantidad de compuestos, difícilmente sus parámetros cumplirán con las ecuaciones propuestas para ellos.

ii) La expresión para m en la ecuación de α se obtiene haciendo regresión de datos de presión de vapor desde el punto de ebullición normal hasta el punto crítico para diferentes compuestos puros. Nuevamente tal expresión será difícil que cumpla en el caso de fracciones pesadas.

Para eliminar las anteriores suposiciones, Ahmed trabajó con muestras a las que se les conocía sus propiedades físicas (densidad, presión de vapor, etc.) y que se podían considerar como fracciones pesadas. Mediante regresión se encontraron unas expresiones para calcular a , b y m para las fracciones pesadas y se encontró que:

$$a(T_c) \circ b = C_0 + C_1 D + C_2 D^2 + C_3 D^3 + C_4 / D + C_5 g + C_6 g^2 + C_7 / g \quad (3.45)$$

donde:

$$D = (M / g)_{C_7^+} \quad (3.46)$$

M es el peso molecular y γ la gravedad específica de la fracción pesada.

$$a = \left(1 + m \left(1 - \left(\frac{520}{T} \right)^{0.5} \right) \right)^2 \quad (3.47)$$

donde:

$$m = (D / (C_0 + C_1 D)) + C_2 M + C_3 M^2 + C_4 / M + C_5 g + C_6 g^2 + C_7 / g \quad (3.48)$$

La temperatura T esta dada en grados Rankin y los coeficientes $C_0 - C_7$ son los que se presentan en la tabla 1.

Tabla 1. Constantes C_0 a C_7 para las ecuaciones 3.45 y 3.48.

C	a(T_c)	b	m
C_0	-2.433525E7	-6.8453198	-36.91776
C_1	8.3201587E3	1.730243E-2	-5.2393763E-2
C_2	-18.444102	-6.2055064E-6	1.7316235E-2
C_3	3.6003101E-2	9.0910383E-9	-1.3743308E-5
C_4	3.4992796E7	13.378698	12.718844
C_5	2.838756E7	7.9492922	10.246122
C_6	-1.1325365E7	-3.1779077	-7.6697942
C_7	6.418828E6	1.7190311	-2.6078099

Fuente: AHMED, T. H. "A Practical Modification of the Peng-Robinson Equation of State".¹⁵

Además, utilizando la EDE PR en términos de Z (ecuación 3.5) se reportaron los siguientes valores de a , b y m para el metano y nitrógeno:

Tabla 2. Valores de a , b , y m para CH_4 y N_2 optimizados.

Compuesto	a(T_c)	b	m
Metano	7709.7080	0.46749727	-0.549765
Nitrógeno	4569.3589	0.46825820	-0.97962859

Fuente: AHMED, T. H. "A Practical Modification of the Peng-Robinson Equation of State".¹⁵

1. Siguiendo el método de Peneloux, Rauzy y Freze (1982) para mejorar las predicciones volumétricas introduciendo un tercer parámetro en la ecuación de estado de dos parámetros, Jhaveri y Youngren (1984)¹⁶

trabajaron con un parámetro de corrección de volumen c_i para la EDE PR. Este tercer parámetro tiene las mismas unidades del segundo parámetro b_i de la EDE PR original y se define mediante la siguiente relación:

$$c_i = S_i b_i \quad (3.49)$$

donde S_i es un parámetro de traslación dimensional.

El parámetro de corrección c_i no cambia las condiciones de equilibrio vapor - líquido determinadas por la ecuación de estado original, pero modifica los volúmenes de líquido y vapor efectuando ciertas traslaciones a lo largo de los ejes de volumen. Los volúmenes de las fases de hidrocarburos corregidos están dados por las siguientes expresiones:

$$V_{corr}^L = V^L - \sum_{i=1} (x_i c_i) \quad (3.50)$$

$$V_{corr}^V = V^V - \sum_{i=1} (y_i c_i) \quad (3.51)$$

Donde V^L y V^V son los volúmenes de las fases líquida y vapor calculados por la EDE PR original, en pies cúbicos por mol, y V_{corr}^L , V_{corr}^V son los volúmenes corregidos de las fases líquida y vapor.

La tabla 3 muestra los valores del parámetro de traslación para los hidrocarburos definidos y el Dióxido de carbono:

Tabla 3. Valores del parámetro de translación para C₁-C₆ y CO₂.

Componente	S _i
C ₁	-0.1540
C ₂	-0.1002
C ₃	-0.08501
i-C ₄	-0.07935
n-C ₄	-0.06413
i-C ₅	-0.04350
n-C ₅	-0.04183
n-C ₆	-0.01478
CO ₂	0.2463

Fuente: JHAVERI, B. S., " Three - Parameter Modification of the Peng-Robinson Equation of State to Improve Volumetric Predictions".¹⁶

Con el fin de caracterizar el parámetro de traslación para pseudocomponentes que describen la fracción heptano-plus, los autores primero calcularon el parámetro de traslación para hidrocarburos bien definidos más pesados que el n-heptano (n-alcanos, n-alkilciclohexanos y n-alkilbencenos) igualando el volumen molar a condiciones estándar y lo correlacionaron con el peso molecular, extendiendo su aplicabilidad para los pseudocomponentes de tal manera que la correlación desarrollada tiene la siguiente presentación,

$$S_i = 1 - \frac{d}{(M_i)^e} \quad (3.52)$$

Donde M_i es el peso molecular de cada pseudocomponente, d y e son coeficientes positivos de la correlación.

Estos se pueden obtener igualando algunos datos de densidad del heptano-plus obtenido en el laboratorio a condiciones estándar. Los autores utilizaron los siguientes tipos de datos:

- Cuando el laboratorio reporta datos de densidad de varios cortes de la fracción de heptano-plus provenientes de una prueba de destilación de punto de ebullición verdadero, se calculan las densidades del seudocomponente utilizando una regla de mezcla y se utilizan para ajustar los coeficientes d y e.
- Algunas pruebas del laboratorio, como el agotamiento a volumen constante y la liberación diferencial, reportan la densidad de la fracción heptano- plus en el fluido original del yacimiento así como también en el fluido agotado (fase líquida) a una presión más baja. Esos dos datos son suficientes para ajustar los dos coeficientes (d, e).
- En ausencia de los anteriores, un dato que está normalmente disponible, es la densidad de la fracción heptano-plus en el fluido del yacimiento. Entonces, para igualar esta densidad, ajustamos únicamente el coeficiente d y al coeficiente e se le asigna un valor de 0.2051, que es un promedio de los coeficientes de los tres grupos (n-alcanos, n-alkilciclohexanos y n-alkilbencenos) dados así:

Tabla 4. Valores de d y e para la corrección de Jhaveri y Youngren.

Tipo de componente	d	e	% desviación
n-Alcanos	2.258	0.1823	0.19
n-Alkil-Ciclohexanos	3.004	0.2324	0.28
n-Alkil-Bencenos	2.516	0.2008	0.24

Fuente: JHAVERI, B. S., " Three - Parameter Modification of the Peng-Robinson Equation of State to Improve Volumetric Predictions".¹⁶

En el procedimiento de ajuste de los coeficientes, se toman estimativos iniciales para $d=2.5927$ y $e=0.2051$, promediados de los valores para los tres grupos. Los límites para estos coeficientes son:

$$\begin{aligned} 2.2 < d < 3.2 \\ 0.18 < e < 0.25 \end{aligned}$$

3.2. RELACIONES DE EQUILIBRIO

La relación de equilibrio K_i de un componente dado se define como la relación existente entre la fracción molar del componente en la fase vapor y_i , y la fracción molar del componente en la fase líquida x_i . Tal relación se expresa de la siguiente manera:

$$K_i = \frac{y_i}{x_i} \quad (3.53)$$

A presiones bajas (menores de 100 psia), se puede suponer que un sistema de hidrocarburos en el que estén presentes la fase vapor y la fase líquida cumple las leyes de Raoult y Dalton para soluciones ideales, o sea que para un componente i de la mezcla:

$$p_i = x_i p_{vi} \quad (\text{Ley de Raoult}) \quad (3.54)$$

$$p_i = y_i p \quad (\text{Ley de Dalton}) \quad (3.55)$$

Donde p_i es la presión parcial del componente i en psia, p_{vi} es la presión de vapor del componente i en psia y p es la presión total del sistema, psia.

La ecuación 3.54 requiere que el vapor esté en equilibrio con la fase líquida. En equilibrio y de acuerdo con las leyes de estado anunciadas, la presión parcial ejercida por un componente en la fase vapor debe ser igual a la presión ejercida por el mismo componente en la fase líquida. Por lo tanto combinando las ecuaciones 3.54 y 3.55 se tiene:

$$x_i p_{vi} = y_i P \quad (3.56)$$

Reagrupando esta ecuación e introduciendo el concepto de equilibrio,

$$\frac{y_i}{x_i} = \frac{p_{vi}}{P} = K_i \quad (3.57)$$

De la segunda definición para K_i dada en la ecuación 3.57, se concluye que para una temperatura y presión dada el valor de K_i es una constante y no depende de la cantidad del componente i presente en la mezcla ni de los otros componentes que puedan estar presentes en ella, o sea, K_i no depende de la composición; sin embargo es importante tener en cuenta que la relación de equilibrio se cumple solo para cuando el sistema esta a presiones bajas y el comportamiento del gas se puede considerar como ideal. Por otro lado, a presiones altas K_i depende de la composición y en equilibrio cada componente de la mezcla esta asociado con su relación de equilibrio. Las suposiciones hechas por la ley de Dalton (la fase vapor es un gas ideal) y la ley de Raoult (la fase líquida es una solución ideal) no son reales y conducen a predicciones inexactas de relaciones de equilibrio a altas presiones.

Para una solución real, las relaciones de equilibrio ya no están solo en función de la presión y la temperatura sino también de la composición de la mezcla de hidrocarburos. Esta observación se puede enunciar matemáticamente como:

$$K_i = K(p, T, z_i) \quad (3.58)$$

Las relaciones de equilibrio son fundamentales para los cálculos de comportamiento de fases y por tanto es necesario tener claridad sobre los métodos para obtenerlas.

3.2.1. Método de Varotsis para obtener relaciones de equilibrio. Nikos Varotsis¹⁷ (1987) desarrolló un nuevo método para calcular las relaciones de equilibrio de todos los componentes presentes en un aceite de yacimiento incluyendo los no-hidrocarburos (CO₂, H₂S, N₂) y la fracción pesada.

El autor construyó un diagrama X-Y de tal manera que cada posición del fluido sobre el mapa, se determina por las coordenadas dadas por la siguiente ecuación:

$$X \text{ o } Y = C_3 T^3 + C_2 T^2 + C_1 T + C_0 \quad (3.59)$$

Donde el coeficiente C_i (i = 0,..., 3) está expresado como una función de una presión aparente del sistema por la siguiente expresión:

$$C_i = \frac{b_{1i} P_A^2 + b_{2i} P_A + b_{3i}}{b_{4i} P_A^2 + b_{5i} P_A + 1} \quad (3.60)$$

Los coeficientes b_{ji}, donde j = 1,...,5, están dados en la tabla 5. La presión aparente P_A es la presión corregida del sistema para explicar las diferencias entre el fluido bajo prueba y el tipo de muestras empleadas para derivar la correlación (P = 5000 psi).

Cuando la presión de convergencia P_k de un fluido es menor de 5000 psi, P_A está dado por la ecuación:

$$P_A = P \left(\frac{5000}{P} \right)^Q \quad (3.61)$$

Donde

$$Q = \exp[(2.91936 + 2337/(5000 - P)) \ln(P/P_k)] \quad (3.62)$$

Cuando la presión de convergencia es mayor de 5000 psi, entonces:

$$P_A = P - \frac{Q (1 - P/P_k)^{0.3}}{\left(1 + 3.101818 \text{ ABS } \left((1 - P/5000)^{2.717366} \right) \right)} \quad (3.63)$$

Donde:

$$Q = 6.0943 (P_k - 5000)^{0.6} (P/5000)^{2.2} \quad (3.64)$$

Para calcular la presión de convergencia P_k de aceites, el autor recomienda la correlación de Standing:

$$P_k = 60M_{C_{7+}} - 4200 \quad (3.65)$$

Para gases condensados propone la siguiente ecuación:

$$P_k = \exp \{ 10.540064 - (1.475 + 6.15 \text{ } C_6/z_{C1}) (1 - z_{N_2}) \} \quad (3.66)$$

Donde C_6 es la composición molecular de la fracción z_{C6} del fluido incluyendo los no-hidrocarburos y z_{C1} , z_{N2} son las composiciones molares del metano y del nitrógeno respectivamente.

Tabla 5. Coeficientes de la Ecuación 3.60.

P < 2000 psia				
X				
b _{1 i}	b _{2 i}	b _{3 i}	b _{4 i}	b _{5 i}
6.86975E-5	-0.048146	10.5467	-7.31624E-7	0.037488
-5.61969E-7	0.001521	0.00481	1.94450E-5	0.024891
2.89064E-9	-5.76324E-6	7.15409E-8	3.20288E-5	0.030499
-5.64312E-12	9.08963E-9	-2.90899E-8	5.50798E-5	0.043878
Y				
b _{1 i}	b _{2 i}	b _{3 i}	b _{4 i}	b _{5 i}
3.46988E-5	1.42366	13.0098	6.45032E-5	0.157406
1.68769E-6	-0.003977	-0.07666	1.05736E-4	0.122398
-6.85363E-9	1.22156E-5	2.79665E-4	1.84848E-4	0.127791
9.77563E-12	-1.44605E-8	-3.45563E-7	2.77283E-4	0.13797
P > 2000 psia				
X				
b _{1 i}	b _{2 i}	b _{3 i}	b _{4 i}	b _{5 i}
3.80737E-9	0.001279	0.553423	0.0	1.12899E-4
1.74465E-10	-2.80582E-6	0.0093099	0.0	-1.46291E-4
3.84480E-12	-2.53417E-8	3.560230E-5	0.0	-9.19080E-5
-1.15588E-14	8.09514E-11	-1.27115E-7	0.0	3.76260E-15
Y				
b _{1 i}	b _{2 i}	b _{3 i}	b _{4 i}	b _{5 i}
1.96564E-8	-1.54607E-4	6.71666	0.0	1.19442E-4
-3.77314E-10	2.62402E-6	-0.0034873	0.0	-5.57765E-4
-1.39143E-12	1.03696E-8	-1.73708E-5	0.0	-1.12448E-4
4.37478E-15	-3.35722E-11	5.72960E-8	0.0	3.90625E-5

Fuente: VAROTSIS, N. "A Robust Prediction Method for Rapid Phase Behavior Calculations".¹⁷

Debido a la poca precisión que ofrece la ecuación 3.66 al predecir la presión de convergencia se optó por reemplazarla por una propuesta por Rzasa,

Glass y Opfell (1952)¹⁸ para determinar la presión de convergencia de mezclas de hidrocarburos livianos:

$$P_k = \exp \left\{ 10.540064 - \left(1.475 + 6.15 \frac{C_6}{z_{C1}} \right) (1 - z_{N_2}) \right\} \quad (3.67)$$

Donde M_{C7+} y g_{C7+} son el peso molecular y la gravedad específica del C_{7+} respectivamente, y $a_1 - a_3$ son los coeficientes de la correlación cuyos valores son:

$$a_1 = 6124.3049$$

$$a_2 = -2753.2538$$

$$a_3 = 415.42049$$

La ecuación de Rzasz, se puede usar para determinar la P_k de mezclas de hidrocarburos con rango de temperatura entre 50 y 300 °F.

En el método de Varotsis, cada componente puro de hidrocarburo se representa sobre el mapa mediante sus propias coordenadas (X_i , Y_i) las cuales son función del factor acéntrico de los componentes y se expresan así:

$$X_i = \frac{-92.779w_i^2 + 9.8631w_i + 8.5905}{3.524w_i^2 + 3.3073w_i + 1} \quad (3.68)$$

$$Y_i = \frac{30.869w_i^2 - 9.0277w_i + 2.1606}{14.3127w_i^2 - 4.1398w_i + 1} \quad (3.69)$$

Para los compuestos no-hidrocarburos se tienen las siguientes coordenadas: CO_2 (7.3949, 2.1566), H_2S (7.9482, 1.2306), N_2 (9.5041, 2.1596)

Las coordenadas de las fracciones pesadas se determinan con las siguientes correlaciones:

$$\begin{aligned} X_i &= -0.0473594 M_{C7+} + 4.1557 \\ Y_i &= 0.828721 \times 10^{-4} M_{C7+} + 2.1271 \end{aligned} \quad (3.70)$$

Donde M_{C7+} es el peso molecular de la fracción pesada.

La línea que une las coordenadas de presión y temperatura (X,Y) del fluido con la posición de cada componente sobre el mapa (X_i , Y_i), intercepta los ejes del valor K en un punto que corresponde al valor de equilibrio del compuesto seleccionado. Los valores de K_i se determinan con la siguiente expresión:

$$\log K_i = -2.9993 + 0.671284 \left[\frac{\frac{Y}{Y-Y_i} - \frac{X}{X-X_i} - 3.97812 \frac{1}{Y-Y_i}}{-\frac{1}{X-X_i} + 0.005515 \frac{1}{Y-Y_i}} \right] \quad (3.71)$$

3.3. CÁLCULOS FLASH

Un cálculo de las condiciones de equilibrio de una mezcla dada a una presión y temperatura dadas se denomina un cálculo flash-(P,T) y se utiliza para determinar las cantidades y composiciones molares de cada fase.

Considerando una mol de un sistema de hidrocarburos que tiene una composición z_i , $i = 1, 2, \dots, n$ (fracción molar) donde n es el número de componentes. Este sistema a unas condiciones de temperatura y presión dadas estará compuesto por V moles de vapor de composición y_i y L moles de líquido de composición x_i . El cálculo de equilibrio vapor - líquido consiste en determinar los valores de V , L , x_i y y_i , los cuales deben cumplir con un

balance de masa, un balance molar y con las condiciones de equilibrio requeridas. Dichas condiciones se presentaron en el capítulo anterior.

3.3.1. Método de Nghiem.⁹ L. X. Nghiem, creó un algoritmo robusto para cálculos flash que combina una versión especial del método de sustituciones sucesivas con el procedimiento de Powell si se observa una convergencia lenta. Con esta metodología se busca asegurar la convergencia del cálculo flash. Se ha demostrado que este algoritmo de solución converge en las cercanías al punto crítico y también detecta la región de una sola fase sin calcular la presión de saturación.

El cálculo del equilibrio líquido - vapor usando ecuaciones de estado genera un sistema de ecuaciones no lineales que debe resolverse iterativamente. Comúnmente el procedimiento usado es el de sustituciones sucesivas, pero exhibe una tasa de convergencia pobre cerca al punto crítico. Para solucionar este problema, Fusell y Yanosik (1989) utilizaron un método de Newton para resolver el sistema. Su desventaja es la necesidad de calcular una complicada matriz Jacobiana y su inversa en cada iteración.

Además su radio de convergencia es relativamente pequeño comparado con el del método de Sustituciones Sucesivas, por lo tanto requiere una buena aproximación inicial para que se pueda alcanzar la convergencia.

La metodología de Nghiem presenta un algoritmo iterativo para cálculos flash utilizando ya sea la ecuación de estado de Soave-Redlich-Kwong o la de Peng - Robinson, el cual combina el esquema de sustituciones sucesivas con el de Powell, que a su vez es un procedimiento híbrido entre uno Cuasi-Newton y otro denominado método del descenso más rápido, de esta forma se aprovecha la simplicidad del primero y la fortaleza del segundo.

El esquema de sustituciones sucesivas se modificó de manera que la región de una sola fase pueda ser detectada sin tener que calcular presión de saturación.

En el enfoque tradicional, al resolver las ecuaciones resultantes de los balances de masa, molar y la condición de equilibrio termodinámico, no necesariamente asegura que se efectúe una minimización de la energía libre de Gibbs, por lo tanto no se puede asegurar que exista un verdadero equilibrio entre las fases presentes. Por esta razón, pueden existir soluciones de dicho sistema de ecuaciones donde β no se encuentre en el rango normal de cero a uno, siendo β la fracción de gas en la mezcla:

$$b = \frac{V}{F}$$

Entonces:

$b \approx 0$: Fase líquida Y $b \approx 1$: Fase gaseosa

Para que se cumpla el balance molar:

$$g(b) = \sum_{i=1}^{N_c} (Y_i - X_i) = \sum_{i=1}^{N_c} \frac{Z_i \times (K_i - 1)}{1 + b \times (K_i - 1)} = 0 \quad (3.71a)$$

Esta condición se debe cumplir siempre que se calculen nuevos valores de K_i y en equilibrio de fases $0 \leq \beta \leq 1$.

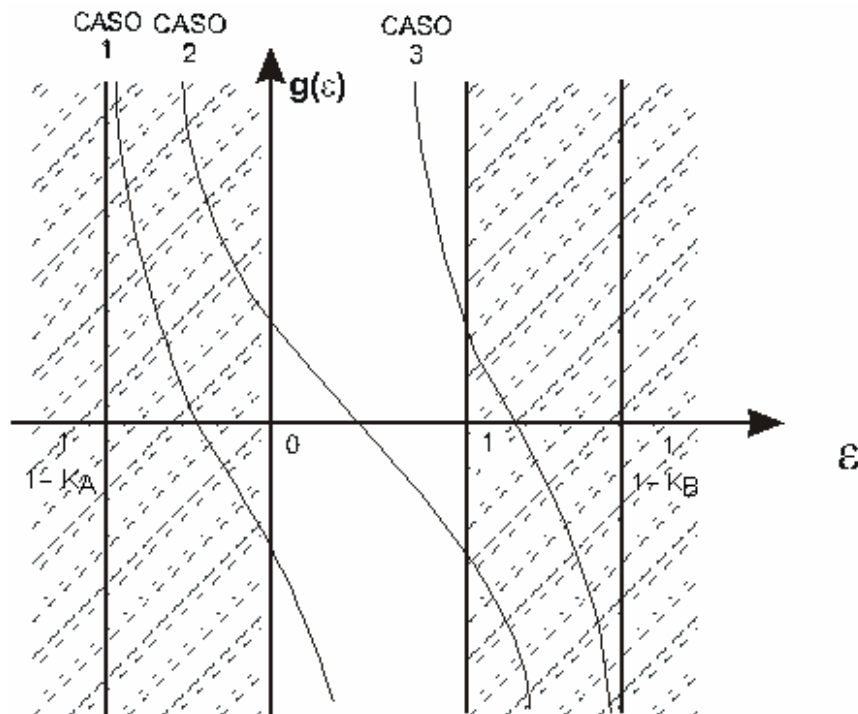
Si definimos $\epsilon \equiv \beta$, entonces la función $g(\epsilon)$ es monotónica decreciente (La derivada de g con respecto a ϵ es menor que cero), para cualquier ϵ con asintotas:

$$e' = \frac{1}{1 - K_j} \quad (3.72)$$

Donde la asíntota izquierda se presenta cuando K_j es el menor valor (K_B) y la asíntota derecha cuando K_j es el mayor valor (K_A).

Una condición necesaria pero no suficiente para la existencia de la función $g(\varepsilon) \in [0,1]$ es que $\frac{1}{1-K_A} < 0 < 1 < \frac{1}{1-K_B}$. En la Figura 10 se observan tres casos:

Figura 10. Condiciones para la existencia de la solución de $g(\varepsilon)$ en $[0,1]$.



Fuente: SAN MIGUEL, Luis Alberto. "Métodos alternos para el ajuste automático de las ecuaciones de estado".²⁹

CASO 1: $g(0) < 0$ y $g(\varepsilon_0) < 0$

En este caso la mezcla se encuentra en estado monofásico (líquido).

β se asume igual a 0.5 en la primera iteración y a 0 en la subsecuentes iteraciones.

CASO 2: $g(0) \geq 0$ y $g(1) \leq 0$ entonces $g(\varepsilon_0) \in [0,1]$

Existen más de una fase en equilibrio.

β se obtiene utilizando un método iterativo como el de Rachford & Rice¹⁹.

CASO 3: $g(1) > 0$ y $g(\varepsilon_0) > 1$

La mezcla se encuentra en estado monofásico (gaseoso).

β se asume como 0.5 en la primera iteración y como 1 en las iteraciones subsecuentes.

En la práctica se encuentra que este procedimiento produce raíces ε_0 perteneciente al intervalo $[0,1]$ para un sistema de dos fases y para un sistema monofásico la relación entre la fugacidad del componente i en la fase líquida y su fugacidad en la fase vapor, converge a un valor que podría ser diferente a la unidad.

Para terminar, este esquema ha sido probado extensamente y ha predicho la región monofásica correctamente. Recientemente se ha mostrado que este procedimiento es equivalente a una prueba de estabilidad de una fase basada en el criterio del plano tangente.

3.3.1.1. Algoritmo del método de Sustituciones Sucesivas modificado de Nghiem.

PRIMERA ITERACIÓN

1. Encontrar un estimativo inicial de K_i .
2. Encontrar $g(\beta=0)$ y $g(\beta=1)$ con la ecuación 3.71a.
 - 2.1. Si $g(\beta=0) > 0$ y $g(\beta=1) < 0$ coexisten dos fases, entonces encontrar β por medio del método iterativo (Rachford & Rice)

- 2.2. Si $g(\beta=0) \leq 0$ ó $g(\beta=1) \geq 0$ existe una sola fase, entonces $\beta=0.5$
3. Calcular las fracciones de gas y líquido con las ecuaciones 2.6 y 2.7.

ITERACIONES SUBSECUENTES

4. Computar los coeficientes de fugacidad por medio de la ecuación de estado.
5. Calcular: $f_{i,m} = \phi_{i,m} X_{i,m} P$.

Siendo $f_{i,m}$, $\phi_{i,m}$, $X_{i,m}$ la fugacidad, el coeficiente de fugacidad y la fracción molar del componente i en la fase m.

$$F_i = \frac{f_{i,L}}{f_{i,g}} \quad y \quad K_i = F_i \frac{X_{i,g}}{X_{i,L}}$$

6. Chequear la convergencia del proceso utilizando una norma Euclidiana:

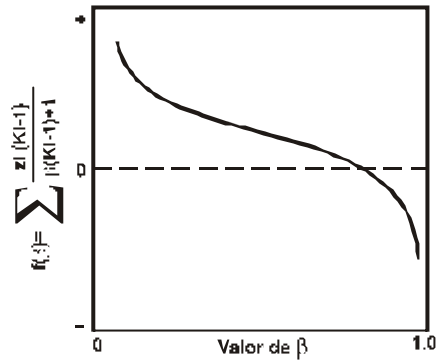
$$\sum_{i=1}^n (F_i - 1)^2 \leq TOL \quad TOL = 1 \times 10^{-7}$$

Si la norma Euclidiana se cumple, los valores calculados (K_i y β) son los correspondientes al equilibrio de fases. De lo contrario continuar con el procedimiento.

7. Computar $g(\beta=0)$ y $g(\beta=1)$ con la ecuación 3.71a
- 2.1. Si $g(\beta=0) < 0$ entonces $\beta=0$ (Sólo fase líquida).
- 2.2. Si $g(\beta=1) > 0$ entonces $\beta=1$ (Sólo fase vapor)
- 2.3. Si $g(\beta=0) > 0$ y $g(\beta=1) < 0$ entonces encontrar β por medio de método iterativo.
8. Actualizar $X_{i,L}$ y $X_{i,g}$ con los valores de β y K_i recientemente calculados e ir al paso 4.

3.3.1.2. Método de Rachford & Rice.¹⁹

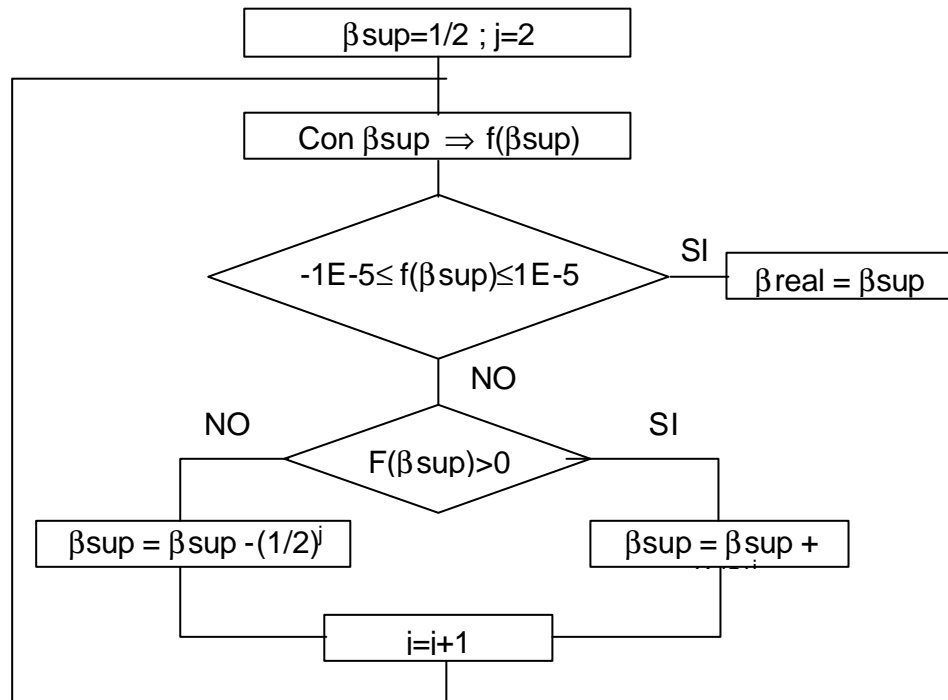
Figura 11. Método de Rachford & Rice para encontrar el valor de β .



Fuente: RACHFORD, H. H. "Procedure for use of Electronic Digital Computers in Calculating Flash Vaporization Hydrocarbon Equilibrium".¹⁹

El método de Rachford & Rice se basa en el siguiente algoritmo iterativo:

Figura 12. Diagrama de flujo del proceso para encontrar la β en un estado determinado utilizando el método de Rachford & Rice.



Fuente: RACHFORD, H. H. "Procedure for use of Electronic Digital Computers in Calculating Flash Vaporization Hydrocarbon Equilibrium".¹⁹

En la figura 12 la función $f(\beta)$ está definida como:

$$f(\mathbf{b}) = \sum \frac{Z_i (K_i - 1)}{b(K_i - 1) + 1} \quad (3.73)$$

3.4. CÁLCULO DE PRESIONES DE SATURACIÓN

Conocer el valor de las presiones de saturación es de gran importancia para obtener un estimativo del punto en el cual la mezcla comenzará a presentar una separación de fases. El objetivo primordial de este cálculo es predecir un rango aproximado de presiones en el cual se presenta el equilibrio de fases buscando mejorar la precisión de las lecturas de laboratorio en esta región.

3.4.1. Determinación de la Presión de Rocío.⁵ Es el valor de presión en el cual existe vapor saturado a una temperatura dada y aparece una cantidad infinitesimal de líquido. Este valor se expresa comúnmente como P_d .

La presión de rocío de una mezcla se describe matemáticamente por las siguientes condiciones:

$$n_v = 1.0$$

$$y_i = z_i \quad 1 \leq i \leq n$$

$$\sum_{i=1}^n \left[\frac{z_i}{K_i} \right] = 1 \quad (3.74)$$

Aplicando la definición de K_i en términos del coeficiente de fugacidad en la ecuación 3.74, tenemos:

$$\sum_{i=1}^n \left[\frac{z_i f_i^v}{f_i^L} \right] = \sum_{i=1}^n \left[\frac{z_i f_i^v}{z_i P_d f_i^L} \right] = 1 \quad (3.75)$$

Ó

$$P_d = \sum_{i=1}^n \left[\frac{f_i^v}{f_i^L} \right] \quad (3.76)$$

La ecuación 3.76 se arregla de la siguiente manera:

$$f(P_d) = \sum_{i=1}^n \left[\frac{f_i^v}{f_i^L} \right] - P_d^A = 0 \quad (3.77)$$

Donde:

P_d = Presión en el Punto de Rocío, psia.

P_d^A = Presión asumida en el punto de rocío, psia

f_i^v = Fugacidad del componente i en la fase vapor, psia

ϕ_i^L = Coeficiente de fugacidad del componente i en la fase líquida.

La ecuación 3.76 se puede resolver para la presión en el punto de rocío usando el método iterativo de Newton-Raphson. Para usar el método iterativo, se requiere la derivada de la ecuación 3.77 con respecto a la presión. Esta derivada se obtiene de la siguiente expresión:

$$\frac{df}{dP_d} = \sum_{i=1}^n \left[\frac{f_i^L \left(\frac{f_i^v}{f_i^L} \right) - f_i^v \left(\frac{f_i^L}{f_i^L} \right)}{(f_i^L)^2} \right] - 1 \quad (3.78)$$

Estas derivadas pueden ser determinadas numéricamente así:

$$\frac{f_i^v}{f_i^L} = \left[\frac{f_i^v(P_d + \Delta P_d) - f_i^v(P_d - \Delta P_d)}{2\Delta P_d} \right] \quad (3.79)$$

$$\frac{f_i^L}{P_d} = \left[\frac{f_i^L(P_d + \Delta P_d) - f_i^L(P_d - \Delta P_d)}{2\Delta P_d} \right] \quad (3.80)$$

Donde:

ΔP_d = Incremento de Presión, 5 psia, por ejemplo.

$f_i^V(P_d + \Delta P_d)$ = Fugacidad del componente i a $(P_d + \Delta P_d)$.

$f_i^V(P_d - \Delta P_d)$ = Fugacidad del componente i a $(P_d - \Delta P_d)$.

$\phi_i^L(P_d + \Delta P_d)$ = Coeficiente de fugacidad del componente i a $(P_d + \Delta P_d)$.

$\phi_i^L(P_d - \Delta P_d)$ = Coeficiente de fugacidad del componente i a $(P_d - \Delta P_d)$.

ϕ_i^L = Coeficiente de fugacidad del componente i a P_d .

El procedimiento computacional para determinar la P_d se resume en los siguientes pasos:

Paso 1: Se asume un valor inicial para la presión del punto de rocío P_d^A .

$$P_d = \frac{1}{\sum_i \frac{z_i}{P_{ci} \exp \left[5.37 \left(1 + w_i \right) \left(1 - \frac{T_{ci}}{T} \right) \right]}} \quad (3.81)$$

Paso 2: Usando el valor de P_d^A , calcular las razones de equilibrio para los componentes de la mezcla usando la correlación de Wilson donde se obtiene K_i a partir de:

$$K_i = \frac{P_{ci}}{P} \exp \left[5.37 \left(1 + w_i \right) \left(1 - \frac{T_{ci}}{T} \right) \right] \quad (3.82)$$

Donde:

P_{ci} = Presión crítica del componente i, psia.

P = Presión del sistema " P_d^A ", psia.

T_{ci} = Temperatura crítica del componente i, °R.

T = Temperatura del sistema, °R.

ω_i = Factor acéntrico del componente i.

Paso 3: Calcular la composición de las gotas de líquido presentes aplicando la definición matemática de K_i , dada:

$$x_i = \frac{z_i}{K_i} \quad (3.83)$$

Vale la pena anotar que $y_i = z_i$

Paso 4: Usando la composición de la fase gas z_i y la de la fase líquida x_i , calcular f_i^V y ϕ_i^L a P_d , $(P_d + \Delta P_d)$, $(P_d - \Delta P_d)$, ya sea por la ecuación de SRK o de PR.

Paso 5: Evaluar las ecuaciones 3.77 y 3.78.

Paso 6: Usando los valores de la función $f(P_d)$ y la de la derivada $\partial f / \partial P_d$ como se determinó en el paso 5, calcular una nueva presión de punto de rocío aplicando la fórmula de Newton-Raphson:

$$P_d = P_d^A - \frac{f(P_d)}{\partial f / \partial P_d} \quad (3.84)$$

Paso 7: El valor calculado de P_d es evaluado numéricamente contra el valor asumido aplicando la siguiente condición:

$$\left| P_d - P_d^A \right| \leq 5 \quad (3.85)$$

Si la condición anterior se cumple, entonces la presión del punto de rocío P_d correcta ha sido encontrada. Si no, repetir los procedimientos desde el paso 2 al 6, usando la P_d calculada como el nuevo valor para la próxima iteración.

3.4.2. Determinación de la presión de burbuja.⁵ La presión en el punto de burbuja P_b se define como la presión a la cual se forma la primera burbuja de gas.

De acuerdo con esto dicha presión se describe matemáticamente por las siguientes condiciones:

$$x_i = z_i \quad 1 \leq i \leq n$$

$$n_L = 1.0$$

$$\sum_{i=1}^n [z_i K_i] = 1 \quad (3.86)$$

Aplicando la definición de K_i en términos del coeficiente de fugacidad en la ecuación 3.86, tenemos:

$$P_b = \sum_{i=1}^n \left[\frac{f_i^L}{f_i^V} \right] \quad (3.87)$$

La ecuación 3.87 se arregla de la siguiente manera:

$$f(P_b) = \sum_{i=1}^n \left[\frac{f_i^L}{f_i^V} \right] - P_b^A = 0 \quad (3.88)$$

Donde:

P_b = Presión en el Punto de burbuja, psia.

P_b^A = Presión asumida en el punto de burbuja, psia.

f_i^L = Fugacidad del componente i en la fase líquida, psia.

ϕ_i^V = Coeficiente de fugacidad del componente i en la fase vapor.

La ecuación 3.88 se puede resolver para la presión en el punto de burbuja utilizando el método iterativo de Newton-Raphson. Para usar el método iterativo, se requiere la derivada de la ecuación 3.87 con respecto a la presión. Esta derivada se obtiene de la siguiente expresión:

$$\frac{df}{dP_b} = \sum_{i=1}^n \left[\frac{f_i^V (f_i^L / P_b) - f_i^L (f_i^V / P_b)}{(f_i^V)^2} \right] - 1 \quad (3.89)$$

Estas derivadas pueden ser determinadas numéricamente así:

$$\frac{f_i^L}{P_b} = \left[\frac{f_i^L (P_b + \Delta P_b) - f_i^L (P_b - \Delta P_b)}{2\Delta P_b} \right] \quad (3.90)$$

$$\frac{f_i^V}{P_b} = \left[\frac{f_i^V (P_b + \Delta P_b) - f_i^V (P_b - \Delta P_b)}{2\Delta P_b} \right] \quad (3.91)$$

Donde:

ΔP_b = Incremento de Presión, 5 psia, por ejemplo.

$f_i^L (P_b + \Delta P_b)$ = Fugacidad del componente i a $(P_b + \Delta P_b)$.

$f_i^L (P_b - \Delta P_b)$ = Fugacidad del componente i a $(P_b - \Delta P_b)$.

$\phi_i^V (P_b + \Delta P_b)$ = Coeficiente de fugacidad del componente i a $(P_b + \Delta P_b)$.

$\phi_i^V (P_b - \Delta P_b)$ = Coeficiente de fugacidad del componente i a $(P_b - \Delta P_b)$.

ϕ_i^V = Coeficiente de fugacidad del componente i a P_b .

El procedimiento computacional para determinar la P_b se resume en los siguientes pasos:

Paso 1: Se asume un valor inicial para la presión del punto de rocío P_b^A .

$$P_b = \sum_i Z_i P_{ci} \exp \left[5.37 \left(1 + w_i \right) \left(1 - \frac{T_{ci}}{T} \right) \right] \quad (3.92)$$

Paso 2: Usando el valor de P_b^A , calcular las razones de equilibrio para los componentes de la mezcla usando la correlación de Whitson y Torp⁵, donde se obtiene K_i a partir de:

$$K_i = \left[\frac{P_{ci}}{P_k} \right]^{A-1} \left[\frac{P_{ci}}{P} \right] \text{EXP} \left[5.37(1 + w_i) \left(1 - \frac{T_{ci}}{T} \right) \right] \quad (3.93)$$

Donde P_k está dada por la ecuación 3.64 y:

$$A = 1 - \left[\frac{P - 14.7}{P_k - 14.7} \right]^{0.6} \quad (3.94)$$

P_{ci} = Presión crítica del componente i, psia.

P = Presión del sistema " P_b^A ", psia.

P_k = Presión de convergencia, psia.

T_{ci} = Temperatura crítica del componente i, °R.

T = Temperatura del sistema, °R.

ω_i = Factor acéntrico del componente i.

Paso 3: Calcular la composición de la fase vapor, aplicando la definición matemática de K_i , dada:

$$y_i = z_i K_i \quad (3.95)$$

Vale la pena anotar que $x_i = z_i$

Paso 4: Usando la composición de la fase gas y y la de la fase líquida z , calcular f_i^L y ϕ_i^V a P_b , $(P_b + \Delta P_b)$, $(P_b - \Delta P_b)$, ya sea por la ecuación de SRK o de PR.

Paso 5: Evaluar las ecuaciones 3.88 y 3.89.

Paso 6: Usando los valores de la función $f(P_b)$ y la de la derivada $\partial f/\partial P_b$ como se determinó en el paso 5, calcular una nueva presión de punto de burbuja aplicando la fórmula de Newton-Raphson :

$$P_b = P_b^A - \frac{f(P_b)}{\partial f / \partial P_b} \quad (3.96)$$

Paso 7: El valor calculado de P_b es evaluado numéricamente contra el valor asumido aplicando la siguiente condición:

$$\left| P_b - P_b^A \right| \leq 5 \quad (3.97)$$

Si la condición anterior se cumple, entonces la presión del punto de burbuja P_b correcta ha sido encontrada. Si no, repetir los procedimientos desde el paso 2 al 6, usando la P_b calculada como el nuevo valor para la próxima iteración.

3.5. DETERMINACIÓN DE LAS PROPIEDADES FÍSICAS DE LAS FRACCIONES DE HIDROCARBUROS

Existen numerosas correlaciones para determinar las propiedades físicas de las fracciones pesadas o petróleos crudos, la mayoría de las cuales emplean como parámetros de correlación la temperatura de ebullición, el peso molecular y la gravedad específica, de los cuales los dos últimos generalmente se conocen con más frecuencia.

Una selección adecuada de los parámetros antes mencionados es importante debido a que pequeños cambios en ellos causan variaciones significativas en los resultados obtenidos.

3.5.1. Método de Katz-Firoozabadi. Katz y Firoozabadi⁵ (1978) presentaron un grupo de propiedades físicas para componentes con número de carbono simple (NCS) de 6 a 45. Las propiedades tabuladas incluyen punto de ebullición promedio, gravedad específica y peso molecular.

Whitson²⁰ (1983) modificó las propiedades físicas tabuladas originalmente para hacer su uso más consistente. La modificación realizada empleó la correlación de Riazi y Daubert⁵ para extrapolar los datos de peso molecular desde el C₂₂ al C₄₅, en la cual los coeficientes a, b y c fueron re-calculados mediante un modelo de regresión no lineal que permite ajustar el peso molecular en el intervalo antes mencionado. La ecuación posteriormente fue utilizada para calcular los pesos moleculares desde el C₂₃ al C₄₅, calculándose también las propiedades críticas y el factor acéntrico desde el C₆ al C₄₅ en términos de sus temperaturas de ebullición, gravedades específicas y el valor modificado del peso molecular. Las propiedades generalizadas están dadas en la tabla 6.

Ahmed²¹ (1985) utilizando un modelo de regresión correlacionó las propiedades físicas tabuladas por Katz-Firoozabadi-Whitson (KFW) con el número de átomos de carbono de la fracción. La ecuación generalizada tiene la siguiente forma:

$$q = a_1 + a_2 n + a_3 n^2 + a_4 n^3 + a_5 / n \quad (3.98)$$

Donde θ es cualquier propiedad física, n es el número de átomos de carbono (6,7...45) y $a_1 - a_5$ son los coeficientes dados en la tabla 7.

Tabla 6. Propiedades Físicas Generalizadas (KFW).

Grupo	T _b (°R)	g	M	T _c (°R)	P _c (psia)	w	V _c (pie ³ /lb)
C ₆	607	0.690	84	923	483	0.250	0.06395
C ₇	658	0.727	96	985	453	0.280	0.06289
C ₈	702	0.749	107	1036	419	0.312	0.06264
C ₉	748	0.768	121	1085	383	0.348	0.06258
C ₁₀	791	0.782	134	1128	351	0.385	0.06273
C ₁₁	829	0.793	147	1166	325	0.419	0.06291
C ₁₂	867	0.804	161	1203	302	0.454	0.06306
C ₁₃	901	0.815	175	1236	286	0.484	0.06311
C ₁₄	936	0.826	190	1270	270	0.516	0.06316
C ₁₅	971	0.836	206	1304	255	0.550	0.06325
C ₁₆	1002	0.843	222	1332	241	0.582	0.06342
C ₁₇	1032	0.851	237	1360	230	0.613	0.06350
C ₁₈	1055	0.856	251	1380	222	0.638	0.06362
C ₁₉	1077	0.861	263	1400	214	0.662	0.06372
C ₂₀	1101	0.866	275	1421	207	0.690	0.06384
C ₂₁	1124	0.871	291	1442	200	0.717	0.06394
C ₂₂	1146	0.876	300	1461	193	0.743	0.06402
C ₂₃	1167	0.881	312	1480	188	0.768	0.06408
C ₂₄	1187	0.885	324	1497	182	0.793	0.06417
C ₂₅	1207	0.888	337	1515	177	0.819	0.06431
C ₂₆	1226	0.892	349	1531	173	0.844	0.06438
C ₂₇	1244	0.896	360	1547	169	0.868	0.06443
C ₂₈	1262	0.899	372	1562	165	0.894	0.06454
C ₂₉	1277	0.902	382	1574	161	0.915	0.06459
C ₃₀	1294	0.905	394	1589	158	0.941	0.06468
C ₃₁	1310	0.909	404	1603	143	0.897	0.06469
C ₃₂	1326	0.912	415	1616	138	0.909	0.06475
C ₃₃	1341	0.915	426	1629	134	0.921	0.06480
C ₃₄	1355	0.917	437	1640	130	0.932	0.06489
C ₃₅	1368	0.920	445	1651	127	0.942	0.06490
C ₃₆	1382	0.922	456	1662	124	0.954	0.06499
C ₃₇	1394	0.925	464	1673	121	0.964	0.06499
C ₃₈	1407	0.927	475	1683	118	0.975	0.06506
C ₃₉	1419	0.929	484	1693	115	0.985	0.06511
C ₄₀	1432	0.931	495	1703	112	0.997	0.06517
C ₄₁	1442	0.933	502	1712	110	1.006	0.06520
C ₄₂	1453	0.934	512	1720	108	1.016	0.06529
C ₄₃	1464	0.936	521	1729	105	1.026	0.06532
C ₄₄	1477	0.938	531	1739	103	1.038	0.06538
C ₄₅	1487	0.940	539	1747	101	1.048	0.06540

Fuente: WHITSON, C. H., "Characterizing Hydrocarbon Plus Fractions".²⁰

Tabla 7. Propiedades Físicas Generalizadas (KFW).

q	a₁	a₂	a₃	a₄	a₅
M	-131.11375	24.96156	-0.34079022	2.4941184E-3	468.32575
T _c , °R	915.53747	41.421337	-0.7586859	5.8675351E-3	-1.3028769E3
P _c , psia	275.56275	-12.522269	0.29926384	-2.8452129E-3	1.7117226E3
T _b , °R	434.38878	50.125279	-0.9027283	7.0280657E-3	-601.85651
ω	-0.50862704	8.700211E-2	-1.8484814E-3	1.4663890E-5	1.8518106
γ	0.86714949	3.4143408E-3	-2.839627E-5	2.4943308E-8	-1.1627984
V _c , pie ³ /lb	5.223458E-2	7.8709139E-4	-1.9324432E-5	1.7547264E-7	4.4017952E-2

Fuente: WHITSON, C. H., "Characterizing Hydrocarbon Plus Fractions".²⁰

3.5.2. Método de Riazi & Daubert²² para la caracterización de las propiedades físicas de la fracción pesada. En 1987 los autores propusieron la siguiente correlación, que permite predecir las propiedades físicas de las fracciones pesadas en sistemas hidrocarburos:

$$q = a_1 \left(M_{c_7^+} \right)^{a_2} \left(g_{c_7^+} \right)^{a_3} EXP \left[a_4 M_{c_7^+} + a_5 g_{c_7^+} + a_6 M_{c_7^+} g_{c_7^+} \right] \quad (3.99)$$

Tabla 8. Constantes para la Ecuación 3.99.

q	a₁	a₂	a₃	a₄	a₅	a₆
T _c , °R	544.4	0.2998	1.0555	-1.3478E-4	-0.61641	0
P _c , psia	4.5203E4	-0.8063	1.6015	-1.8078E-3	-0.3084	0
T _b , °R	6.77857	0.401673	-1.58262	3.7741E-3	2.984036	-4.25288E-3
V _c , pie ³ /lb	1.206E-2	0.20378	-1.3036	-2.657E-3	0.5287	2.6012E-3

Fuente: RIAZI, M. R. and DAUBERT, T. E. "Characterization Parameters for Petroleum Fractions".²²

3.5.3. Correlación de Edmister para determinar el Factor Acéntrico (w).

Edmister (1958)²³ propuso una correlación para estimar el factor acéntrico ω de fluidos puros y fracciones de petróleo. La ecuación, que es usada con frecuencia en la industria del petróleo requiere que se conozcan la temperatura de ebullición, la temperatura crítica y la presión crítica. La expresión esta dada por la siguiente relación:

$$w = \frac{3 [\log(P_c/14.7)]}{7 [(T_c/T_b)-1]} - 1 \quad (3.100)$$

Donde la presión crítica P_c esta dada en psia, la temperatura crítica T_c y temperatura de ebullición T_b están dadas en $^{\circ}\text{R}$.

3.6. CARACTERIZACIÓN DE LA FRACCIÓN PESADA (C_{7+})

Generalmente, todos los sistemas de hidrocarburos encontrados contienen una gran cantidad de fracciones que no están bien definidas y no son mezclas de componentes unidos discretamente. Estas fracciones pesadas son a menudo agrupadas como una especie con características propias conocida como C_n^+ .

Una descripción adecuada de las propiedades físicas de estas fracciones y de los componentes no definidos en la mezcla de hidrocarburo, es esencial para obtener cálculos de comportamiento de fases y modelamiento composicional confiables. La información que normalmente se posee de la fracción pesada (C_{7+}) es su peso molecular y su fracción molar en la mezcla y en caso menos frecuente se tienen análisis de laboratorio para obtener la curva de destilación y contenido de parafínicos, nafténicos y aromáticos

(PNA), pero aun así la obtención de las propiedades críticas y el factor acéntrico, fundamentales para aplicar la ecuación de estado, se hace difícil.

Para usar cualquier modelo de predicción de propiedades termodinámicas en la predicción del comportamiento de fases de hidrocarburos, es necesario proporcionar el factor acéntrico, temperatura crítica y presión crítica tanto de la fracción pesada definida como de los componentes no definidos en la mezcla. Se han propuesto numerosos esquemas de caracterización, algunos más ampliamente aceptados que otros.

El comportamiento de fases para aceites volátiles y gases condensados es muy sensible a la composición y propiedades de los componentes más pesados.

El manejo de la fracción pesada tiene dos aspectos importantes, el primero tiene que ver con la forma como se va a considerar ésta; es decir, si se va a tratar como un solo pseudocomponente, si se va a fraccionar en pseudocomponentes de número de carbono simple (NCS) o si se va a tratar como varios pseudocomponentes de número de carbono múltiple (NCM); y el segundo aspecto tiene que ver con los métodos que se apliquen para calcular las propiedades físicas y el factor acéntrico de los pseudocomponentes, ya que pequeñas variaciones en los valores de estas propiedades pueden representar cambios grandes en los resultados de cálculos de fases.

3.6.1. Método de Ahmed para Caracterización de la fracción pesada.²⁴

Ahmed (1985) presenta un método simple y práctico para caracterizar y extender la distribución molar de la fracción C_7^+ a partir de propiedades físicas del heptano-plus que normalmente se encuentran disponibles, tales como el peso molecular y la densidad.

El procedimiento se basa en el cálculo de la fracción molar z_{Cn} a un número de átomos de carbono progresivamente más alto. El proceso de cálculo continúa hasta que la suma de la fracción molar de los pseudocomponentes sea igual a la fracción molar total del heptano-plus (z_{C7+}).

Para obtener la expresión final de z_{Cn} se utiliza la ecuación de balance de materiales de la fracción molar y la ecuación de peso molecular promedio para la fracción pesada:

$$z_{C_n} + z_{C_{(n+1)+}} = z_{C_{n+}} \quad (3.101)$$

$$M_{C_n} \cdot z_{C_n} + M_{C_{(n+1)+}} \cdot z_{C_{(n+1)+}} = M_{C_{n+}} \cdot z_{C_{n+}} \quad (3.102)$$

Solucionando para z_{Cn} se tiene la siguiente expresión final:

$$z_{Cn} = z_{C_{n+}} \left[\frac{M_{C_{(n+1)+}} - M_{C_{n+}}}{M_{C_{(n+1)+}} - M_{C_n}} \right] \quad (3.103)$$

Donde z_{Cn} es la fracción molar del pseudocomponente con n número de átomos de carbono, M_{Cn} es el peso molecular del grupo hidrocarburo con n átomos de carbono, y $M_{C_{n+}}$ es el peso molecular de la fracción C_n^+ calculado por la siguiente expresión:

$$M_{C_{n+}} = M_{C_{7+}} + S(n - 7) \quad (3.104)$$

Donde n es el número de átomos de carbono y S es la pendiente de la ecuación la cual dependiendo del tipo de aceite puede tomar los siguientes valores:

Tabla 9. Pendiente de la ecuación 3.104 para el cálculo de M_{Cn+} .

Número de átomos de carbono	Condensado	Sistemas de Aceite
	S	
n=8	15.5	16.5
n≥9	17.0	20.1

Fuente: AHMED T.H. "A Generalized Correlation for Characterizing the Hydrocarbon Heavy Fractions".²⁴

Para calcular la gravedad específica de la fracción n^+ se tiene la siguiente ecuación:

$$g_{C_{n+}} = g_{C_7^+} \left[1 + ae^{-bn} \left(\frac{M_{C_n^+}}{M_{C_7^+}} - 1 \right) \right] \quad (3.105)$$

Donde γ_{Cn+} es la gravedad específica de la fracción C_{n+} y γ_{C7+} es la gravedad específica del C_7^+ . Los coeficientes a y b dependiendo del tipo de sistema de hidrocarburo empleado son:

Tabla 10. Coeficientes para la ecuación 3.105.

Coeficientes	Condensado	Sistemas de Aceite
a	0.06773	0.247308
b	0.008405	0.063241

Fuente: AHMED T.H. "A Generalized Correlation for Characterizing the Hydrocarbon Heavy Fractions".²⁴

El procedimiento para desarrollar el método es el siguiente:

1. Seleccionar un valor apropiado de S de acuerdo al tipo de sistema de hidrocarburo con el que se esté trabajando.

2. Conociendo el peso molecular de la fracción C_{7+} , con la ecuación 3.104 calcular el peso molecular de la fracción octano-plus (M_{C8+}).
3. Calcular la fracción molar del z_{C7} con la ecuación 3.103.
4. Repetir los pasos 2 y 3 para cada componente en el sistema, hasta que la suma de la fracción molar calculada sea igual a la fracción molar del C_{7+} .

4. SIMULACIÓN DE PRUEBAS PVT

Los estudios PVT de laboratorio y los análisis de comportamiento de fases son necesarios para caracterizar los fluidos de yacimiento a diferentes niveles de presión. Existen diversas pruebas que pueden realizarse dependiendo de la cantidad de datos que se desee obtener.

En el laboratorio experimental PVT se efectúan tres tipos de tipos de pruebas, que permiten verificar las relaciones volumétricas entre las fases presentes en un fluido de yacimiento petrolífero⁵:

1. Pruebas primarias: Estas pruebas involucran la medición de la gravedad específica y la relación gas - aceite del hidrocarburo producido.
2. Pruebas detalladas de laboratorio: Estas pruebas involucran la medición de la composición del fluido, peso molecular, viscosidad, compresibilidad, presión de saturación, factor volumétrico de formación y resultados de la depleción a volumen constante y de la expansión a masa constante.
3. Pruebas de barrido: Son pruebas utilizadas para aplicaciones muy específicas. Si el yacimiento va ser producido bajo un esquema de gas, por ejemplo, este tipo de pruebas puede ser de gran utilidad.

Debido al costo relativamente alto en la realización de una prueba PVT y a la incertidumbre en la precisión de los resultados, las ecuaciones de estado ofrecen una aproximación atractiva para generar este tipo de datos.

En este capítulo se describe la forma como se simula la prueba de Expansión a Masa Constante (E.M.C.). Con los resultados de esta prueba se ajustarán los parámetros de la ecuación de estado utilizando redes neuronales buscando mejorar las predicciones efectuadas por medio de la ecuación de estado.

4.1. PROCEDIMIENTO PARA SIMULAR LA EXPANSIÓN A COMPOSICIÓN CONSTANTE

Los procedimientos de Expansión a Composición Constante, se realizan tanto para gases condensados como para aceites buscando simular las relaciones presión - volumen de los sistemas hidrocarburos. Su propósito principal es la determinación de:

- Presión de saturación (Punto de burbuja o de Rocío).
- Coeficientes de compresibilidad isotérmica del fluido en una fase.
- Factores de compresibilidad de la fase gas.
- Volumen de hidrocarburo total como una función de la presión.

Dentro del procedimiento experimental, se considera importante tener en cuenta que:

- Temperatura inicial = Temperatura de yacimiento.
- Presión inicial > Presión de yacimiento.
- La temperatura se mantiene constante en cada reducción de presión.
- El cambio en volumen de hidrocarburo, V , debe ser registrado para cada cambio de presión.
- La presión de saturación y su volumen correspondiente debe ser registrada.

- El volumen a presión de saturación es usado como volumen de referencia.
- A valores de presión mayores que la de saturación, el volumen de hidrocarburo se registra como un volumen relativo, expresado como:

$$V_{rel} = \frac{V}{V_{sat}} \quad (4.1)$$

Donde: V_{rel} : Volumen relativo
 V : Volumen del sistema hidrocarburo
 V_{sat} : Volumen a presión de saturación.

- Por debajo de la presión de saturación, el volumen total (V_t) de las dos fases debe ser registrado, y el volumen relativo se expresa como:

$$V_{rel} = \frac{V_t}{V_{sat}} \quad (4.2)$$

- Por encima de la presión de saturación el coeficiente de compresibilidad isotérmico es determinado a partir de la expresión como:

$$C = -\frac{1}{V_{rel}} \left[\frac{\partial V_{rel}}{\partial P} \right]_T \quad (4.3)$$

donde C es el coeficiente de compresibilidad isotérmico en psi^{-1} .

- Además, para sistemas de gas condensado, el factor de compresibilidad debe ser determinado.

4.1.1. Algoritmo para simular la prueba. Los siguientes son los pasos a seguir para la simulación utilizando la ecuación de estado de Peng-Robinson:

Paso 1: Dada la composición total del sistema hidrocarburo, z_i , y la presión de saturación (P_b para sistemas de aceite ó P_d para gases), calcular el volumen total ocupado por una mol de sistema. Este volumen corresponde al volumen de referencia V_{sat} (Volumen a presión de saturación). Matemáticamente este valor se calcula a partir de la relación:

$$V_{sat} = \frac{(1) Z R T}{P_{sat}} \quad (4.4)$$

Donde: V_{sat} = Volumen a presión de saturación (ft³/mol)
 P_{sat} = Presión de saturación (psia)
 T = Temperatura del sistema, °R
 Z = Factor de compresibilidad utilizando la ecuación 3.5

Paso 2: En cada cambio de presión, determinar el factor de compresibilidad Z , resolviendo la ecuación 6.5, para usarlo en la determinación del volumen de fluido:

$$V = \frac{(1) Z R T}{P} \quad (4.5)$$

Donde: V = Volumen de líquido o gas comprimido, ft³/ mol
 Z = Factor de compresibilidad del líquido o gas comprimido.
 P = Presión del sistema, $P > P_{sat}$, psia.
 R = Constante del gas

Calcular el volumen relativo correspondiente utilizando la ecuación 4.1.

Paso 3: Cuando la presión sea menor que la presión de saturación, calcular las constantes de equilibrio y realizar el correspondiente cálculo flash. Los datos resultantes incluyen K_i , x_i , y_i , n_v , Z^V y Z^L . Ya que no se ha removido ningún fluido de la celda, el número de moles original ($n_i = 1$) y la

composición z_i permanecen constantes. El volumen de las fases líquido y gas pueden ser calculados a partir de las siguientes expresiones:

$$V_L = \frac{(1) (n_L) Z^L R T}{P} \quad (4.6)$$

$$V_g = \frac{(1) n_V Z^V R T}{P} \quad (4.7)$$

$$V_t = V_L + V_g \quad (4.8)$$

Donde: n_L y n_V = moles de gas y líquido obtenidas del cálculo flash.
 Z^L y Z^V = Factores de compresibilidad de gas y líquido obtenidos de la ecuación 3.5.
 V_t = Volumen total del sistema hidrocarburo.

Paso 4: Calcular el volumen total relativo utilizando la ecuación 4.2.

Este procedimiento se debe efectuar a cada paso de presión hasta que se alcance la presión mínima del sistema, generalmente asociada a la presión del primer tren de separación.

Los valores numéricos de los volúmenes predichos por la Ecuación de Estado se calculan a partir de un flash isotérmico efectuado a la presión de cada etapa, dichos cálculos se describen en el capítulo dos del presente proyecto.

5. AJUSTE DE ECUACIONES DE ESTADO

En yacimientos de hidrocarburos donde la presión y la temperatura se encuentran en una región lo suficientemente cercana al punto crítico de la mezcla, pequeñas variaciones en la presión o en la composición - variables termodinámicas - repercuten en cambios significativos en el estado del sistema. En consecuencia las cantidades de los productos en superficie y las propiedades que determinan su calidad y precio – gravedad API, GOR – también se ven considerablemente afectadas.

La simulación de este tipo de yacimientos debe ser composicional, esto es, auxiliada por una expresión que determine las condiciones de equilibrio termodinámico entre las fases, como lo es una ecuación de estado - EdE -. No obstante, la expresión original de cualquier EdE es insuficiente para describir acertadamente el comportamiento de una mezcla de sustancias, por lo que se hace necesaria la modificación de ciertos parámetros de dicha ecuación o del fluido con el fin de representar con mayor exactitud los datos experimentales obtenidos de una muestra representativa del yacimiento.

Se requiere de una EdE para la estimación inicial de reservas así como la predicción de las cantidades de los fluidos recuperables bajo un esquema dado de producción, buscando determinar la estrategia más apropiada para la óptima explotación del yacimiento. Si la EdE no ha sido convenientemente ajustada – respetando los criterios termodinámicos en los cuales se basa –, las estrategias seleccionadas no serán las adecuadas para la economía del proyecto. Por tal razón el uso de una EdE ajustada otorgará mayor confiabilidad en las predicciones y certeza en la toma de decisiones. La EdE ajustada también permite describir escenarios de recuperación

mejorada del yacimiento como desplazamientos con nitrógeno, gas pobre o dióxido de carbono, por medio de la simulación.

En este capítulo se mostraran las metodologías tradicionales de ajuste de parámetros de ecuaciones de estado, con sus ventajas y sus fallas.

5.1. ANTECEDENTES

El proceso de ajuste ha sido utilizado con anterioridad para mejorar las predicciones realizadas con ecuaciones de estado y lograr reproducir el comportamiento real de mezclas de hidrocarburos de yacimiento. Aunque los trabajos son diversos, la mayoría se fundamenta en la minimización de una función error aplicando métodos numéricos y estadísticos para variar propiedades de componentes como la fracción pesada, coeficientes de interacción binaria - CIB - o parámetros propios de las ecuaciones de estado.

Entre estos trabajos se encuentra el realizado por Coats y Smart²⁵ donde se presenta un programa de ajuste de parámetros - Ω_a y Ω_b de cualquier componente así como los CIB -, y se busca reproducir simultáneamente los diferentes tipos de información experimental disponible para un fluido. Se utilizan las EdE de Zudkevitch-Joffe-Redlich-Kwong y la de Peng-Robinson. En este caso la función a minimizar E viene dada por:

$$E(\bar{I}) = \sum_{i=1}^{ND} W_i \left| \frac{(d_i - dc_i)}{d_i} \right| \quad (5.1)$$

Donde \bar{I} el conjunto de variables de regresión, d_i es el dato experimental, dc_i el dato calculado y W_i el factor de peso que se asigna a la propiedad a la cual

se esté aplicando la ecuación y ND el número de datos experimentales con que se cuenta. El método de minimización que se usa es una extensión de la técnica de mínimos cuadrados, donde se definen regiones o rangos para cada variable y a cada iteración estas se acotan en subregiones cada vez más pequeñas hasta lograr que todas las variables alcancen un valor estable permaneciendo en la última subregión alcanzada. Es importante anotar que en este estudio se concluye que los parámetros que son más importantes y deben estar presentes en cualquier ajuste son: Ω_a y Ω_b del metano y la fracción pesada junto con el CIB entre estos componentes.

Otro estudio realizado acerca de técnicas de regresión para representar el comportamiento de fases por Agarwal et al²⁶ utilizaron un proceso de selección y descarte de variables, que les permite ser tenidas en cuenta o no como parámetros de ajuste dependiendo de la sensibilidad que a éstas presente la función objetivo, - función a minimizar -, durante el ajuste. Dentro del conjunto de posibles parámetros de regresión se encuentran las propiedades críticas de los componentes y los CIB. La función objetivo utilizada es:

$$E(\bar{I}) = \sum_{i=1}^{ND} \left| W_i \frac{(dc_i - d_i)}{d_i} \right|^2 \quad (5.2)$$

Para minimizar la función objetivo se utiliza una modificación del algoritmo de mínimos cuadrados de Dennis y colaboradores que trabaja con conceptos de optimización no lineal para seleccionar la dirección y el incremento de las variables. La EdE utilizada para realizar los cálculos fue la de Peng-Robinson.

Da Silva y Rodríguez²⁷ realizan un estudio donde se presenta una metodología diferente para hacer el ajuste. Se divide el proceso en dos

partes, una inicial donde mediante la variación del CIB entre el metano y la fracción pesada se reproduce la presión de saturación y posteriormente se aplica un proceso de regresión no lineal utilizando la técnica de mínimos cuadrados ajustando los valores de Ω_a y Ω_b del metano y la fracción pesada hasta lograr reproducir con el menor error posible los datos experimentales de una fase. La función objetivo en este caso es:

$$E(\vec{I}) = \sum_{i=1}^{ND} \left| \frac{dc_i - d_i}{d_i} \right|^2 \quad (5.3)$$

Los cálculos se llevan a cabo con la EDE de Peng-Robinson. Mediante esta técnica se logran reproducir los datos experimentales satisfactoriamente pero se presentan desviaciones en el cálculo de la presión de saturación.

Este inconveniente fue posteriormente superado por Trujillo²⁸ que utiliza en su metodología el mismo esquema de Da Silva y Rodríguez para llevar a cabo el ajuste, pero lo hace en forma iterativa. Incluye además un conjunto alternativo de variables de regresión conformado por las propiedades críticas: P_c , T_c y el factor acéntrico de la fracción pesada. Trujillo trabaja con la siguiente función objetivo:

$$E(\vec{I}) = \sum_{i=1}^{ND} (dc_i - d_i)^2 \quad (5.4)$$

En este trabajo se utilizan datos experimentales de las regiones monofásica y bifásica obteniéndose resultados similares. Es importante apuntar que el tiempo de cómputo empleado al utilizar datos de la región bifásica se incrementa notablemente debido a complejidad de los cálculos de equilibrio.

En un estudio posterior San Miguel²⁹ retoma la metodología de Da Silva y Rodríguez aplicándola en forma iterativa pero utilizando técnicas de

minimización diferentes a la de mínimos cuadrados, como son: mínimo valor absoluto (LAV) y mínimo valor absoluto modificado (MLAV), para cada una de las cuales utiliza una función objetivo diferente. Para la técnica LAV se minimiza la sumatoria de las desviaciones de los valores absolutos:

$$E(\vec{I}) = \sum_{i=1}^{ND} |d_i - dc_i| \quad (5.5)$$

Mientras que para la técnica MLAV se minimiza:

$$E(\vec{I}) = \sum_{i=1}^{ND} W_i (d_i - dc_i)^2 \quad (5.6)$$

$$W_i = \frac{1}{|d_i - dc_i|} \quad (5.7)$$

El ajuste se lleva a cabo con datos experimentales de la región monofásica y bifásica alcanzando resultados satisfactorios con pequeñas desviaciones y reduciendo el tiempo de cómputo empleado cuando se trabaja en la zona de dos fases.

5.2. METODOLOGÍA DEL AJUSTE

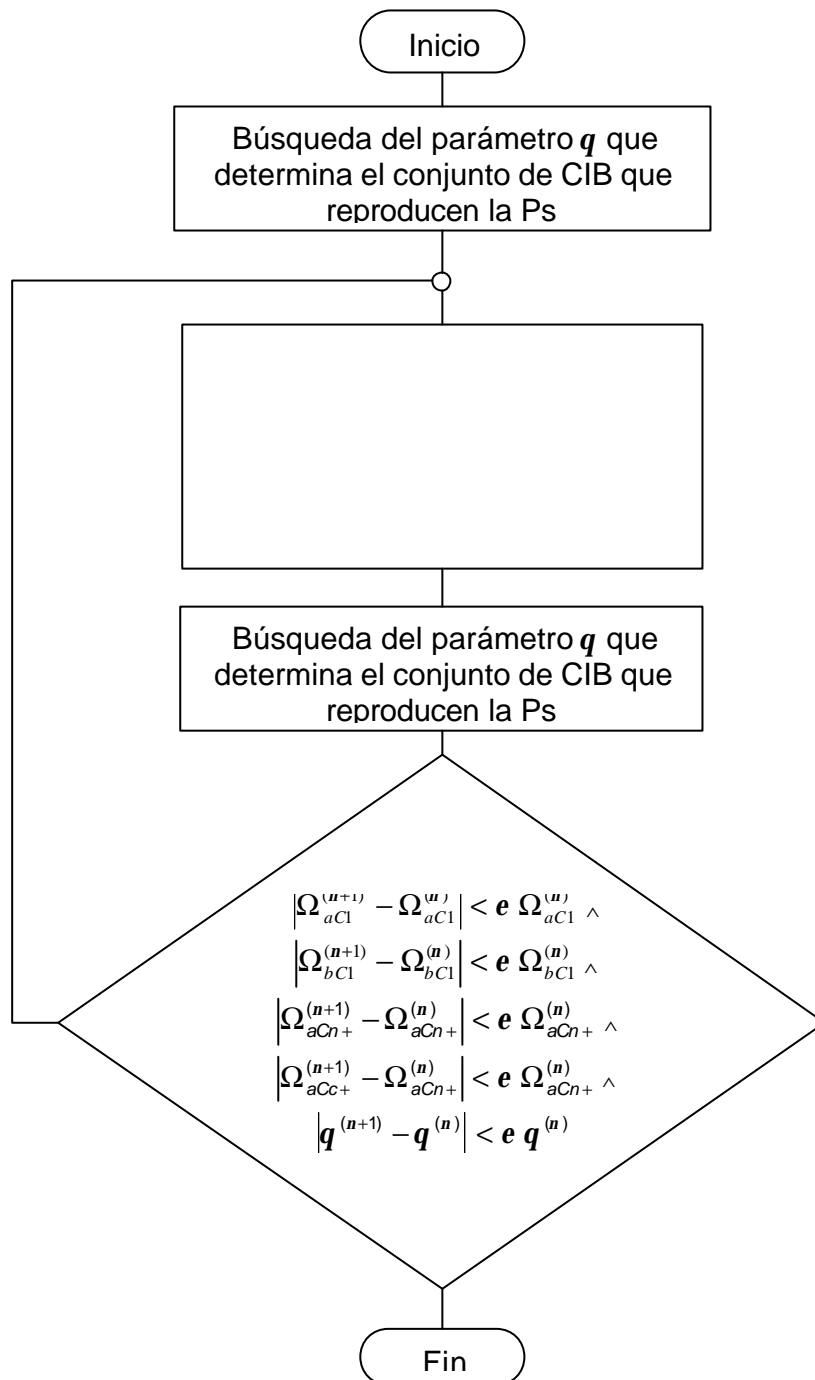
El proceso de ajuste puede ser iniciado a partir de cualquiera de las metodologías anteriormente mencionadas. En este caso dará como punto de partida el esquema utilizado por Da Silva y Rodríguez, efectuando las modificaciones necesarias para optimizar y aplicar el método cuando se trabaja con fluidos cercanos al punto crítico.

El proceso consta entonces de dos partes claramente definidas e importantes, una donde se busca el conjunto de CIB que conducen a reproducir con la menor desviación posible la presión de saturación de la mezcla en estudio y otra en la que se modifican los valores de Ω_a y Ω_b para el metano (C_1) y la fracción pesada (C_{n+}) mediante regresión múltiple, de tal forma que se minimice la desviación de los valores calculados de una propiedad del fluido con la EdE respecto a los datos experimentales. En este trabajo, la propiedad con la que se realiza el proceso de regresión es el factor de compresibilidad Z de la región monofásica con datos obtenidos de pruebas EMC.

El procedimiento descrito anteriormente se hace en forma repetitiva hasta lograr que todas las variables utilizadas alcancen un valor estable, es decir que su variación no exceda cierta tolerancia, definida como un valor proporcional ϵ de la variable, - se sugiere para ϵ un valor de 5×10^{-4} -. Esta tolerancia es generalmente alcanzada con suficiencia cuando se termina el ajuste y al ser dependiente del valor que tenga la variable en la iteración inmediatamente anterior garantiza que estas en realidad se hayan estabilizado, lo cual puede no ocurrir al asignar a esta tolerancia un valor fijo. En la figura 13 se presenta el diagrama de flujo del proceso.

La principal característica de esta metodología (propuesta por Da Silva y Rodríguez) y que la diferencia de las anteriores, es que para lograr reproducir la presión de saturación, no se emplea tan solo el CIB entre el C_1 y el C_{n+} sino la totalidad de los coeficientes, lo cual se hace posible mediante la búsqueda de un parámetro óptimo q , al cual están sujetos los valores de todos los CIB. La forma por medio de la cual se hacen depender los CIB de este parámetro puede estar dada tanto por el esquema desarrollado por Slot-Petersen, - basado en pesos moleculares -, como por el esquema de Chueh-Prausnitz, basado en volúmenes críticos. Estos esquemas son los siguientes:

Figura 13. Diagrama de flujo del proceso de ajuste de EdE de Peng-Robinson.



Fuente: DA SILVA M. B., RODRIGUEZ, F., "Automatic fitting of equation of state for phase behavior matching".²⁷

- ✦ **Esquema de Slot-Petersen¹¹.** Este método para estimar los CIB se fundamenta en las diferencias relativas entre los pesos moleculares de los componentes presentes en la mezcla. Además, todos los coeficientes están sujetos en forma proporcional al valor del coeficiente entre el metano y la fracción pesada δ_{C1-Cn+} .

Los pasos para calcular los CIB entre parejas de hidrocarburos son los siguientes:

- Adoptar un valor para el δ_{C1-Cn+}
- Calcular los CIB entre los demás componentes (C2, C3, ..., Cn-1) y la fracción pesada mediante la siguiente expresión, donde k es el número de átomos de carbono del componente

$$d_{Ck-Cn+} = r d_{C(k-1)-Cn+} \quad (5.8)$$

- Calcular los demás CIB utilizando la siguiente relación

$$d_{i-j} = d_{i-Cn+} \cdot \frac{(Mw_j)^n - (Mw_i)^n}{(Mw_{Cn+})^n - (Mw_i)^n} \quad (5.9)$$

Los valores de r y n son escogidos de tal forma que permitan reproducir la presión de saturación. El factor r puede variar entre cero y uno mientras el exponente n entre uno y diez.

- ✦ **Esquema de Chueh-Prausnitz³⁰.** Este método a diferencia del anterior está basado en las diferencias existentes entre los volúmenes críticos de los componentes presentes en la mezcla.

La expresión que se propone para calcular los coeficientes entre parejas de componentes hidrocarburos es:

$$d_{i-j} = A \left[1 - \left(\frac{2(V_{C_i} - V_{C_j})^{1/6}}{V_{C_i}^{1/3} - V_{C_j}^{1/3}} \right)^B \right] \quad (5.10)$$

Los valores de A y B se escogen para reproducir de la mejor forma el comportamiento real, y los autores sugieren los siguientes valores: para el factor A de 0.15 a 0.25 y el exponente B puede tomarse inicialmente como 6 pero se puede modificar.

Se recomienda también que el proceso termine con el procedimiento de ajuste de la presión de saturación para asegurar una desviación mínima en su cálculo, ya que es de vital importancia al momento de hacer predicciones de las demás propiedades del fluido en la zona de dos fases con la EdE.

5.2.1. Planteamiento del ajuste de los CIB. Teniendo en cuenta las expresiones para el cálculo de la presión de saturación de un fluido, donde se tiene un conjunto de N_c+1 incógnitas conformado por fracción molar de cada uno de los componentes en la fase infinitesimal y la presión de saturación del fluido, se plantea un esquema similar del tipo Newton-Raphson, donde se reemplaza la variable presión de saturación, por el parámetro q , que regula el comportamiento de los CIB, ya que la presión de saturación es un dato experimental conocido y necesario para desarrollar el ajuste.

Como en todo proceso iterativo se necesitan unos estimados iniciales para cada variable, en el caso de las composiciones de la fase infinitesimal, estos se obtienen a partir de la composición original de la muestra y usando las constantes de equilibrio correlacionadas mediante el método de Wilson o el método de Varotsis¹⁷.

El valor estimado inicial de q se obtiene a partir de un valor correlacionado para el CIB entre el C_1 y C_{n+} . En la literatura se encuentran expresiones desarrolladas para tal fin por diversos autores, Amhed¹⁵ propone una expresión basado en la temperatura del sistema,

$$d_{C1-Cn+} = 0.00189(T - 460) - 0.297659 \quad (5.11)$$

Katz y Firoozabadi presentan en su trabajo³¹ una correlación para estimarlo a partir de la densidad de la fracción pesada a 60°F.

$$d_{C1-Cn+} = 0.1275r_{C+} - 0.0579 \quad (5.12)$$

El parámetro q tiene un significado diferente dependiendo del esquema de cálculo de CIB que se utilice. Si se trabaja con el esquema de Slot-Petersen q representará el valor del CIB entre el metano y la fracción pesada d_{C1-Cn+} , - del cual dependen los demás CIB -, mientras que si se prefiere el método de Chueh-Prausnitz, q representa el valor del factor A en la expresión (ecuación 5.10).

Una vez determinado d_{C1-Cn+} el valor inicial del parámetro q se obtiene para cada esquema de cálculo de CIB en forma diferente. Para Slot-Petersen se tiene:

$$q^{(0)} = d_{C1-Cn+} \quad (5.13)$$

Por otra parte para Chueh-Prausnitz se tiene:

$$q^{(0)} = \frac{d_{C1-Cn+}}{\left[1 - \left(\frac{2(V_{C1} - V_{Cn+})^{1/6}}{V_{C1}^{1/3} - V_{Cn+}^{1/3}} \right)^B \right]} \quad (5.14)$$

De esta forma, con los valores iniciales de cada variable se aplica el siguiente esquema iterativo de solución para N_C+1 variables, donde a cada iteración se debe resolver el siguiente sistema lineal de ecuaciones.

$$[J]^v [\Delta I]^{v+1} = -[F]^v \quad (5.15)$$

Donde J es la matriz jacobiana, F el vector de funciones de residuos y ΔI el vector de incógnitas que representa el incremento correspondiente a cada variable. A cada iteración se resuelve el siguiente sistema:

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_1}{\partial x_n} & \frac{\partial F_1}{\partial q} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \dots & \frac{\partial F_2}{\partial x_n} & \frac{\partial F_2}{\partial q} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \dots & \frac{\partial F_n}{\partial x_n} & \frac{\partial F_n}{\partial q} \\ \frac{\partial F_{n+1}}{\partial x_1} & \frac{\partial F_{n+1}}{\partial x_2} & \dots & \frac{\partial F_{n+1}}{\partial x_n} & \frac{\partial F_{n+1}}{\partial q} \end{bmatrix}^v \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \\ \Delta q \end{bmatrix}^{v+1} = - \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \\ F_{n+1} \end{bmatrix}^v \quad (5.16)$$

Donde las N_C primeras funciones de residuo están dadas por la diferencia de las fugacidades entre el líquido y vapor para cada componente y como función de residuo N_C+1 se utiliza la ecuación

$$\sum_i \left(\frac{f_i^P}{\Phi_i^N} \right) - P_{sat} \quad (5.17)$$

Los superíndices P y N hacen referencia a las fases presente e infinitesimal respectivamente. Las derivadas de las funciones de residuo respecto al parámetro q se evalúan numéricamente, las demás en forma analítica.

De igual manera que en el proceso del cálculo de la presión de saturación, cada c_i representa la fracción molar de los componentes en la fase infinitesimal y mediante el proceso iterativo se buscan aquellos valores que satisfacen la condición termodinámica de equilibrio. Por otra parte el parámetro q determina a cada iteración un nuevo conjunto de CIB entre parejas de hidrocarburos. Al finalizar el proceso iterativo se han determinado las composiciones y los CIB que satisfacen el equilibrio y además permiten reproducir la presión de saturación.

Este proceso iterativo termina cuando se cumplen 3 condiciones simultáneamente:

- Norma euclidiana de las funciones de residuo menor que una tolerancia ϵ_1 . Normalmente se utiliza un valor de 10^{-6} .
- Valor de la función de residuo N_{c+1} (ecuación 3.33) menor que una tolerancia ϵ_2 . Normalmente se utiliza un valor de 10^{-2} .
- Porcentaje de variación del valor del parámetro q inferior a una tolerancia ϵ_3 . Normalmente se utiliza un 2%.

La imposición de las tres condiciones garantiza en la mayoría de los casos encontrar soluciones adecuadas, ya que pueden hallarse soluciones triviales en las que las K_i y las relaciones de fugacidad tienden a la unidad. En este caso aunque la primera condición se cumple la variación del parámetro no es uniforme. En la figura 14 se presenta el diagrama de flujo del proceso de ajuste de los CIB para reproducir la presión de saturación.

5.2.2. Método de regresión múltiple. El procedimiento de ajuste de los parámetros Ω_a y Ω_b del metano y la fracción pesada se lleva a cabo mediante el empleo de la técnica de mínimos cuadrados no lineal, donde se busca minimizar la sumatoria de las diferencias al cuadrado entre los valores

calculados por la EdE y los datos experimentales. Siguiendo la anterior definición la función a minimizar $E(\vec{I})$ es:

$$E(\vec{I}) = \sum_{m=1}^{ND} \left[\frac{(dc_m - d_m)}{d_m} \right]^2 \quad (5.18)$$

Donde dc_m es el valor de la propiedad calculado con la EdE de Peng-Robinson, d_m es el dato experimental, N_D el número de datos con que se cuenta y el vector \vec{I} está conformado por las variables de regresión:

$$\vec{I} = (I_1, I_2, I_3, I_4)^T \quad (5.19)$$

Donde

$$I_1 = \Omega_{aC1}$$

$$I_2 = \Omega_{bC1}$$

$$I_3 = \Omega_{aCn+}$$

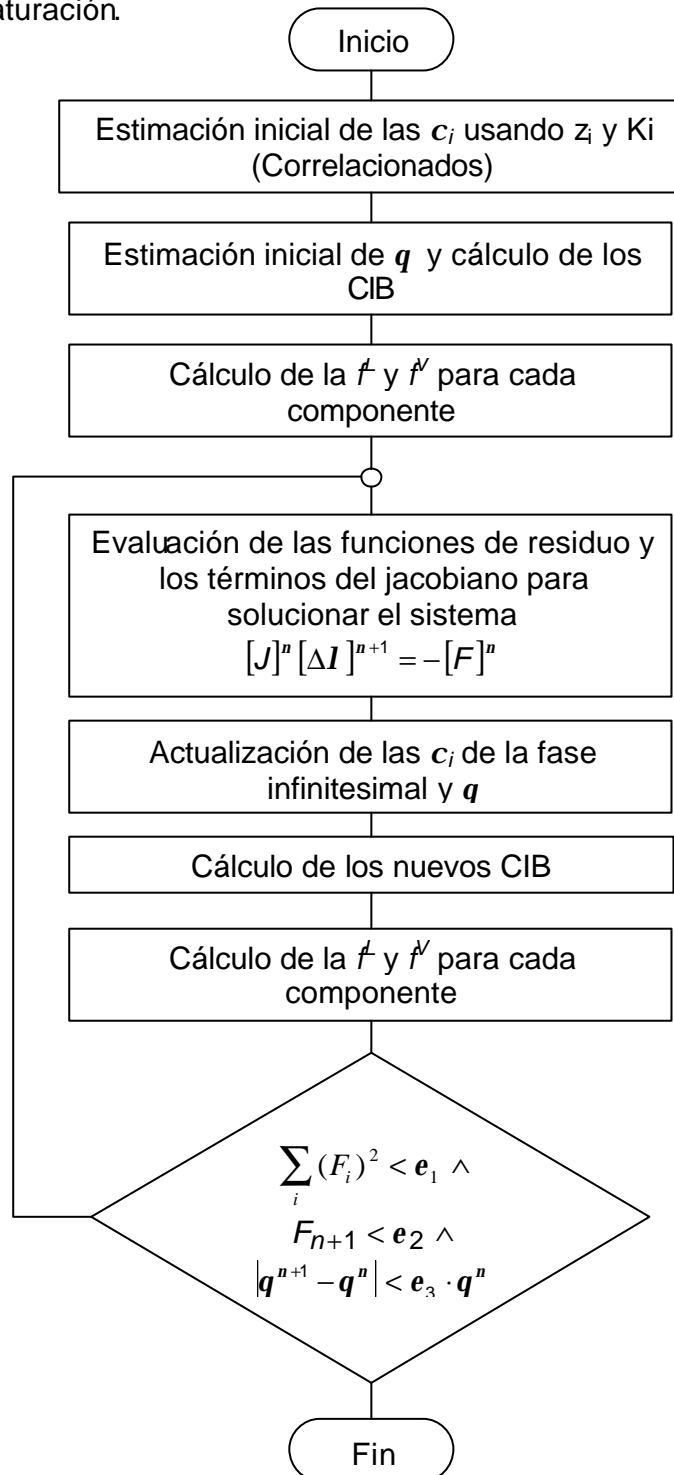
$$I_4 = \Omega_{bCn+}$$

Para obtener un mínimo de la función $E(\vec{I})$ es condición necesaria que su primera derivada respecto a cada variable del vector \vec{I} sea igual a cero, es decir:

$$\frac{\partial E(\vec{I})}{\partial I_k} = 0; k = 1, 2, 3, 4 \quad (5.20)$$

El conjunto de ecuaciones representadas por la ecuación 5.20 conforman un sistema de 4 ecuaciones no lineales en Ω_{aC1} , Ω_{bC1} , Ω_{aCn+} , Ω_{bCn+} . Este sistema puede resolverse aplicando el método de Newton-Raphson, para lo cual se definen las siguientes funciones de residuo:

Figura 14. Diagrama del proceso de búsqueda de los CIB que ajustan la Presión de saturación.



Fuente: DA SILVA M. B., RODRIGUEZ, F., "Automatic fitting of equation of state for phase behavior matching".²⁷

$$F_k = \frac{\partial E(\vec{I})}{\partial I_k} = 0; k= 1, 2, 3, 4 \quad (5.21)$$

Y se tiene el siguiente esquema iterativo de solución:

$$\begin{bmatrix} \frac{\partial F_1}{\partial I_1} & \frac{\partial F_1}{\partial I_2} & \frac{\partial F_1}{\partial I_3} & \frac{\partial F_1}{\partial I_4} \\ \frac{\partial F_2}{\partial I_1} & \frac{\partial F_2}{\partial I_2} & \frac{\partial F_2}{\partial I_3} & \frac{\partial F_2}{\partial I_4} \\ \frac{\partial F_3}{\partial I_1} & \frac{\partial F_3}{\partial I_2} & \frac{\partial F_3}{\partial I_3} & \frac{\partial F_3}{\partial I_4} \\ \frac{\partial F_4}{\partial I_1} & \frac{\partial F_4}{\partial I_2} & \frac{\partial F_4}{\partial I_3} & \frac{\partial F_4}{\partial I_4} \end{bmatrix}^v \begin{bmatrix} \Delta I_1 \\ \Delta I_2 \\ \Delta I_3 \\ \Delta I_4 \end{bmatrix}^{v+1} = - \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix}^v \quad (5.22)$$

Las expresiones correspondientes a las funciones de residuo son:

$$\frac{\partial E(\vec{I})}{\partial I_k} = 2 \sum_{m=1}^{nd} \frac{dc_m - d_m}{d_m} \cdot \frac{\partial dc_m}{\partial I_k} \quad (5.23)$$

Mientras que para los términos del Jacobiano se tiene:

$$J_{ij} = \frac{\partial F_i}{\partial I_j} \quad (5.24)$$

$$\frac{\partial F_i}{\partial I_j} = 2 \sum_{m=1}^{ND} \left[\frac{dc_m - d_m}{d_m} \cdot \frac{\partial}{\partial I_j} \left(\frac{\partial dc_m}{\partial I_k} \right) + \frac{1}{d_m} \left(\frac{\partial dc_m}{\partial I_j} \right) \left(\frac{\partial dc_m}{\partial I_k} \right) \right] \quad (5.25)$$

Como se puede apreciar para evaluar tanto en la ecuación 5.24 como la 5.25, se necesita conocer el valor de la primera y segunda derivada de la propiedad calculada con la EDE respecto a cada una de las variables de regresión. Estas derivadas pueden obtenerse en forma numérica o en forma analítica según el grado de dificultad que presente.

Los valores de las variables de regresión al iniciar el proceso iterativo son los propuestos originalmente para la EdE³²:

$$\Omega_{aC1} = 0.45723552\ 8921$$

$$\Omega_{bC1} = 0.0777960739039$$

$$\Omega_{aCn+} = 0.457235528921$$

$$\Omega_{bCn+} = 0.07779607\ 39039$$

A cada iteración se resuelve el sistema y se actualizan los valores de las variables de regresión. Este proceso termina cuando se cumple cualquiera de las siguientes condiciones:

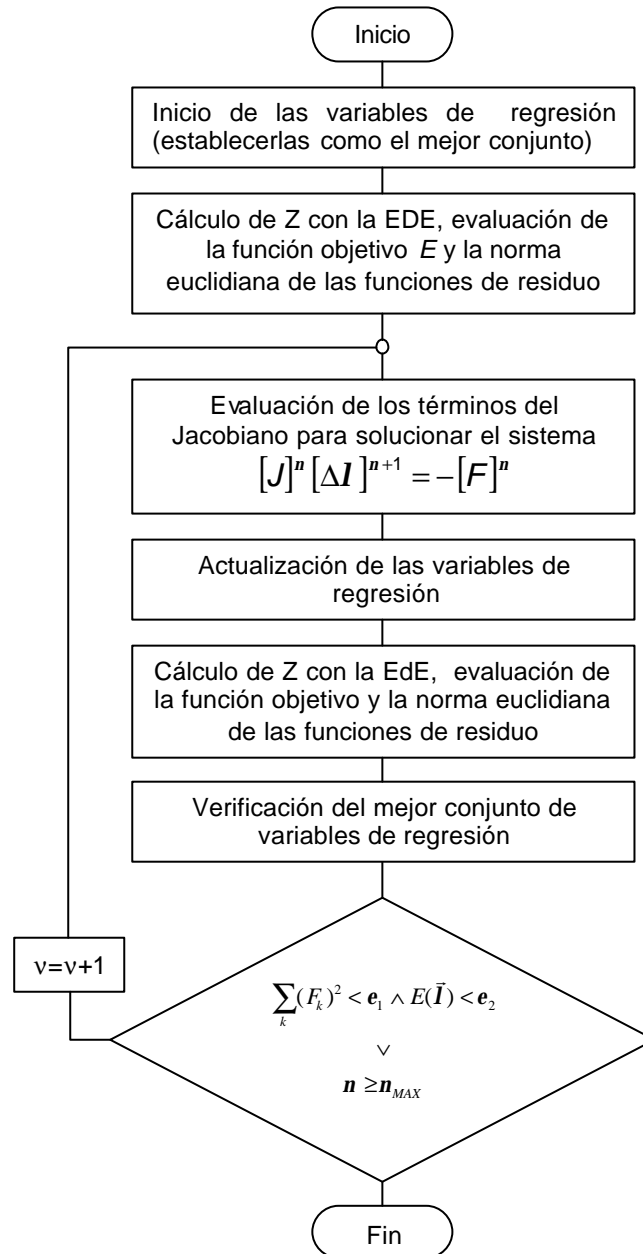
- Norma euclidiana de las funciones de residuo menor que una tolerancia ε_1 , - fijada en 10^{-6} -, y la función objetivo menor que una tolerancia ε_2 . Se sugiere 2×10^{-5} .
- Después de un determinado numero de iteraciones - v_{MAX} -. Se utilizan máximo15.

Estas condiciones han sido probadas con resultados satisfactorios, en el caso de la primera, generalmente se cumple cerca de la sexta iteración, alcanzándose siempre la tolerancia ε_1 primero, mientras que la tolerancia ε_2 suele tardar algunas iteraciones más. En tal caso si el número de iteraciones supera el valor máximo definido sin alcanzarse convergencia, se aplica la segunda condición de salida y el proceso se detiene. En este último caso se escoge el conjunto de valores con el cual el error absoluto promedio (EAP) de los datos calculados respecto a los datos experimentales haya alcanzado el valor más bajo. El EAP se define como:

$$EAP = \frac{\sum_{m=1}^{ND} \frac{dc_m - d_m}{d_m}}{ND} \quad (5.26)$$

Este proceso de ajuste se describe en la figura 15.

Figura 15. Diagrama del proceso de regresión de variables.



Fuente: Autor.

6. REDES NEURONALES ARTIFICIALES

Las redes neuronales son un elemento importante de las denominadas tecnologías de Inteligencia Artificial (IA). La IA es "la disciplina científica y técnica que se ocupa del estudio de las ideas que permiten ser inteligentes a los ordenadores" (definición de H. Winston)³³.

A pesar de que el objetivo final de la IA, dotar de auténtica inteligencia a las máquinas, queda todavía muy lejos (e incluso hay autores que defienden que esto nunca será posible), la ciencia de la Inteligencia Artificial ha generado numerosas herramientas prácticas, entre las que se encuentran las redes neuronales. Las redes neuronales, también llamadas "redes de neuronas artificiales", son modelos bastante simplificados de las redes de neuronas que forman el cerebro. Y, al igual que este, intentan "**aprender**" a partir de los datos que se le suministran.

En el desarrollo de este capítulo se esbozarán los conceptos básicos acerca del funcionamiento de las redes neuronales, sus diferentes características y la forma como se deben entrenar para poder ser aplicadas a problemas de ingeniería, con el fin de utilizar la capacidad que tiene estas de clasificar conjuntos de datos de acuerdo a patrones específicos y previamente dados a conocer, para con esta información clasificar y predecir un comportamiento nuevo y específico.

6.1. BREVE HISTORIA³³

La era moderna de las redes neuronales se cree que comenzó en 1943 gracias a McCullough y Pitts, quienes describieron el cálculo lógico de las redes neuronales y perfilaron el primer módulo formal de una neurona elemental.

En la década de los cincuenta, Minsky comienza a construir la primera neurocomputadora (basada en modelos de redes neuronales que imitan al cerebro).

En 1962, Frank Rosenblatt presenta los resultados de una máquina a la que denominó "Perceptrón", la cual reproducía una estructura neuronal muy simplificada, capaz de aprender a reconocer y clasificar determinadas figuras.

En la misma década, Minsky y Pappert (autoridades de la IA clásica) publicaron un libro en el que se ponían de manifiesto las limitaciones de los perceptrones de una capa. Esto hará que se pierda interés en el campo de las redes neuronales hasta la década de los 80, en que el estudio de nuevas arquitecturas de redes y la mayor potencia de los ordenadores permiten el diseño de redes muy eficientes en tareas en las que otros procedimientos de tipo simbólico encuentran dificultades.

Así, en 1982 J. Hopfield describe las "redes de Hopfield" en las que se utilizan funciones de energía para entender las redes dinámicas. Cohen y Grossberg desarrollan en el 83 el principio de la memoria direccional. En 1986 Rumulhart, Hinton y Williams redescubren el algoritmo de "back-propagation" (desarrollado en 1974 por Paul Werbor) para el aprendizaje de

redes neuronales. Por estas fechas, y gracias a las nuevas tecnologías de fabricación de microchips, comienzan a construirse redes neuronales implementadas en silicio (mucho más rápidas que las de software).

Actualmente, el uso de redes neuronales se ha extendido bastante en el mercado de software doméstico, dejando de estar restringidas a los entornos de investigación y a las grandes empresas. De esta forma, se pueden encontrar modelos de redes neuronales en programas de reconocimiento de voz, en juegos de ordenador, programas de contabilidad, tutores, y muchos otros³³.

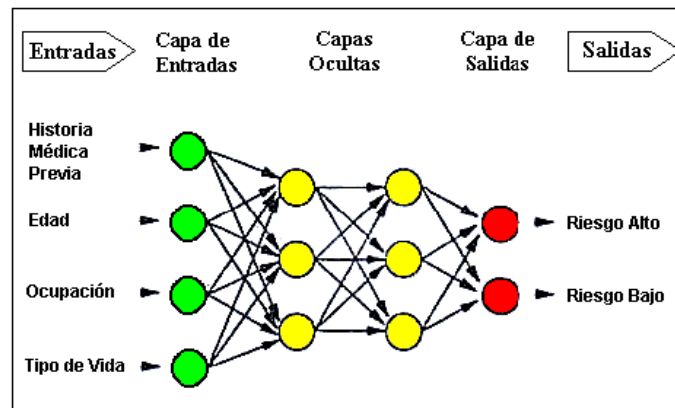
6.2. FUNCIONAMIENTO BÁSICO³⁴

Las redes neuronales están formadas por un conjunto de neuronas artificiales interconectadas. Las neuronas de la red se encuentran distribuidas en diferentes capas de neuronas, de manera que las neuronas de una capa están conectadas con las neuronas de la capa siguiente, a las que pueden enviar información.

La arquitectura más utilizada, en la actualidad, de una red neuronal (como la presentada en la figura 16) consistiría en:

- Una primera capa de entradas, que recibe información del exterior.
- Una serie de capas ocultas (intermedias), encargadas de realizar el trabajo de la red.
- Una capa de salidas, que proporciona el resultado del trabajo de la red al exterior.

Figura 16. Estructura de una red neuronal.



Fuente: Autor

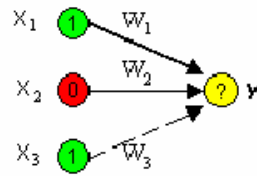
El número de capas intermedias y el número de neuronas de cada capa dependerá del tipo de aplicación al que se vaya a destinar la red neuronal.

6.2.1. Neuronas y conexiones. Cada neurona de la red es una **unidad de procesamiento** de información; es decir, recibe información a través de las conexiones con las neuronas de la capa anterior, procesa la información, y emite el resultado a través de sus conexiones con las neuronas de la capa siguiente, siempre y cuando dicho resultado supere un valor "umbral".

En una red neuronal ya entrenada, las conexiones entre neuronas tienen un determinado peso ("peso sináptico").

Un ejemplo de una neurona sobre la que convergen conexiones de diferente peso sináptico (W_i) sería el de la Figura 17.

Figura 17. Pesos sinápticos de una neurona .



Fuente: Autor

El procesamiento de la información llevado a cabo por cada neurona Y, consiste en una función (F) que opera con los valores recibidos desde las neuronas de la capa anterior (X_i , generalmente 0 o 1), y que tiene en cuenta el peso sináptico de la conexión por la que se recibieron dichos valores (W_i). Así, una neurona dará mas importancia a la información que le llegue por una conexión de peso mayor que no a aquella que le llegue por una conexión de menor peso sináptico.

Un modelo simple de la función F sería:

$$F = X_1W_1 + X_2W_2 + \dots + X_iW_i \quad (6.1)$$

Si el resultado de la función F es mayor que el valor umbral (U), la neurona se activa y emite una señal (1) hacia las neuronas de la capa siguiente. Pero, si por el contrario, el resultado es menor que el valor umbral, la neurona permanece inactiva (0) y no envía ninguna señal:

$$X_1W_1 + X_2W_2 + \dots + X_iW_i \leq U \Leftrightarrow \text{Inactivación} \Leftrightarrow Y = 0 \quad (6.2)$$

$$X_1W_1 + X_2W_2 + \dots + X_iW_i > U \Leftrightarrow \text{Activación} \Leftrightarrow Y = 1 \quad (6.3)$$

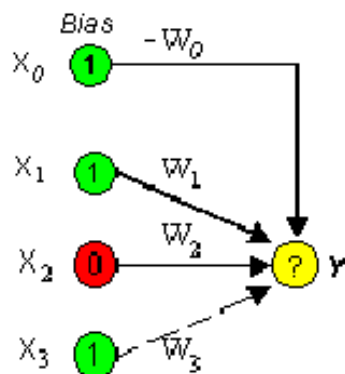
De esta forma, definido un conjunto inicial de pesos en las conexiones, al presentar un "estímulo" (conjunto de ceros y unos que representa un dato,

perfil u objeto) a la capa de entradas, cada neurona en cada capa realiza la operación descrita anteriormente, activándose o no, de manera que al final del proceso las neuronas de la capa de salidas generan un resultado (otro conjunto de ceros y unos), que puede coincidir o no con el que se desea asociar el estímulo.

En el entrenamiento de una red neuronal tanto el peso sináptico de las conexiones como el valor umbral para cada neurona se modifican (según un algoritmo de aprendizaje), con el fin de que los resultados generados por la red coincidan con (o se aproximen a) los resultados esperados.

Y para simplificar el sistema de entrenamiento, el valor umbral (U) pasa a expresarse como un peso sináptico más ($-W_0$), pero asociado a una neurona siempre activa (X_0). Esta neurona siempre activa, se denomina "bias", y se sitúa en la capa anterior a la neurona Y , tal como se muestra en la figura 18.

Figura 18. Neurona, bias y su peso sináptico.



Fuente: Autor

Así, la condición de activación puede re-escribirse como:

$$\underline{X_0W_0} + X_1W_1 + X_2W_2 + \dots + X_iW_i > 0 \Leftrightarrow \text{Activación} \Leftrightarrow Y = 1 \quad (6.4)$$

De esta manera el algoritmo de aprendizaje puede ajustar el umbral como si ajustará un peso sináptico más.

6.2.2. Tipos de aprendizaje básicos. Para poder aprender, las redes neuronales se sirven de un algoritmo de aprendizaje. Estos algoritmos están formados por un conjunto de reglas que permiten a la red neuronal aprender (a partir de los datos que se le suministran), mediante la modificación de los pesos sinápticos de las conexiones entre las neuronas (recordar que el umbral de cada neurona se modificará como si fuera un peso sináptico más).

Generalmente los datos que se usan para entrenar la red se le suministran de manera aleatoria y secuencial.

Los tipos de aprendizaje pueden dividirse básicamente en tres, atendiendo a como se guíe este aprendizaje:

- **Aprendizaje supervisado:** se introducen unos valores de entrada a la red, y los valores de salida generados por esta se comparan con los valores de salida correctos. Si hay diferencias, se ajusta la red en consecuencia.
- **Aprendizaje de refuerzo:** se introducen valores de entrada, y lo único que se le indica a la red si las salidas que ha generado son correctas o incorrectas.
- **Aprendizaje no supervisado:** no existe ningún tipo de guía. De esta manera lo único que puede hacer la red es reconocer patrones en los datos de entrada y crear categorías a partir de estos patrones. Así

cuando se le entre algún dato, después del entrenamiento, la red será capaz de clasificarlo e indicará en que categoría lo ha clasificado.

6.3. MODELO SIMPLE DE RED NEURONAL

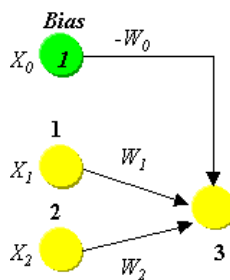
Se considerará una red neuronal formada por 2 capas:

- Una capa de entradas formada por 2 neuronas: **1** y **2**.
- Una capa de salidas formada por una sola neurona: **3**.

Las conexiones entre las dos neuronas de entrada y la neurona de salida presentan pesos sinápticos ajustables mediante el entrenamiento. Y a su vez, el valor umbral (U) de la neurona 3 puede ser ajustado como un peso sináptico más, al considerar $U = -W_0$ asociado a una neurona bias (siempre activa: $X_0 = 1$).

El esquema de este modelo sería el presentado en la figura 19.

Figura 19. Modelo Simple de tres neuronas y una neurona bias.



Fuente: FREEMAN J. A., SKAPURA D. M. "Redes neuronales algoritmos, aplicaciones y técnicas de programación".³⁴

Y la condición de activación de la neurona 3 sería:

$$X_0W_0 + X_1W_1 + X_2W_2 > 0 \Leftrightarrow \text{Activación} \quad (6.5)$$

Además, en este modelo, cada neurona de una capa "transmite" su estado de activación (0 o 1) a la siguiente capa de neuronas, y después deja de estar activa.

Ahora, entrenaremos la red de manera que la neurona 3 se active sólo si las neuronas 1 y 2 están activas, y permanezca inactiva en caso contrario. Hay así cuatro posibles casos:

Tabla 11. Casos de entrenamiento en un sistema de tres neuronas.

	X_1	X_2	Capa de salida (X_3)
Caso 1	0	0	0
Caso 2	0	1	0
Caso 3	1	0	0
Caso 4	1	1	1

Fuente: FREEMAN J. A., SKAPURA D. M. "Redes neuronales algoritmos, aplicaciones y técnicas de programación".³⁴

Para entrenar esta red, usaremos una regla de aprendizaje derivada del método de **aprendizaje Hebbiano**.

6.3.1. Regla de aprendizaje de Hebb. También llamada aprendizaje de coincidencia o aprendizaje Hebbiano. Y fue sugerida por Donald Hebb en 1949, a partir de estudios con neuronas reales.

Hebb observó que la sinapsis (conexión entre 2 neuronas reales) se reforzaba sí la neurona de entrada (o presináptica) y la neurona de salida (o

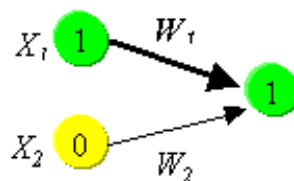
postsináptica) eran activadas de manera continua. De esta forma, las conexiones que se usan son las que se refuerzan.

Así, según este método de aprendizaje aplicado a las redes neuronales artificiales, las conexiones entre las neuronas de entrada activas y las neuronas de salida activas se refuerzan durante el entrenamiento: coincidencias entre actividad de entrada y actividad de salida se intensifican. Mientras que las conexiones entre neuronas de entrada inactivas y neuronas de salida (activas o inactivas) no se refuerzan.

Este método de aprendizaje puede ser tanto supervisado como no supervisado. Cuando es supervisado, la respuesta correcta para el dato de entrada es introducida para cada neurona de salida, y los pesos sinápticos entre las neuronas activas se incrementan, mientras que los pesos entre neuronas que no estén activas simultáneamente permanecen igual que estaban.

El problema de este método es que no tiene en cuenta la eficacia de la red. Así, aunque la red ya esta entrenada y los valores de entrada generen valores de salida correctos, la regla de aprendizaje continua incrementando los pesos sinápticos entre neuronas activas.

Figura 20. Intensificación del peso sináptico entre dos neuronas activas.



Fuente: FREEMAN J. A., SKAPURA D. M. "Redes neuronales algoritmos, aplicaciones y técnicas de programación".³⁴

6.3.2. Entrenamiento de la red. Debido al problema de la regla de aprendizaje de Hebb, usaremos una regla de aprendizaje derivada, y que si tenga en cuenta la eficacia de la red en cada momento:

1. Si la salida generada por la neurona 3 (para los valores de las otras dos neuronas) es la correcta, **no** se realizan ajustes de los pesos sinápticos.
2. Si la salida es **1** pero debería ser **0**, se **reducen** sólo los pesos de las conexiones activas según una constante **C**.
3. Si la salida es **0** pero debería ser **1**, entonces se **aumentan** sólo los pesos de las conexiones activas según la misma constante **C**.

Así, en cada ciclo de entrenamiento:

- Se presenta un dato de entrada (formado por los valores de las neuronas 1 y 2) del conjunto de datos de entrenamiento.
- La red, a partir del dato de entrada generará un dato de salida.
- Se aplica la regla anterior, que mide la eficiencia de la red y actúa en consecuencia.

Y se realizan diferentes ciclos de entrenamiento con los valores de entrenamiento hasta que la red responda correctamente a todos los datos de entrada (en todos los casos de entrenamiento) o hasta que la red ya no pueda aprender más.

De esta forma, si partimos de unos pesos sinápticos iguales a 0 ($W_1 = W_2 = W_0 = 0$) y de una constante de refuerzo igual a 1 ($C = 1$):

- Introducimos el dato de entrada del **caso 1**:
 $(X_1, X_2) = (0, 0)$
- Puesto que todos los pesos valen 0:
 $X_0W_0 + X_1W_1 + X_2W_2 = 0 \Leftrightarrow \text{Inactivación}$
- Así el valor de salida de la neurona 3 es 0, y como el valor de la neurona 3 tenía que ser 0 (según el caso 1 de entrenamiento), no se realizan ajustes de los pesos.
- Introducimos el dato de entrada del **caso 4**:
 $(X_1, X_2) = (1, 1)$
- Puesto que todos los pesos valen 0:
 $X_0W_0 + X_1W_1 + X_2W_2 = 0 \Leftrightarrow \text{Inactivación}$
- Pero, como el valor de salida de la neurona 3 es 0, y tenía que ser 1 (según el caso 4 de entrenamiento), los pesos de las conexiones activas se aumentan en un valor de C (tal como indica el apartado 3 de la regla de aprendizaje):
 - $W_1 = W_1 + C = 0 + 1 = 1$
 - $W_2 = W_2 + C = 0 + 1 = 1$
 - $W_0 = W_0 + C = 0 + 1 = 1$

Y se continuaría de esta manera hasta terminar el entrenamiento, en cuyo caso obtendríamos los siguientes pesos sinápticos ajustados:

Tabla 12. Casos de entrenamiento en un sistema de tres neuronas.

Peso sináptico	Valor
W_1	1
W_2	2
W_0	-2

Fuente: FREEMAN J. A., SKAPURA D. M. "Redes neuronales algoritmos, aplicaciones y técnicas de programación".³⁴

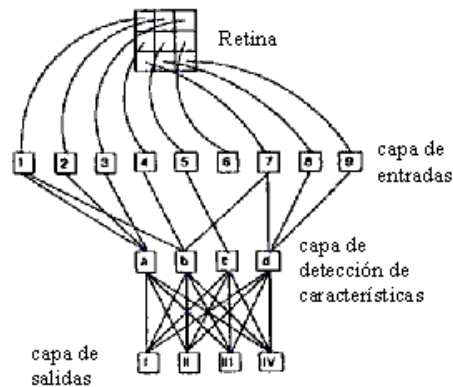
6.3.3. Modelo del Perceptrón. Este modelo de red simple, sería un modelo reducido del modelo del Perceptrón, desarrollado por Rosenblatt entre 1958 y 1962, a partir de los modelos de red de McCulloch-Pitts.

En el ojo, la imagen proyectada sobre la retina es convertida en señales eléctricas gracias a las células sensitivas a la luz (conos y bastones). Estas señales, a través de los axones de las células ganglionares, son transmitidas al cerebro, que interpreta la imagen.

El modelo del Perceptrón está basado en esta idea del ojo: las células sensibles de la retina serían detectores conectados con neuronas de la capa de entrada; estas neuronas se activan (1) cuando se activa el detector correspondiente. Cada neurona de la capa de entrada está conectada con diferentes neuronas de la capa de detección de características. Y cada neurona de esta capa estaría conectada a su vez con diferentes neuronas de la capa de salidas del Perceptrón (figura 21).

Y la regla de aprendizaje usada para entrenar la red es una versión simplificada de la regla de aprendizaje del Perceptrón.

Figura 21. Modelo del Perceptrón.



Fuente: FREEMAN J. A., SKAPURA D. M. "Redes neuronales algoritmos, aplicaciones y técnicas de programación".³⁴

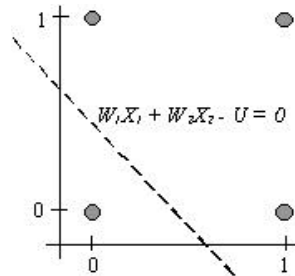
6.3.4. Limitaciones del modelo. El modelo simple de red neuronal (al igual que el modelo del Perceptrón en el que se basa), presenta bastantes limitaciones en las tareas que pueden llegar a aprender.

Así, por ejemplo, el modelo sería incapaz de aprender a realizar la operación XOR (O exclusivo): No puede ser entrenado con éxito para que la neurona 3 se active si las neuronas 1 o 2 están activas, pero no las dos a la vez. Esto es debido a que el método de entrenamiento lo que hace es mover la siguiente recta sobre un plano:

$$X_1W_1 + X_2W_2 - U = 0 \quad (6.6)$$

De manera que los puntos (X_1, X_2) del plano por encima de la recta hacen que la neurona 3 se active, y los que quedan por debajo de la recta hacen que la neurona 3 permanezca inactiva. Y no se puede dividir el plano en 2 partes de manera que solo los puntos $(0,1)$ y $(1,0)$ queden en la zona de activación de la neurona 3.

Figura 22. Plano de activación.



Fuente: Autor

Para realizar este tipo de tareas "complejas" es precisa una red neuronal que contenga más de una capa de pesos sinápticos ajustables. Para ello, se añaden capas ocultas de neuronas entre la capa de entradas y la capa de salidas.

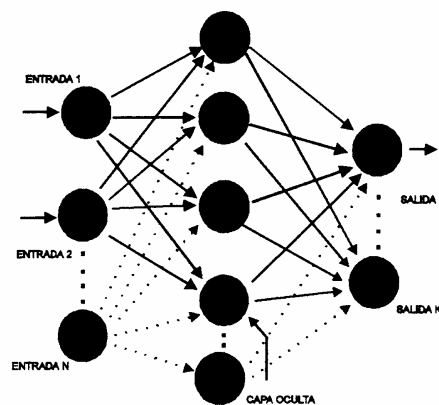
6.4. PERCEPTRÓN MULTICAPA³⁵

Es una estructura paralela, que procesa información en forma distribuida y consistente de elementos de procesamiento - que pueden tener una memoria local y pueden realizar procesamiento de información localizado - interconectados con canales unidireccionales. También se puede definir como un clasificador que determina cual de las M clases es la más representativa de un patrón estático de entrada con N característica.

En la figura 23, se presenta una estructura general de una **red neuronal artificial (RNA) multicapa**, que incluye tres capas con diferente número de nodos comunicados entre sí. En la parte izquierda se encuentra la capa de entrada, que es la encargada de percibir los estímulos exteriores, normalmente cada nodo en la capa de entrada se encarga de capturar y

procesar un estímulo o característica crítica en el proceso de aprendizaje; en la capa de salida, al lado derecho de la figura 23, se presentan los resultados o las respuestas a los estímulos sensoriales, en la parte intermedia se encuentra una o varias capas llamadas ocultas, con un numero determinado de nodos adecuado para cada problema específico.

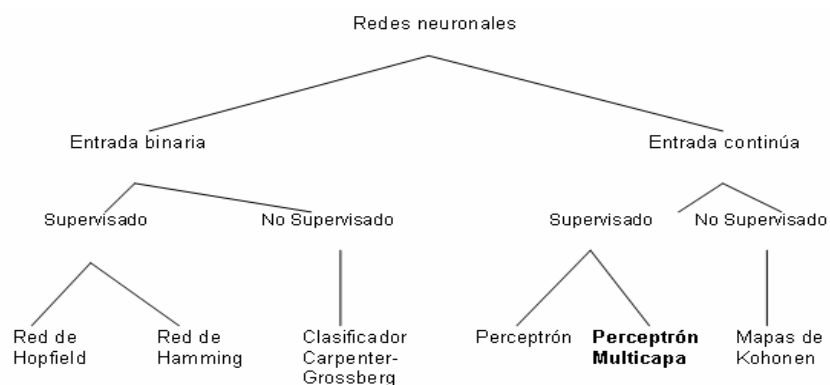
Figura 23. Arquitectura de una red neuronal multicapa.



Fuente: WERBOS P. J. "Backpropagation past and future".³⁵

Las redes neuronales como clasificadores se pueden agrupar de acuerdo a la figura 24.

Figura 24. Tipos de redes neuronales como clasificadores y según su tipo de aprendizaje.



Fuente: FREEMAN J. A., SKAPURA D. M. "Redes neuronales algoritmos, aplicaciones y técnicas de programación".³⁴

6.4.1. Formulación vectorial. En el modelo de red que se va a tratar resulta útil describir ciertas magnitudes en forma de vectores. Considérese una red neuronal compuesta por varias **capas** de elementos de procesamiento idénticos. Si una cierta capa contiene n unidades, las salidas de esa capa se pueden considerar como un vector n -dimensional, $\mathbf{x} = (x_1, x_2, \dots, x_n)^t$, en donde el índice t denota traspuesto. En nuestra notación los vectores escritos en negrilla, como \mathbf{x} , se supondrá que son vectores columna. Cuando se escriban en forma de filas, se añadirá el símbolo de transposición. En otras palabras, la notación \mathbf{x}^t denota un vector fila.

Suponiendo que el vector n -dimensional de salida del párrafo anterior proporciona los valores de entrada de todas las unidades de una capa m -dimensional (una capa de m unidades). Cada una de las unidades de la capa m -dimensional poseerá n pesos asociados a las conexiones procedentes de la capa anterior. De esta manera, hay m vectores de pesos n -dimensionales asociados a esta capa; hay un vector de pesos n -dimensional para cada una de las m unidades. El vector de pesos de la i -ésima unidad se puede escribir en la forma $\mathbf{w}_i = (w_{i1}, w_{i2}, \dots, w_{in})^t$. Se puede añadir un índice a la notación de los pesos para distinguir entre los pesos de diferentes capas.

La entrada neta de la i -ésima unidad se puede escribir en términos del producto interno, o producto escalar, del vector de entradas por el vector de pesos. Para vectores *de iguales dimensiones*, el producto interior se define como la suma de los productos de los componentes correspondientes de ambos vectores.

$$Net_{ai} = \sum_{j=1}^n x_j w_{ij} \quad (6.7)$$

En donde n es el número de conexiones que haya en la i -ésima unidad. Esta ecuación se puede escribir de forma resumida en notación vectorial en la forma:

$$Net_i = \mathbf{x} \mathbf{w}_i \quad (6.8)$$

ó bien

$$Net_i = \mathbf{x}^t \mathbf{w}_i \quad (6.9)$$

Téngase en cuenta que, como consecuencia de las reglas de multiplicación de vectores:

$$\mathbf{x}^t \mathbf{w}_i = \mathbf{w}_i \mathbf{x}^t \quad (6.10)$$

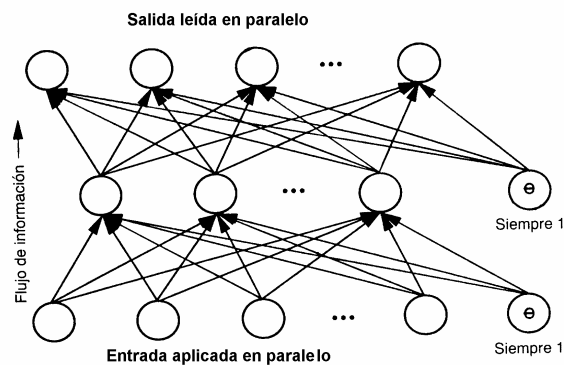
6.5. RETROPROPAGACIÓN^{35, 36}

Son muchas las posibles aplicaciones de los computadores que resultan difíciles de realizar porque hay muchos problemas cuya resolución no es adecuada mediante procesos secuenciales. Las aplicaciones que deben realizar complejas traducciones de datos, y sin embargo no poseen una función de correspondencia predefinida que describa el proceso de traducción, o aquellas que deben proporcionar una «mejor aproximación» como salida cuando se les presentan unos datos de entrada con ruido son sólo dos ejemplos de problemas de este tipo.

Una RNA que ha resultado útil para resolver problemas que requieren el reconocimiento de tramas complejas y la realización de funciones de correspondencia no triviales es la red de retropropagación (BPN), que fue descrita formalmente en primer lugar por Werbos y posteriormente por Parker y por Rumelhart y McClelland. Esta red, que se ilustra de forma

general en la Figura 25, está diseñada para que funcione como red multicapa, con propagación hacia adelante, empleando el modo supervisado de aprendizaje.

Figura 25. Arquitectura general de una red de retropropagación.



Fuente: WERBOSP. J. "Backpropagation past and future".³⁵

6.5.1. Funcionamiento de una BPN. Corresponde hacer aquí una descripción de funcionamiento de la red, para ilustrar la forma en que se puede utilizar una BPN para resolver problemas complejos de reconocimiento de tramas. Para empezar, la red aprende un conjunto predefinido de pares de entradas y salidas dados como ejemplo, empleando un ciclo propagación-adaptación de dos fases. Una vez que se ha aplicado una trama de entrada como estímulo para la primera capa de unidades de la red, ésta se va propagando a través de todas las capas superiores hasta generar una salida. La señal de salida se compara entonces con la salida deseada, y se calcula una señal de error para cada unidad de salida.

Las señales de error se transmiten entonces hacia atrás, partiendo de la capa de salida, hacia todos los nodos de la capa intermedia que contribuyan

directamente a la salida. Sin embargo, las unidades de la capa intermedia sólo reciben una fracción de la señal total de error, basándose aproximadamente en la contribución relativa que haya aportado la unidad a la salida original. Este proceso se repite, capa por capa, hasta que todos los nodos de la red hayan recibido una señal de error que describa su contribución relativa al error total. Basándose en la señal de error percibida, se actualizan los pesos de conexión de cada unidad, para hacer que la red converja hacia un estado que permita codificar todas las tramas de entrenamiento.

La importancia de este proceso consiste en que, a medida que se entrena la red, los nodos de las capas intermedias se organizan a si mismos de tal modo que los distintos nodos aprenden a reconocer distintas características del espacio total de entradas. Después del entrenamiento, cuando se les presente una trama arbitraria de entrada que contenga ruido o que esté incompleta, las unidades de las capas ocultas de la red responderán con una salida activa si la nueva entrada contiene una trama que se asemeje a aquella característica que las unidades individuales hayan aprendido a reconocer durante su entrenamiento. Y a la inversa, las unidades de las capas ocultas tienen una tendencia a inhibir sus salidas si la trama de entrada no contiene la característica para reconocer.

A medida que las señales se propagan a través de las diferentes capas de la red, la trama de actividad que esta presente en todas las capas superiores se puede ver como una trama con características que son reconocidas por las unidades de la capa subsiguiente. La trama de salida que se genera se puede ver como un mapa de características que ofrece una indicación de la presencia o ausencia de muchas combinaciones distintas de características dentro de la entrada. El resultado global de este comportamiento es que la BPN constituye un medio eficiente para permitir a un sistema de

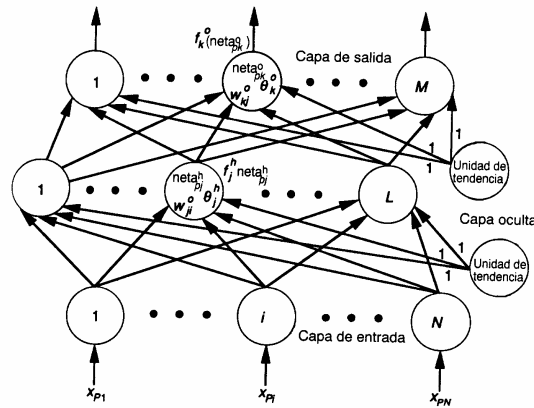
computadores examinar tramas de datos que pueden ser ruidosas, o estar incompletas, y reconocer tramas sutiles a partir de entradas parciales.

Hay varios investigadores que han demostrado, que durante el entrenamiento, las BPN tienden a desarrollar relaciones internas entre nodos con el fin de organizar los datos de entrenamiento en clases de tramas. Esta tendencia se puede extrapolar, para llegar a la hipótesis consistente en que todas las unidades de capas ocultas de una BPN son asociadas de alguna manera a características específicas de la trama de entrada como consecuencia del entrenamiento. Lo que sea o no sea exactamente la asociación puede no resultar evidente para el observador humano. Lo importante es que la red ha encontrado una representación interna que le permite generar las salidas deseadas cuando se le dan las entradas de entrenamiento. Esta misma representación interna se puede aplicar a entradas que no fueran utilizadas durante el entrenamiento. La BPN clasificará estas entradas que no había visto hasta el momento según las características que compartan con los ejemplos de entrenamiento.

6.5.2. La regla delta generalizada. En esta sección se presenta la descripción matemática formal del funcionamiento de una BPN. Presentaremos una derivación detallada **de La Regla Delta Generalizada** (RDG), que es el algoritmo de aprendizaje de la red.

La Figura 26 sirve como referencia para la mayoría de las descripciones. La BPN es una red formada por capas, con propagación hacia adelante, que está completamente interconectada entre capas. Por tanto, no hay conexiones de realimentación ni conexiones que salten una capa para ir directamente a una capa anterior. Aunque sólo se utilizan tres capas en nuestro tratamiento, se puede admitir más de una capa oculta.

Figura 26. Arquitectura de una BPN de tres capas.



Fuente: WERBOS P.J. "Backpropagation and neurocontrol: A review and prospectus".³⁶

Se dice que una red neuronal es una red de correspondencias si se demuestra capaz de calcular alguna relación funcional entre su entrada y su salida. Por ejemplo, si la entrada de la red es el valor de un ángulo θ y la salida es el coseno del ángulo, la red establece la correspondencia $\theta \rightarrow \cos(\theta)$. Para una función tan sencilla, no se necesita una red neuronal; sin embargo, podría ser necesario llevar a cabo una correspondencia complicada en la que no se conociera la forma de describir la relación funcional por anticipado, pero fueran conocidos ejemplos de correspondencias correctas. En esta situación, la potencia de una red neuronal para descubrir su propio algoritmo es de enorme utilidad.

Suponiendo que se tiene un conjunto de P pares de vectores (x_1, y_1) , (x_2, y_2) , ..., (x_p, y_p) , que son ejemplos de una correspondencia funcional $y = \phi(x) : x \in \mathbb{R}^N, y \in \mathbb{R}^M$. Se desea entrenar a la red para que aprenda una aproximación $\mathbf{o} = \mathbf{y}' = \mathbf{f}'(\mathbf{x})$. A continuación se deriva un método para hacer este entrenamiento que suele funcionar, suponiendo que los pares de vectores de

entrenamiento se hayan seleccionado adecuadamente y que haya un número suficiente de ellos. Debe recordarse que el aprendizaje, en una red neuronal, significa hallar un conjunto adecuado de pesos. La técnica de aprendizaje que se describe aquí se asemeja al problema de hallar la ecuación de una línea que sea la que mejor se ajuste a un cierto número de puntos conocidos. Además, es una generalización de la regla de los mínimos cuadrados. Para un problema de ajuste de líneas, es probable que utilizásemos una aproximación de mínimos cuadrados. Dado que la relación que pretendemos estudiar será, probablemente, no lineal, además de multidimensional, empleamos una versión iterativa del sencillo método de mínimos cuadrados, denominada *técnica del descenso más pronunciado*.

Para empezar, revisemos las ecuaciones para el procesamiento de información que hay en la red de tres capas de la Figura 25. Se aplica un vector de entrada, $\mathbf{x}_p = (x_{p1}, x_{p2}, \dots, x_{pN})^t$, en la capa de entrada de la red. Las unidades de entrada distribuyen los valores a las unidades de la capa oculta. La entrada neta de la j -ésima unidad oculta es:

$$neta_{pj}^h = \sum_{i=1}^n w_{ji}^h x_{pi} + q_j^h \quad (6.11)$$

En donde w_{ji}^h es el peso de la conexión procedente de la i -ésima unidad de entrada, y q_j^h es el término de tendencia o valor umbral el cual tiene implícito su peso. El índice «h» se refiere a magnitudes de la capa oculta. Se supone que la activación de este nodo es igual a la entrada neta; entonces, la salida de este nodo es

$$i_{pj} = f_j^h(neta_{pj}^h) \quad (6.12)$$

Las ecuaciones para los nodos de salida son las que siguen:

$$neta_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj} + q_k^o \quad (6.13)$$

$$o_{pk} = f_k^o(neta_{pk}^o) \quad (6.14)$$

En donde el índice “o” se refiere a magnitudes de la capa de salida.

El conjunto inicial de valores de pesos representa una primera aproximación de los pesos correctos para el problema. A diferencia de otros métodos, la técnica que empleamos aquí no depende de hacer una *buena* primera aproximación.

El procedimiento básico para entrenar la red está plasmado en la siguiente descripción:

1. Se aplica un vector de entrada a la red, y se calculan los correspondientes valores de salida.
2. Se comparan las salidas obtenidas con las salidas correctas, y se determina una medida del error.
3. Se determina en que dirección (+ o -) debe cambiar cada peso con objeto de reducir el error.
4. Se determina la cantidad en que es preciso cambiar cada peso.
5. Se aplican las conexiones a los pesos.
6. Se repiten los pasos del 1 al 5 con todos los vectores de entrenamiento hasta que el error para todos los vectores del conjunto de entrenamiento quede reducido a un valor aceptable.

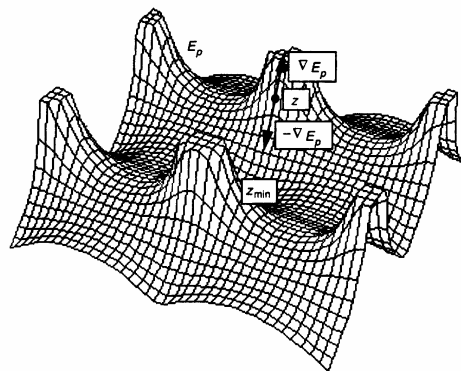
6.5.2.1. Actualización de pesos de la capa de salida. Definiremos el error de una sola unidad de salida en la forma $\delta_{pk} = (Y_{pk} - \alpha_k)$, en donde el

subíndice «P» se refiere al p-ésimo vector de entrenamiento, y «K» se refiere a la k-ésima unidad de salida. En este caso, Y_{pk} es el valor de salida deseado, y o_{pk} es la salida obtenida a partir de la k-ésima unidad. El error que se minimiza por RDG es la suma de los cuadrados de los errores de todas las unidades de salida:

$$E_p = \frac{1}{2} \sum_{k=1}^M d_{pk}^2 \quad (6.15)$$

El factor $\frac{1}{2}$ de la ecuación 6.15 aparece por conveniencia para calcular derivadas más adelante. Dado que aparecerá una constante arbitraria en el resultado final, la presencia de este factor no invalida la derivación. Para determinar el sentido en que se deben cambiar los pesos, se calcula el valor negativo del gradiente de E_p , $-\nabla E_p$, respecto a los pesos W_{kj} . Después se pueden ajustar los valores de los pesos de tal forma que se reduzca el error total. Suele resultar útil pensar que E_p es una superficie en el espacio de pesos.

Figura 27. Superficie hipotética del espacio de pesos.



Fuente: Autor

La Figura 27 muestra un ejemplo sencillo en el que la red tiene varios pesos. El gradiente, ∇E_p en el punto \mathbf{z} , se muestra junto con el valor negativo del gradiente. Los cambios de pesos deberían producirse en la dirección negativa del gradiente, que es la dirección del descenso más pronunciado a lo largo de la superficie en el punto \mathbf{z} . Además, los cambios de los pesos deberían hacerse iterativamente hasta que E_p alcance el punto \mathbf{z}_{\min} .

Para no complicar las cosas, consideramos por separado cada componente de ∇E_p . Partiendo de la Ecuación 6.15 y de la definición de \mathbf{d}^{pk}

$$E_p = \frac{1}{2} \sum_k (y_{pk} - o_{pk})^2 \quad (6.16)$$

Y

$$\frac{\partial (E_p)}{\partial (w_{kj}^o)} = -(y_{pk} - o_{pk}) \frac{\partial f_k^o}{\partial (neta_{pk}^o)} \frac{\partial (neta_{pk}^o)}{\partial (w_{kj}^o)} \quad (6.17)$$

En donde se ha utilizado la Ecuación 6.14 como valor de salida, o_{pk} , y la regla de la cadena para las derivadas parciales. Por el momento, no intentaremos evaluar las derivadas de f_k^o , sino que nos limitaremos a escribirlas en la forma $f_k^{o'}(neta_k^o)$. El último factor de la Ecuación 6.17 es:

$$\frac{\partial (neta_{pk}^o)}{\partial (w_{kj}^o)} = \left(\frac{\partial}{\partial (w_{kj}^o)} \sum_{j=1}^L w_{kj}^o I_{pj} + q_k^o \right) = i_{pj} \quad (6.18)$$

Combinando las ecuaciones 6.17 y 6.18, tenemos lo siguiente para el gradiente negativo:

$$-\frac{\partial E_p}{\partial (w_{kj}^o)} = (y_{pk} - o_{pk}) f_k^{o'}(neta_{pk}^o) i_{pj} \quad (6.19)$$

En la magnitud del cambio de peso, consideramos que será proporcional al gradiente negativo. De esta manera, los pesos de la capa de salida se actualizan según lo siguiente:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \Delta_p w_{kj}^o(t) \quad (6.20)$$

En donde

$$\Delta_p w_{kj}^o = h(y_{pk} - o_{pk}) f_k^{o'}(neta_{pk}^o) i_{pj} \quad (6.21)$$

El factor η se denomina **Rata de Aprendizaje**. Es un valor positivo y que suele ser menor que 1.

Volvamos a examinar la función $f_k^{o'}$. En primer lugar, obsérvese el requisito consistente en que $f_k^{o'}$ sea derivable. Este requisito elimina la posibilidad de utilizar una función discontinua, ya que esta no es derivable en la(s) discontinuidad(es). Aquí hay dos formas de la función de salida que tienen interés:

$$f_k^o(neta_{jk}^o) = neta_{jk}^o$$

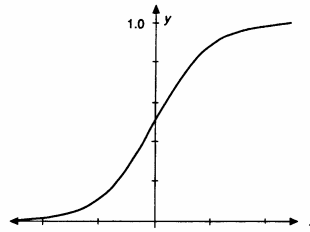
Y

$$f_k^o(neta_{jk}^o) = (1 + e^{-neta_{jk}^o})^{-1}$$

La primera función define la unidad lineal de salida. La segunda función es lo que se denomina una sigmoide, o función logística; se ha representado en la Figura 28. La selección de la función de salida depende de la forma en que se decida representar los datos de salida. Por ejemplo, si se desea que las unidades de salida sean binarias, se utiliza una función de salida sigmoide, puesto que la sigmoide limita la salida y es casi biestable, pero

también es derivable. En otros casos, es tan aplicable una función de salida lineal como una sigmoide.

Figura 28. Función sigmoideal típica de forma exponencial.



Fuente: Autor

En el primer caso, $f_k^{o'} = 1$; en el segundo, $f_k^{o'} = f_k^o(1 - f_k^o) = o_{pk}(1 - o_{pk})$

Para ambos casos, tenemos:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \mathbf{h}(y_{pk} - o_{pk})i_{pj} \quad (6.22)$$

Para la salida lineal, y

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \mathbf{h}(y_{pk} - o_{pk})o_{pk}(1 - o_{pk})i_{pj} \quad (6.23)$$

Para la salida en forma de sigmoide.

Resumiendo las ecuaciones de actualización de pesos definiendo una magnitud:

$$\mathbf{d}_{pk}^o = (y_{pk} - o_{pk})f_k^{o'}(neta_{pk}^o) \quad (6.24)$$

$$\mathbf{d}_{pk} f_k^{o'}(neta_{pk}^o)$$

Entonces se puede escribir la ecuación de actualización de pesos en la forma:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \eta d_{pk}^o i_{pj} \quad (6.25)$$

Independientemente de la forma funcional de la función de salida. El error que minimiza este proceso es:

$$E = \sum_{p=1}^P E_p \quad (6.26)$$

En donde P es el número de tramas del conjunto de entrenamiento. En la práctica, resulta poco ventajoso este seguimiento estricto de la analogía con el método de los mínimos cuadrados. Además, para utilizar este método es preciso almacenar una gran cantidad de información. Recomendamos llevar a cabo las actualizaciones de los pesos a medida que se va procesando cada trama de entrenamiento.

6.5.2.2. Actualizaciones de los pesos de capas ocultas. Es deseable repetir para la capa oculta el mismo tipo de cálculo que se ha realizado para la capa de salida. Surge un problema cuando se intenta determinar una medida del error de las salidas para las unidades de la capa oculta. Se sabe cuál es la salida obtenida, pero no tenemos forma de saber por anticipado cuál debería ser la salida correcta para estas unidades. Intuitivamente, el error E_p debe estar relacionado de alguna manera con los valores de salida de la capa oculta. Podemos comprobar la veracidad de nuestra intuición volviendo a la Ecuación 6.16:

$$E_p = \frac{1}{2} \sum_k (y_{pk} - o_{pk})^2 = \frac{1}{2} \sum_k (y_{pk} - f_k^o(\text{net}_{pk}^o))^2 = \frac{1}{2} \sum_k (y_{pk} - f_k^o(\sum_j w_{kj}^o i_{pj} + q_k^o))^2$$

Se sabe que i_{pj} depende de los pesos de las capas ocultas a través de las Ecuaciones 6.11 y 6.12. Podemos aprovechar este hecho para calcular el gradiente de E_p respecto a los pesos de las capas ocultas.

$$\frac{\partial E_p}{\partial w_{ji}^h} = \frac{1}{2} \sum_k \frac{\partial}{\partial w_{ji}^h} (y_{pk} - o_{pk})^2 = - \sum_k (y_{pk} - o_{pk}) \frac{\partial(o_{pk})}{\partial(neta_{pk}^o)} \frac{\partial(neta_{pk}^o)}{\partial(i_{pj})} \frac{\partial(i_{pj})}{\partial(neta_{pj}^h)} \frac{\partial(neta_{pj}^h)}{\partial(w_{ji}^h)} \quad (6.27)$$

Cada uno de los factores de la Ecuación 6.27 puede calcularse explícitamente a partir de ecuaciones anteriores. El resultado es el que sigue:

$$\frac{\partial(E_p)}{\partial(w_{ij}^h)} = - \sum_k (y_{pk} - o_{pk}) f_k^{o'}(neta_{pk}^o) w_{kj}^o f_j^{h'}(neta_{pj}^h) x_{pi} \quad (6.28)$$

Se actualizan los pesos de la capa oculta proporcionalmente al valor negativo de la Ecuación 6.28:

$$\Delta_p w_{ji}^h = \eta f_j^{h'}(neta_{pj}^h) x_{pi} \sum_k (y_{pk} - o_{pk}) f_k^{o'}(neta_{pk}^o) w_{kj}^o \quad (6.29)$$

En donde η es, una vez más, la tasa de aprendizaje. Se puede utilizar la definición de d_{pk}^o dada en la sección anterior para escribir:

$$\Delta_p w_{ji}^h = \eta f_j^{h'}(neta_{pj}^h) x_{pi} \sum_k d_{pk}^o w_{kj}^o \quad (6.30)$$

Obsérvese que todas las actualizaciones de pesos de la capa oculta dependen de *todos* los términos de error, d_{pk}^o , de la capa de salida. Este resultado es el lugar en que surge la noción de *propagación hacia atrás*. Los errores conocidos de la capa de salida *se propagan hacia atrás*, hacia la capa oculta, para determinar los cambios de peso adecuados en esa capa. Si se define un término de error para la capa oculta:

$$\mathbf{d}_{pj}^h = f_j^h(\mathbf{neta}_{pj}^h) x_{pi} \sum_k \mathbf{d}_{pk}^o w_{kj}^o \quad (6.31)$$

Se da lugar a que las ecuaciones de actualización de pesos pasen a ser análogas a las correspondientes a la capa de salida:

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \mathbf{h} \mathbf{d}_{pj}^h x_i \quad (6.32)$$

Por último, para cerrar los límites puestos a la RDG, obsérvese que tanto la Ecuación 6.25 de la sección anterior como la 6.32 de esta sección tienen la misma forma.

6.5.2.3. Resumen de BPN. Con objeto de reducir la necesidad de pasar páginas con objeto de hallar las ecuaciones correspondientes, se recogen aquí todas las ecuaciones relevantes para el BPN. Se presentan por el orden en que serian utilizadas durante el entrenamiento para un único par de vectores de entrenamiento.

1. Se aplica el vector de entrada $\mathbf{x}_p = (x_{p1}, x_{p2}, \dots, x_{pn})^t$ a las unidades de entrada.
2. Se calculan los valores netos procedentes de las entradas para las unidades de la capa oculta:

$$\mathbf{neta}_{pj}^h = \sum_{i=1}^n w_{ji}^h x_{pi} + \mathbf{q}_j^h$$

3. Se calculan las salidas de la capa oculta:

$$i_{pj} = f_j^h(\mathbf{neta}_{pj}^h)$$

4. Se pasa a la capa de salida. Se calculan los valores netos de las entradas para cada unidad

$$neta_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj} + q_k^o$$

5. Se calculan las salidas:

$$o_{pk} = f_k^o(neta_{pk}^o)$$

6. Se calculan los términos de error para las unidades de salida:

$$d_{pk}^o = (y_{pk} - o_{pk}) f_k^{o'}(neta_{pk}^o)$$

7. Se calculan los términos de error para las unidades ocultas:

$$d_{pj}^h = f_j^{h'}(neta_{pj}^h) x_{pi} \sum_k d_{pk}^o w_{kj}^o$$

Se observa que los términos de error de las unidades ocultas se calculan antes que hayan sido actualizados los pesos de conexión con las unidades de la capa de salida.

8. Se actualizan los pesos de la capa de salida:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \eta d_{pk}^o i_{pj}$$

9. Se actualizan los pesos de la capa oculta:

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \eta d_{pj}^h x_i$$

El orden de actualización de pesos de una capa individual no es importante. Se debe asegurar el cálculo del término de error:

$$E_p = \frac{1}{2} \sum_{k=1}^M d_{pk}^2$$

Debido a que esta magnitud es la medida de lo bien que está aprendiendo la red. Cuando el error resulta aceptablemente pequeño para todos los pares de vectores de entrenamiento, éste se puede dar por concluido.

6.6. CONSIDERACIONES PRÁCTICAS

Hay algunos temas que se han omitido en las secciones anteriores para no apartar nuestra atención de las ideas principales. Los cuales se tratan a continuación.

6.6.1. Datos de entrenamiento. Es muy importante dar una definición de los términos *adecuado* y *suficiente* con respecto a la selección de pares de vectores de entrenamiento para la BPN. Desafortunadamente, no hay una definición sencilla que sea aplicable a todos los casos. Al igual que muchos aspectos de los sistemas de redes neuronales, la experiencia suele ser la mejor maestra. A medida que va uno adquiriendo soltura en la utilización de redes, se adquiere también un conocimiento acerca de la forma de seleccionar y preparar conjuntos de entrenamiento. Por tanto, lo que se va a dar aquí son unas cuantas ideas generales.

En general, se pueden utilizar todos los datos que estén disponibles para entrenar la red, aunque quizá no sea necesario utilizarlos todos. Con cierta frecuencia, lo único que se necesita para entrenar con éxito una red es un pequeño subconjunto de los datos de entrenamiento de los que se dispone. Los datos restantes pueden emplearse para probar la red, con objeto de verificar que la red pueda llevar a cabo la asociación deseada al utilizar vectores de entrada que nunca haya encontrado durante el entrenamiento.

Si se está entrenando una red para que funcione en un entorno con ruido, como por ejemplo el paso de una imagen de píxeles a ASCII, entonces hay que incluir unos cuantos vectores de entrada con ruido en el conjunto de datos. Algunas veces la adición de ruido a los vectores de entrada durante el entrenamiento ayuda a la red a converger, incluso en el caso de que no se espere ruido en las entradas.

La BPN admite bien la generalización. Lo que quiere decir aquí *generalización* es que dados varios vectores de entrada distintos, todos los cuales pertenecen a una misma clase, una BPN aprenderá a adaptarse a las similitudes significativas de los vectores de entrada. Los datos irrelevantes serán ignorados.

En contraste con las generalizaciones, la BPN no extrapolará bien. Si una BPN se entrena de modo inadecuado o insuficiente empleando una clase concreta de vectores de entrada, la posterior identificación de miembros de esa clase puede ser imprecisa. Asegúrese de que los datos de entrenamiento cubran todo el espacio de entradas esperado. Durante el proceso de entrenamiento, seleccione aleatoriamente los pares de vectores de entrenamiento del conjunto, si es que el problema se presta a esta estrategia. En todo caso, no entrene por completo a la red con vectores de una clase, pasando después a otra clase; la red se *olvidará* del entrenamiento original.

Si la función de salida es una sigmoide, entonces será preciso aplicar una escala a los vectores de entrada. Como consecuencia de la forma de la función sigmoidea, las salidas de la red nunca pueden alcanzar el cero ni el uno. Por tanto, hay que utilizar valores como 0,1 y 0,9 para representar los valores de entrada más pequeños y más grandes. También se puede desplazar la sigmoide de tal manera que, por ejemplo, los valores limitantes

pasen a ser $\pm 0,4$. Además, se puede modificar la pendiente de la parte lineal de la curva sigmoide incluyendo una constante multiplicativa en la exponencial. Hay muchas posibilidades como éstas que dependen fuertemente del problema que se esté resolviendo.

6.6.2. Dimensionamiento de la red. El número de capas y nodos que se necesitan, exactamente, para resolver un problema concreto depende de la naturaleza de este. En general, tres capas son suficientes. Hay veces, sin embargo, en que parece que un problema es más fácil de resolver con más de una capa oculta. En este caso, *más fácil* quiere decir que la red aprende más de prisa.

El tamaño de la capa de entrada suele venir dictado por la naturaleza de la aplicación. A menudo, es posible determinar el número de nodos de salida decidiendo si se desean valores analógicos o valores binarios en las unidades de salida.

Determinar el número de unidades que hay que utilizar en la capa oculta no suele ser tan evidente como lo es para las capas de entrada y de salida. La idea principal consiste en utilizar el menor número posible de unidades en la capa oculta, porque cada unidad supone una carga para la CPU durante las simulaciones. Por supuesto, en un sistema que esté construido el hardware en su totalidad (un procesador por cada elemento de proceso), la carga adicional para la CPU no es un factor tan importante (sin embargo, la comunicación entre procesadores sí que puede ser problema). No es recomendable ofrecer unas ideas generales muy concretas, salvo para decir que, en la experiencia, para redes de un tamaño razonable (cientos o miles de entradas), el tamaño de la capa oculta sólo necesita ser una fracción relativamente pequeña del que tiene la capa de entrada. Si la red no llega a converger para llegar a una solución, cabe en lo posible que se necesiten

más nodos ocultos. Si converge, se puede probar con un número inferior de nodos ocultos y determinar un tamaño final basándose en el rendimiento global del sistema.

También es posible eliminar unidades ocultas que resulten superfluas. Si se examinan los valores de los pesos de los nodos ocultos periódicamente, a medida que se entrena la red, se verá que los pesos de ciertos nodos cambian muy poquito con respecto a sus valores iniciales. Estos nodos pueden no estar participando en el proceso de aprendizaje, y quizá baste con un número menor de unidades ocultas. También existe un método automático, desarrollado por Rumelhart, para podar los nodos innecesarios de la red.

6.6.3. Pesos y parámetros de aprendizaje. Los pesos deberían recibir unos valores iniciales pequeños y aleatorios (digamos entre $\pm 0,5$), al igual que los términos de tendencia, θ_i , que aparecen en las ecuaciones de la entrada neta a cada unidad. Se suele tratar a este valor de tendencia como a un peso más, que está conectado a una unidad ficticia cuya salida es siempre 1.

Para ver la forma en que funciona este procedimiento, recuérdese la Ecuación 6.11:

$$neta_{pk}^o = \sum_{j=1}^L w_{kj}^o i_{pj} + q_k^o$$

Haciendo las definiciones $q_k^o \equiv w_{k((L+1))}^o$ e $i_{p(L+1)} \equiv 1$ se puede escribir lo siguiente:

$$neta_{pk}^o = \sum_{j=1}^{L+1} w_{kj}^o i_{pj}$$

Por tanto, q_k^o es tratado exactamente igual que un peso, y como tal participa en el proceso de aprendizaje. Otra posibilidad es, simplemente, eliminar por completo los términos de tendencia; su utilización es opcional.

La selección de un valor para el parámetro de velocidad de aprendizaje η , tiene un efecto significativo en el rendimiento de la red. Normalmente, η debe ser un número pequeño (del orden de 0,05 a 0,25) para asegurar que la red llegue a asentarse en una solución. Un valor pequeño de η significa que la red tendrá que hacer un gran número de iteraciones, pero este es el costo en que se incurre. Suele ser posible incrementar el valor de η a medida que progresa el aprendizaje. Aumentando a medida que disminuye el error de la red suele acelerar la convergencia incrementando el tamaño del paso conforme el error alcanza un valor mínimo, pero la red puede rebotar, alejándose demasiado del valor mínimo verdadero, si η llegara a ser demasiado grande. Otra forma de incrementar la velocidad de convergencia consiste en utilizar una técnica llamada **momento**. Cuando se calcula el valor del cambio de peso, Δ_{pw} , se añade una fracción del cambio *anterior*. Este término adicional tiende a mantener los cambios de peso en la misma dirección: de aquí el término *momento*. Las ecuaciones de cambio de pesos de la capa de salida pasan entonces a ser:

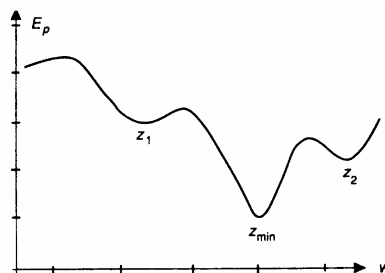
$$w_{kj}^o(t+1) = w_k^o(t) + \mathbf{hd}_{pk}^o i_{pj} + \alpha \Delta_p w_{kj}^o(t-1) \quad (6.33)$$

Se puede plantear una ecuación similar para la capa oculta. En la Ecuación 6.33 α es el parámetro de momento, y suele dársele un valor positivo menor que 1. También es opcional el uso del término de momento.

Para terminar, existe la posibilidad de convergencia a un **mínimo local** del espacio de pesos. La Figura 29 ilustra la idea. Una vez que una red se asienta en un mínimo, sea local o global, cesa el aprendizaje. Si se alcanza

un mínimo local, el error de las salidas de la red puede seguir siendo excesivamente alto. Afortunadamente, este problema no parece causar grandes dificultades en la práctica. Si una red deja de aprender antes de llegar a una solución aceptable, un cambio del número de nodos ocultos o de los parámetros de aprendizaje suele resolver el problema; o bien se puede simplemente volver a empezar con un conjunto distinto de pesos originales. Cuando una red alcanza una solución aceptable, no hay garantía de que *haya* alcanzado el mínimo global en lugar de uno local. Si la solución es admisible desde el punto de vista del error, no importa si el mínimo es local o global, y ni siquiera importa si el entrenamiento se ha detenido en algún momento previo a alcanzar un verdadero mínimo.

Figura 29. Sección transversal de una superficie de error hipotética dentro del espacio de pesos.



Fuente: WERBOS P.J. "Backpropagation and neurocontrol: A review and prospectus".³⁶

7. METODOLOGÍA DE AJUSTE UTILIZANDO REDES NEURONALES ARTIFICIALES

Como se ha propuesto desde el comienzo del trabajo de investigación, se pretende efectuar el ajuste de los parámetros de la ecuación de estado de Peng-Robinson a datos experimentales PVT utilizando redes neuronales. En los capítulos anteriores se explicó la construcción del simulador de pruebas a partir de las ecuaciones, métodos y cálculos utilizados en el comportamiento de fases. Con ello se llegó a la construcción del mismo, este software permite predecir el comportamiento volumétrico de muestras de fluido.

Se mostró en otro capítulo la forma como tradicionalmente se efectúa el ajuste de las ecuaciones de estado a los datos PVT experimentales, su metodología y la forma como diferentes autores trabajan el tema.

También se presentó las bases conceptuales y las metodologías de entrenamiento de las redes neuronales, con el fin mostrar las ventajas que ofrece este tipo de solución a problemas numéricos en ingeniería.

En este capítulo se presentará la metodología que combina la simulación de pruebas PVT con el enfoque metaheurístico de las redes neuronales y que permitirá la predicción de los parámetros que permitan al simulador establecer las relaciones entre las variables físicas más cercanas a los datos de pruebas reales.

7.1. DISEÑO DEL PROYECTO

Para el diseño y planeación de este proyecto de neuro-computación se utilizó la metodología general que rige en este tipo de proyectos, ya que esta permite obtener un lineamiento general guía en su desarrollo. De forma esquemática una guía general de los pasos a seguir para aplicar las redes neuronales en la solución de un problema es la siguiente³⁷:

Figura 30. Ciclo de vida de un proyecto utilizando neurocomputación.



Fuente: INSTITUTO MEXICANO DEL PETRÓLEO. "Uso de Inteligencia Artificial en la Inversión de Registros Eléctricos y Electromagnéticos".³⁷

Los pasos para el desarrollo de este proyecto de neurocomputación específicamente son los siguientes:

1. *Identificación y estudio del problema.* En esta etapa se plantea si el problema es adecuado para intentar resolverlo utilizando redes neuronales.

En el caso del ajuste de datos de pruebas reales, se ha visto que es un problema muy adecuado para el uso de esta técnica. Diversos estudios en la temática han mostrado que el proceso de ajuste se asemeja a una regresión múltiparamétrica y multiobjetivo, ya que se pretende ajustar los valores de varios parámetros de la ecuación de estado de manera que permitan reproducir todos los datos experimentales y a su vez mantener los valores de estos parámetros dentro de rangos que les permitan seguir representando las propiedades físico-químicas de las especies químicas que componen los fluidos^{28,29,30,38,39}.

Existen diversos parámetros en la ecuación de estado que son susceptibles de ser ajustados para que la Ecuación represente fielmente los datos PVT experimentales. Estos parámetros son:

- Ω_a y Ω_b de los compuestos puros.
- Coeficientes de Interacción Binaria (CIB) δ_{ij} .
- Propiedades críticas (P_c , T_c , V_c) y factor acéntrico (ω) de la fracción pesada (C_7^+).

Varios estudios^{25,27,40} hacen énfasis sobre la importancia y peso específico que dos de los parámetros tienen sobre el ajuste: El factor acéntrico de la fracción pesada ($\omega_{C_7^+}$) y el coeficiente de interacción binaria entre el metano y la fracción pesada ($\delta_{C_1-C_7^+}$). Por ello estos parámetros son los parámetros que se buscan ajustar en este proyecto.

La utilización de una topología neuronal adecuada permitirá que los dos objetivos -minimizar el error entre los datos experimentales y los predichos, y la utilización de varias funciones a minimizar- sean manejados de una mejor manera, permitiendo la generalización del ajuste por medio del entrenamiento de la red neuronal propuesta, sin perder en sí el sentido físico que representan los parámetros a ajustar en la ecuación de estado. En la literatura se presentan varias topologías neuronales que permiten efectuar cálculos estadísticos^{41,42,43}. Para el problema planteado se pretende utilizar la *Red Perceptrón Multicapa* como se puede ver en la siguiente tabla:

Tabla 13. Modelos neuronales y técnicas estadísticas.

MODELO NEURONAL	TÉCNICA ESTADÍSTICA PARALELA
Perceptrón Simple (nodo tipo umbral)	Análisis Discriminante
Perceptrón Simple (nodo sigmoideo)	Regresión logística
Adeline	Regresión lineal
<u>Perceptrón Multicapa</u>	<u>Regresión no Lineal Simple</u> <u>Regresión no Lineal Multivariada</u>
Aprendizaje Hebbiano no Supervisado	Análisis de Componentes Principales
Red Simple de Kohonen (competitiva)	Análisis Cluster Mínimos Cuadrados
Cuantificación de Vectores (LVQ)	Análisis Discriminante (Vecindad)
Funciones de Base Radial (RBF)	Métodos de Kernel Regression
Mapas Auto-Organizativos de Kohonen	Escalas Multidimensionales

Fuente: MARTÍN DEL BRÍO B. Y SANZ MOLINA A. "Redes Neuronales y Sistemas Borrosos".⁴⁴

2. *Diseño y construcción de un prototipo que ofrezca soluciones al problema.* En esta etapa se define la arquitectura de la red, el tipo de algoritmo de entrenamiento, el tipo de pre-procesamiento y/o codificación de los datos de entrenamiento. Aquí se determinará con que información se entrenará a la red (proporción de datos sintéticos y datos de campo).

Para el problema planteado se establecieron los siguientes prototipos:

- Redes Perceptrón multicapas: se probaron redes de 3 capas con la siguiente topología: 2-4-2, 2-5-2, 2-6-2 y de dos capas: dos entradas y dos salidas.
 - Normalización y desnormalización de datos.
 - Cálculo de factor de momentum.
 - Datos sin codificar, se trabajaron datos en formato numérico (números reales).
 - La tasa de aprendizaje utilizada es de 0.01 y el error máximo tolerable es del orden de 0.15.
3. *Entrenamiento de la red.* En esta etapa se entrena la red planteada con los datos para que permita estimar los mejores valores de solución al problema.

En concordancia con las observaciones de la independencia del fenómeno físico en sí, en cada muestra de laboratorio, se probó que la mejor manera de establecer los parámetros de regresión era un entrenamiento tipo backpropagation con datos de entrada individuales para cada red. También se estableció que los datos de entrenamiento y la cantidad de los mismos sean a criterio del usuario, quien con su experticia determinará los límites de las variables de entrada.

4. *Optimización, prueba y validación del prototipo.* En esta etapa se determina si se debe modificar la arquitectura de la red, si se requieren más datos de entrenamiento o si es necesario cambiar el modelo original incluyendo el algoritmo de entrenamiento (en esta fase se pone a prueba el diseño y se mejora en caso de ser necesario).

En este paso se observó que la calidad de los resultados depende solamente de la habilidad del usuario para determinar los límites físicos de los datos de entrada. La topología utilizada que mejor resultados ofrece depende de la muestra utilizada.

5. *Mantenimiento del modelo.* Esta fase se debe de realizar continuamente después de que se ha construido el sistema, lo cual permite realizar un ajuste necesario al modelo a medida que el problema va cambiando. Para el ajuste de datos utilizando las redes neuronales el mantenimiento de la red es de forma individual ya que cada muestra entrenada es en sí una red y puede ser reentrenada con otras variables de entrenamiento si fuera necesario.

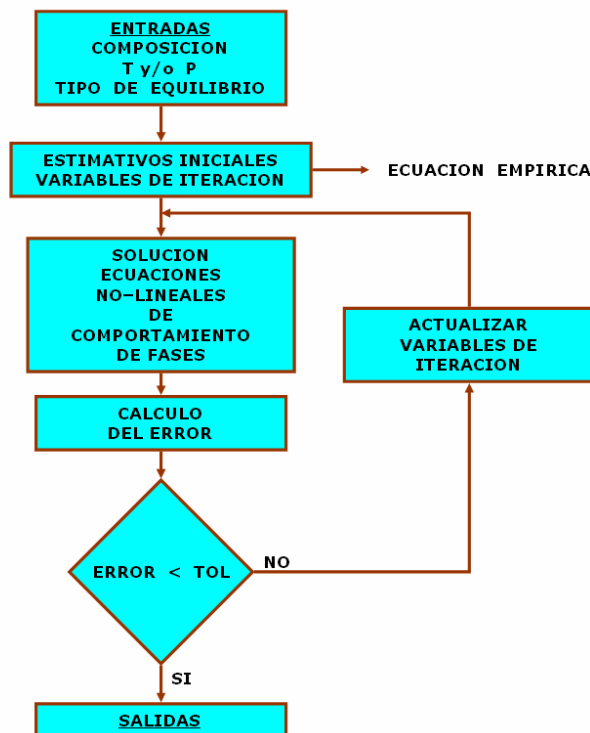
7.2. METODOLOGÍA DESARROLLADA

7.2.1. Planteamiento general. Como se ha esbozado en este proyecto de investigación, la solución numérica del comportamiento de fases de fluidos de yacimientos petrolíferos utilizando la Ecuación de Estado es la solución de un sistema de ecuaciones altamente no-lineales, en el cual existen varios parámetros que representan las moléculas de las especies químicas presentes en los fluidos petrolíferos. Este esquema iterativo de solución se presenta en la figura 31.

Los diferentes métodos de ajuste de los parámetros buscan modificarlos de tal manera que este sistema arroje datos de volúmenes, relaciones de equilibrio y masa que sean iguales o representen los valores estimados de manera experimental (experimentos PVT).

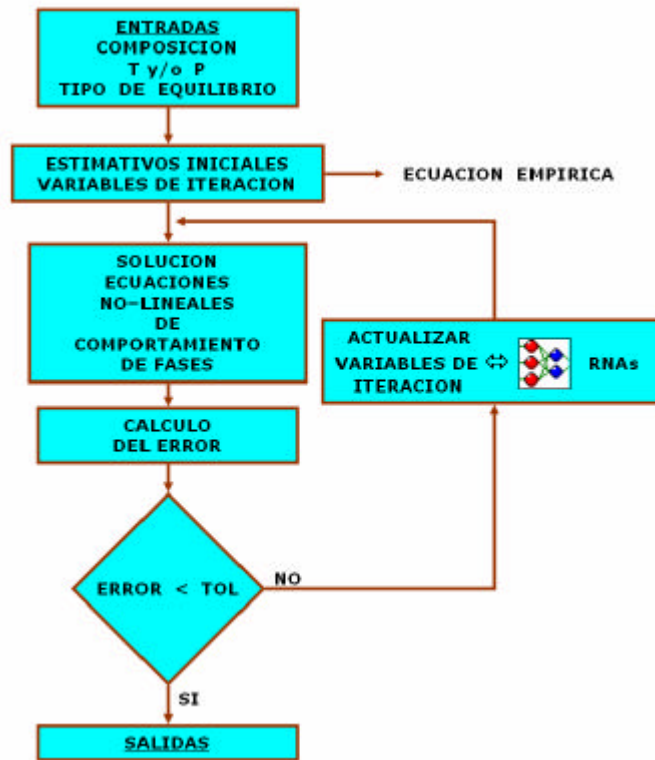
Se propone con esta metodología de ajuste escoger de manera directa y utilizando las redes neuronales los valores de dos de los parámetros que permitan establecer las relaciones de equilibrio, masa y volumen de las fases presentes en los experimentos PVT. Un esquema del cambio propuesto se muestra en la figura 32.

Figura 31. Esquema de solución de las ecuaciones del comportamiento de fases.



Fuente: CALVETE G., FERNANDO E. y RANGEL C., NOHORA P. "Simulación de Experimentos PVT para Yacimientos de Gas Condensado y A ceite Volátil".⁴⁵

Figura 32. Esquema de solución de las ecuaciones del comportamiento de fases utilizando redes neuronales.

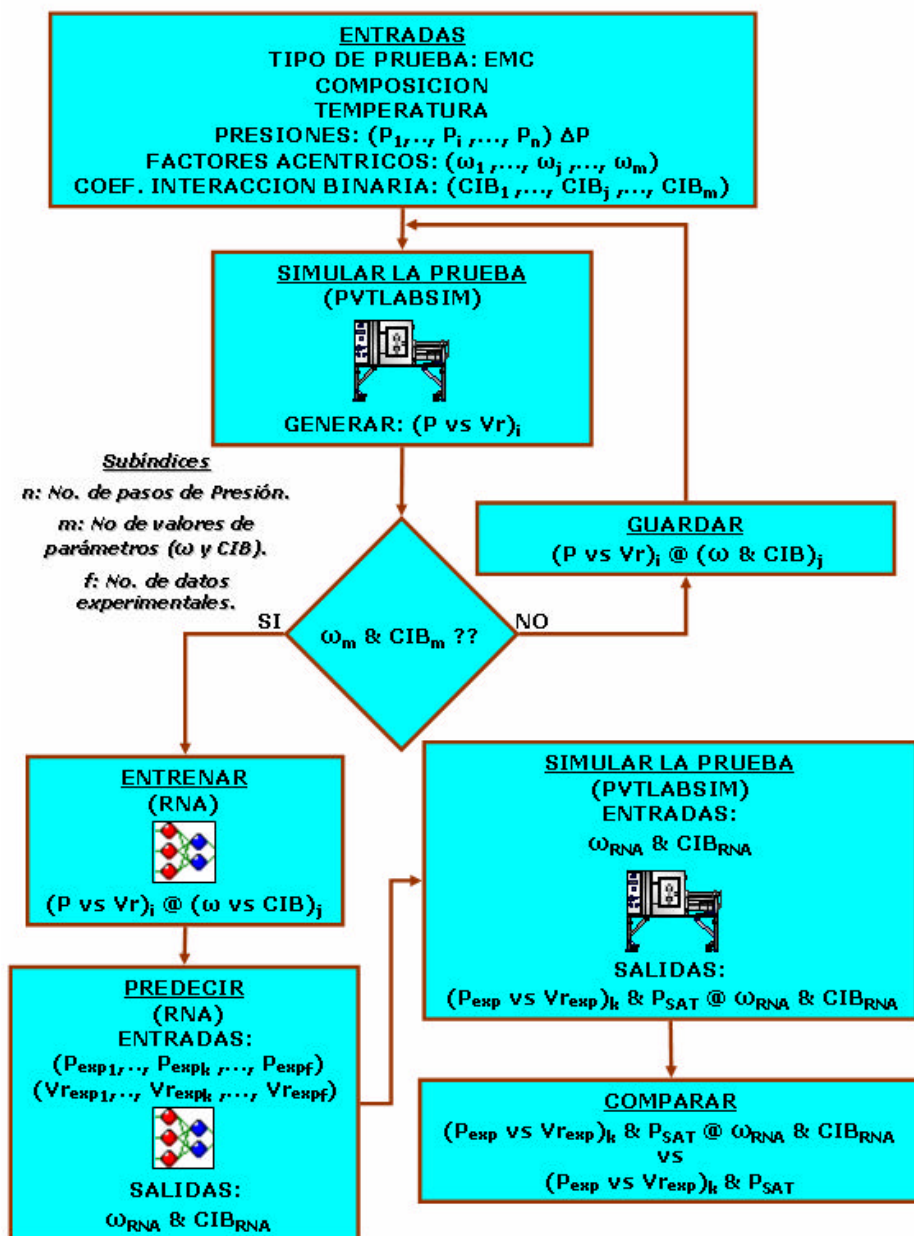


Fuente: Autor

7.2.2. Metodología propuesta. En base a los objetivos del proyecto y en los cambios propuestos para la utilización de las redes neuronales en la solución del sistema de ecuaciones que describen el comportamiento de fases, se plantea un esquema jerárquico que permita al software PVTLABSIM^{45,46} de simulación de pruebas PVT generar datos de volumen relativo de un experimento de expansión a masa constante para que un software de creación y entrenamiento de redes neuronales (GERENE^{47,48} u otro) pueda encontrar los valores óptimos de los parámetros de la Ecuación de Estado que permitan al simulador predecir fielmente los valores

experimentales PVT. Un esquema del proceso se puede ver en detalle en la figura 33.

Figura 33. Esquema de la metodología propuesta para el ajuste de parámetros.



Fuente: Autor

7.2.3. Datos de entrada. Los datos iniciales que se utilizan en la metodología son los requeridos por el simulador de pruebas: composiciones de la muestra, temperatura del experimento y las condiciones experimentales a ser simuladas (presiones experimentales).

Para alimentar la red neuronal se requiere generar las diferentes relaciones que existen entre presión y volumen relativo calculados a diferentes valores del coeficiente de interacción binaria y el factor acéntrico. La tabla 14 muestra los datos de entrenamiento de la red.

Tabla 14. Modelos neuronales y técnicas estadísticas.

PRESIÓN (P)	VOL. RELATIVO (Vr)	FACTOR ACÉNTRICO C_{7+} (wC_{7+})	COEF. DE INTERACCIÓN BINARIA (dC_1-C_{7+})
P_1	$Vr_{1,1}$ $Vr_{1,2}$... $Vr_{1,j}$... $Vr_{1,m}$	$W_{1,1}$ $W_{1,2}$... $W_{1,j}$... $W_{1,m}$	$CIB_{1,1}$ $CIB_{1,2}$... $CIB_{1,j}$... $CIB_{1,m}$
P_2	$Vr_{2,1}$ $Vr_{2,2}$... $Vr_{2,j}$... $Vr_{2,m}$	$W_{2,1}$ $W_{2,2}$... $W_{2,j}$... $W_{2,m}$	$CIB_{2,1}$ $CIB_{2,2}$... $CIB_{2,j}$... $CIB_{2,m}$
... P_i ...	$Vr_{i,1}$ $Vr_{i,2}$... $Vr_{i,j}$... $Vr_{i,m}$	$W_{i,1}$ $W_{i,2}$... $W_{i,j}$... $W_{i,m}$	$CIB_{i,1}$ $CIB_{i,2}$... $CIB_{i,j}$... $CIB_{i,m}$
P_n	$Vr_{f,1}$ $Vr_{f,2}$... $Vr_{f,j}$... $Vr_{n,m}$	$W_{f,1}$ $W_{f,2}$... $W_{f,j}$... $W_{n,m}$	$CIB_{f,1}$ $CIB_{f,2}$... $CIB_{f,j}$... $CIB_{n,m}$

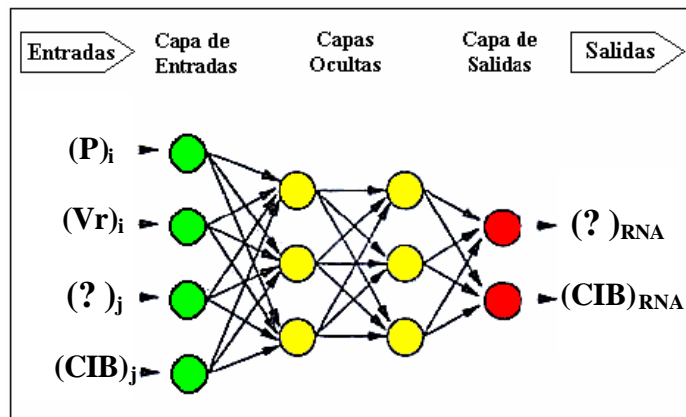
Fuente: MARTÍN DEL BRÍO B. Y SANZ MOLINA A. "Redes Neuronales y Sistemas Borrosos".⁴⁴

1. Definir el rango de presiones de estudio: Se puede escoger cualquier rango de presiones entre la inicial y la final del experimento, para tener la mayor consistencia y poder verificar el comportamiento volumétrico de la muestra se propone utilizar la presión inicial y la final experimental. Las presiones intermedias se generan al cambiar diferencialmente la presión ($\Delta P=10$ psig). ($P_1, P_2, \dots, P_i, \dots, P_f$).
2. Se definen los diferentes datos de factor acéntrico y coeficiente de interacción binaria. Estos deben estar entre los valores del C_7 y del C_{45} puros. Existen multitud de correlaciones y procedimientos para su cálculo, en este estudio se utilizaron dos tipos: Los valores del software PVTi de ECLIPSE300 que son generados a partir de la tabla de valores generalizados de Katz-Firoozabadi-Whitson – KFW - (Tabla 6). Además se utilizaron los valores del software comercial HYSYS. El usuario puede definir sus propios valores a trabajar y según su criterio definir cuantos datos utilizar (dependiendo del tipo de fluido y la caracterización de la fracción pesada). ($W_{1,1} - CIB_{1,1}$) ... ($W_{n,m} - CIB_{n,m}$).

Nota: El software calcula por defecto los valores de estos parámetros para la fracción pesada - C_7^+ -. Se recomienda tener en cuenta este valor ya que es un promedio de las tendencias que se pueden presentar.

7.2.4. Entrenamiento y predicción con redes neuronales. Luego de tener definidos los datos de entrenamiento, se escoge la topología neuronal para ser entrenada. En la figura 34 se observa las entradas y salidas de la red.

Figura 34. Estructura de entrenamiento.



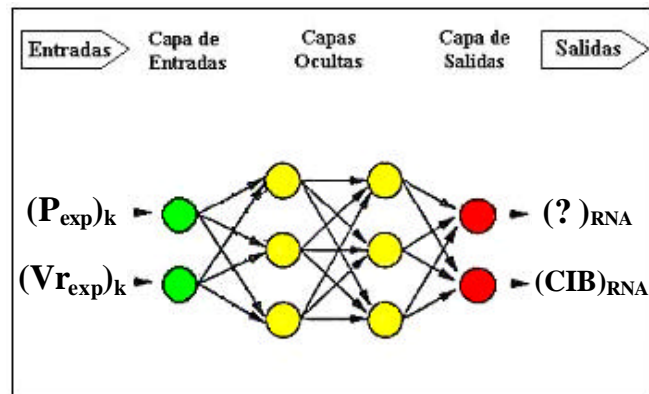
Fuente: Autor

En este estudio se entrenaron redes tipo perceptrón multicapa con las siguientes características:

- Topología: 2-4-2, 2-5-2, 2-6-2 y 2-2.
- Normalización y desnormalización de datos.
- Cálculo de factor de momentum.
- Datos sin codificar, se trabajaron datos en formato numérico (números reales).
- La tasa de aprendizaje utilizada es de 0.01 y el error máximo tolerable es del orden de 0.15.

La fase siguiente luego del entrenamiento es la predicción de datos. Para ello se utilizan los valores reales de la prueba (presión experimental y volumen relativo experimenta) con el fin de utilizar la red entrenada con los valores simulados y encontrar entre estos datos cuales son los que permiten calcular con el mínimo error los volúmenes relativos experimentales a las diferentes condiciones presentes en el laboratorio. La figura 35 representa la entrada y las salidas en la fase de predicción con la red entrenada.

Figura 35. Estructura de predicción.



Fuente: Autor

7.2.5. Validación de los parámetros de la predicción. Luego de predecir los valores de los parámetros por medio de la red neuronal, es necesario verificar su validez. Para ello, se utiliza de nuevo el simulador de pruebas. Con las presiones experimentales y los parámetros predichos por la red se generan los valores de volúmenes relativos y la presión de saturación de la muestra, para efectuar los análisis estadísticos que determinen la aproximación de la predicción.

Si es necesario el usuario puede cambiar de topología neural y variables de entrenamiento o los límites de los datos de entrenamiento (familias de carbonos presentes en la fracción pesada, correlaciones y tablas de datos utilizadas) para lograr un mejor ajuste.

7.3. SOFTWARE DE REDES NEURONALES

Existen muchos software que permiten diseñar, entrenar y optimizar proyectos de redes neuronales. La temática planteada en esta investigación presenta una particularidad muy especial, la habilidad y la experticia del profesional en esta área. Por lo anterior se evitó al máximo una automatización del software desarrollado, y se puede utilizar cualquier herramienta de la que se tenga disposición.

En el desarrollo de este proyecto de investigación se construyó una herramienta software que permite trabajar redes neuronales aplicadas a diferentes problemas de ingeniería (GENERE⁴⁷), el cual fue implementado por estudiantes de pregrado del grupo de investigación en Informática aplicada a Hidrocarburos de la Escuela de Ingeniería de Petróleos de la Universidad Industrial de Santander (UIS). Esta herramienta esta incluida en el software producto de esta investigación PVTLabRNA. Para estudiar su utilización consulte el tutorial que acompaña al software.

Otra aplicación de redes neuronales utilizada en el ajuste de los parámetros de la Ecuación de Estado es ANNBack desarrollada en Borland Builder C++ por los ingenieros de Sistemas Henry Alberto Cárdenas Gelvez y Carlos Alberto Brewer Ramos del grupo de investigación LINCE de la Escuela de Ingeniería de Sistemas UIS. Este software está incluido en la herramienta PVTLabRNA.

7.3.1. Software ANNBack. La herramienta *ANNBack* se divide en tres partes, que permiten cargar datos, entrenar la red y efectuar la evaluación y predicción de datos a partir de la estructura neural entrenada. A continuación

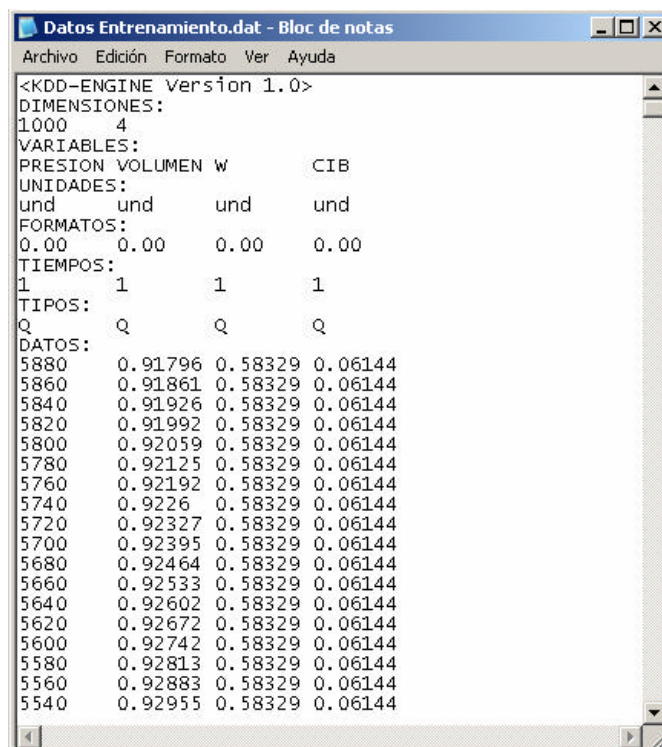
se presenta una descripción breve de su funcionamiento para el usuario interesado en su utilización.

7.3.1.1. Entrada de datos y entrenamiento. Los datos de entrada son la información de la red y los datos para su entrenamiento. La red utilizada es de tipo perceptrón multicapa y puede tener cualquier cantidad de capas, las cuales se deben introducir como el número de neuronas por capa separadas por comas.

Ej. 5 capas: 3;4;2;2;3: *3 neuronas de entrada, tres capas ocultas con 4,2,2 neuronas respectivamente y 3 neuronas de salida.*

Los datos de entrenamiento se deben introducir por medio de un archivo de datos con separación por tabulaciones como se muestra en la figura 36.

Figura 36. Archivo de entrada de datos para el software ANNBack.



```
<KDD-ENGINE Version 1.0>
DIMENSIONES:
1000 4
VARIABLES:
PRESION VOLUMEN w CIB
UNIDADES:
und und und und
FORMATOS:
0.00 0.00 0.00 0.00
TIEMPOS:
1 1 1 1
TIPOS:
Q Q Q Q
DATOS:
5880 0.91796 0.58329 0.06144
5860 0.91861 0.58329 0.06144
5840 0.91926 0.58329 0.06144
5820 0.91992 0.58329 0.06144
5800 0.92059 0.58329 0.06144
5780 0.92125 0.58329 0.06144
5760 0.92192 0.58329 0.06144
5740 0.9226 0.58329 0.06144
5720 0.92327 0.58329 0.06144
5700 0.92395 0.58329 0.06144
5680 0.92464 0.58329 0.06144
5660 0.92533 0.58329 0.06144
5640 0.92602 0.58329 0.06144
5620 0.92672 0.58329 0.06144
5600 0.92742 0.58329 0.06144
5580 0.92813 0.58329 0.06144
5560 0.92883 0.58329 0.06144
5540 0.92955 0.58329 0.06144
```

Fuente: Autor

Como se puede observar, existen varios campos de entrada de datos:

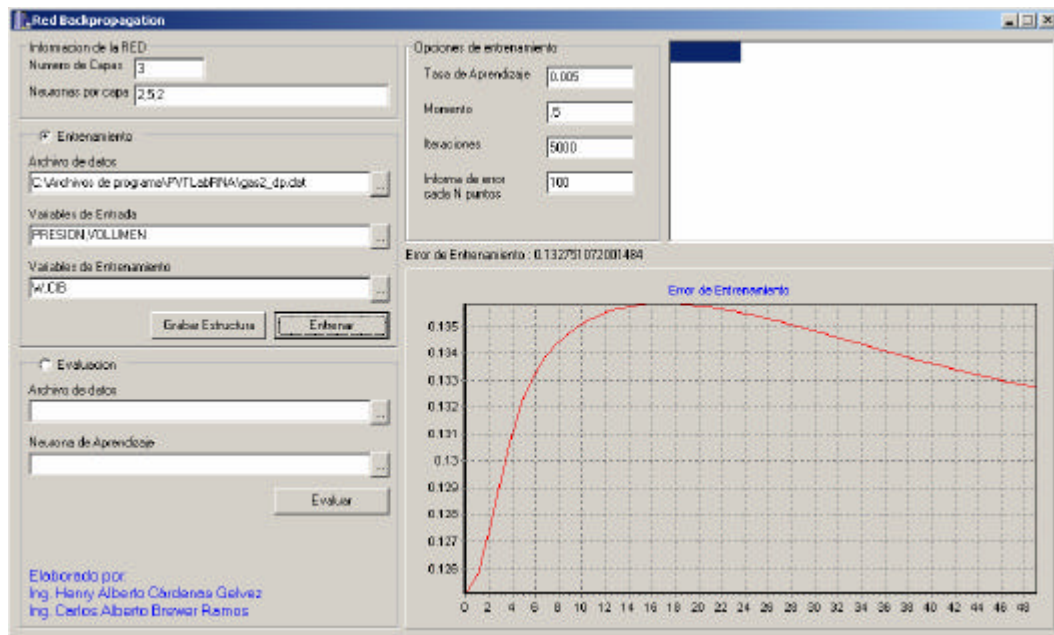
- **DIMENSIONES:** Son respectivamente las dimensiones de los datos. Filas x Columnas.
- **VARIABLES:** El nombre de las variables de entrenamiento
- **UNIDADES:** Las unidades de las variables. Si las hay.
- **FORMATOS:** Son variables de control interno del software, hay que agregarlas sin modificarlas.
- **TIEMPOS:** Para una simulación dinámica se pueden modificar, en nuestro caso se agregan igual.
- **TIPOS:** El tipo de datos a entrenar. Q por defecto trabaja datos reales.
- **DATOS:** Los datos de entrenamiento separados por tabulaciones.

7.3.1.2. Entrenamiento de la red. La red creada y con datos e puede entrenar al ver y modificar los parámetros como se muestra en la figura 37. Los campos que el usuario debe modificar son los siguientes:

- **VARIABLES DE ENTRADA:** El nombre de las variables de entrada (deben estar en el archivo de datos).
- **VARIABLES DE SALIDA:** El nombre de las respectivas variables de salida (guardadas en el archivo de datos).
- **TASA DE APRENDIZAJE:** Factor de entrenamiento que permite graduar la forma como la red va a aprender de los datos de entrada. Se recomienda comenzar a trabajar con 0.01 e irlo disminuyendo a la mitad.
- **MOMENTO:** Factor que permite suavizar la tendencia del error permisible. Este valor fluctúa entre 0 y 1. Se recomienda comenzar con un valor de 0.5 e irlo incrementando agregando cada vez la mitad de su valor actual.
- **ITERACIONES:** Son las iteraciones máximas que debe efectuar la red. Se recomienda un valor de 10.000 para evitar el sobreentrenamiento.

- **INFORME DE ERROR:** Este valor representa cada cuantas iteraciones se mostrará el error en la curva del error permisible.

Figura 37. Entrenamiento de redes con el software ANNBack.

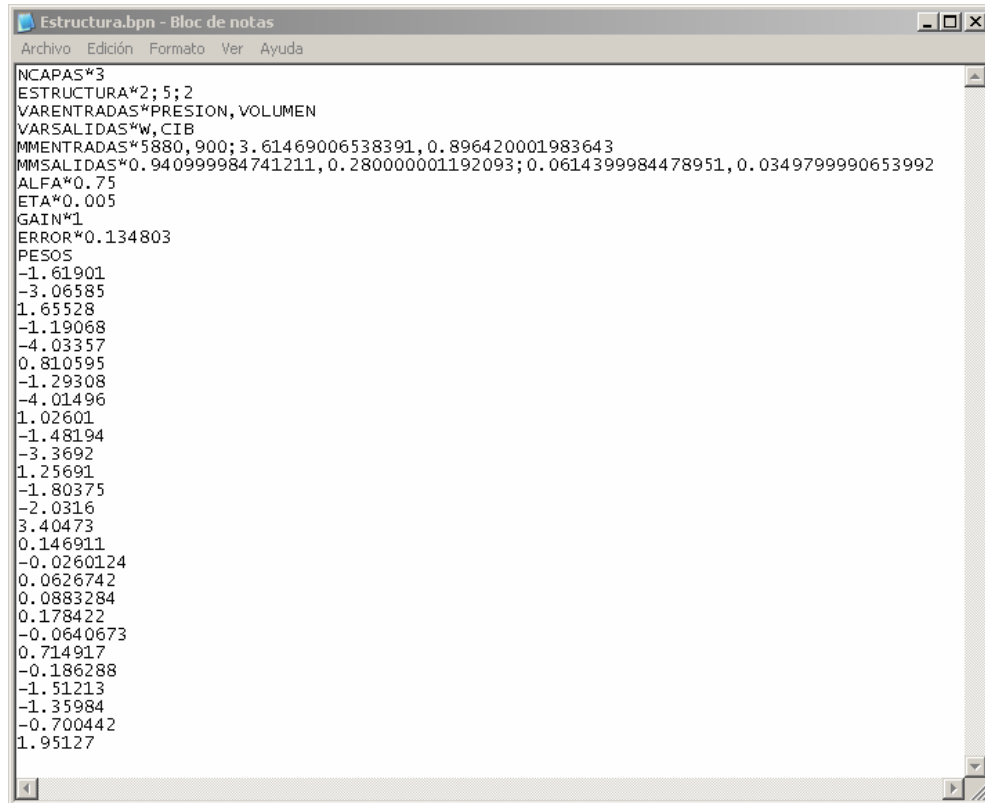


Fuente: Autor

El entrenamiento comienza cuando el usuario oprime el botón ENTRENAR. Esta etapa finaliza cuando el usuario modifica los parámetros de entrenamiento de tal manera que la curva de error permisible mostrada tenga un valor mínimo y se mantenga con tendencia decreciente a través de las iteraciones. Para grabar la estructura de la red al finalización de la sesión se debe oprimir el botón GRABAR ESTRUCTURA.

El archivo de la estructura entrenada y grabada tiene la forma mostrada en la figura 38.

Figura 38. Entrenamiento de redes con el software ANNBack.



```
Estructura.bpn - Bloc de notas
Archivo Edición Formato Ver Ayuda

NCAPAS*3
ESTRUCTURA*2;5;2
VARENTRADAS*PRESION,VOLUMEN
VARSALIDAS*W,CIB
MMENTRADAS*5880,900;3.61469006538391,0.896420001983643
MMSALIDAS*0.940999984741211,0.280000001192093;0.0614399984478951,0.0349799990653992
ALFA*0.75
ETA*0.005
GAIN*1
ERROR*0.134803
PESOS
-1.61901
-3.06585
1.65528
-1.19068
-4.03357
0.810595
-1.29308
-4.01496
1.02601
-1.48194
-3.3692
1.25691
-1.80375
-2.0316
3.40473
0.146911
-0.0260124
0.0626742
0.0883284
0.178422
-0.0640673
0.714917
-0.186288
-1.51213
-1.35984
-0.700442
1.95127
```

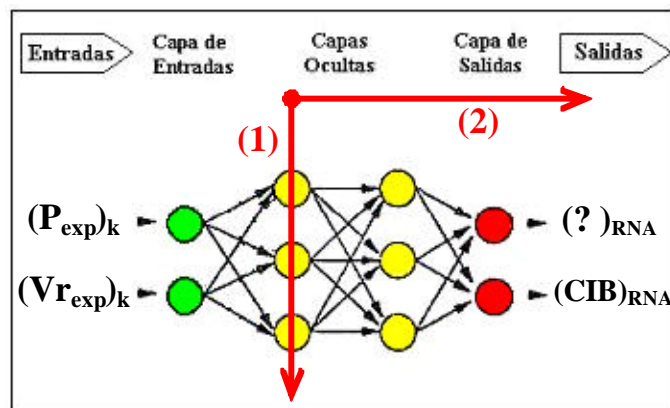
Fuente: Autor

Los campos presentes en el archivo de estructura son los siguientes:

- NCAPAS: Número de capas.
- ESTRUCTURA: Topología de la red.
- VARENTRADAS: Variables a entrenar.
- VARSALIDAS: Salidas de la red.
- MMENTRADAS: Dato máximo y mínimo de cada variable de entrada para efectuar la normalización.
- MMSALIDAS: dato máximo y mínimo de cada variable de salida para efectuar la normalización.
- ALFA: Momento utilizado.

- ETA: Tasa de aprendizaje utilizada.
 - GAIN: Indica la función de transferencia utilizada en la red (Sigmoidal=1).
 - ERROR: Error permisible final.
 - PESOS: Lista de los pesos finales de cada conexión entre neuronas.
- Antes de los pesos de cada neurona se coloca las *bias* obtenidas en la neurona. La forma de recorrer la red para guardar los pesos se puede ver en la figura 39.

Figura 39. Dirección del recorrido para guardar los pesos de cada conexión luego del entrenamiento.

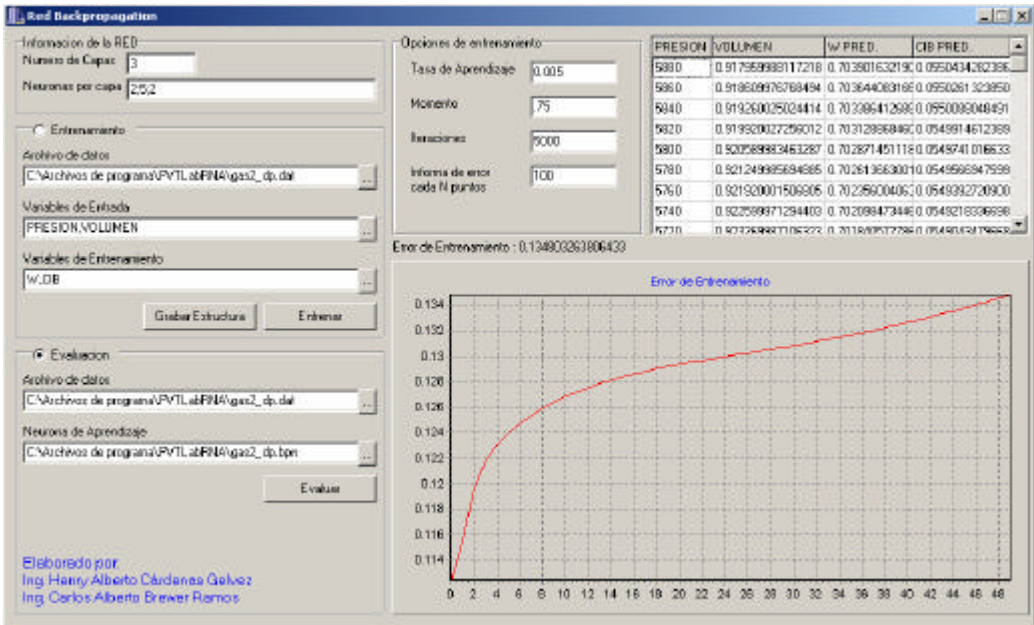


Fuente: Autor

Se puede observar que primero se recorre cada capa oculta en el sentido de la flecha (1) desde la primera (oculta) y se continúa de capa en capa hasta terminar en la capa de salida como lo indica la flecha (2).

7.3.1.3. Evaluación y datos de salida de la red. Para efectuar la evaluación de la estructura que se entreno se utiliza la sección del software llamada EVALUACION. Ver figura 40.

Figura 40. Evaluación de la red entrenada con el software ANNBack.

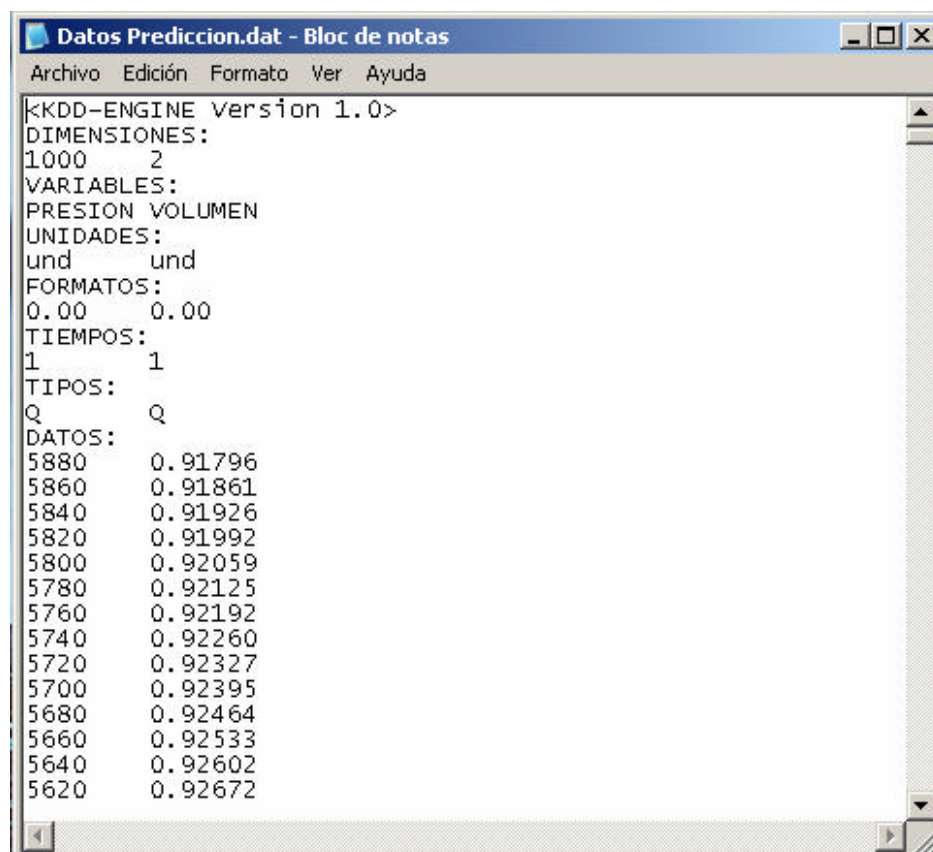


Fuente: Autor

Se introducen el nombre del archivo de datos de predicción (donde están los datos a los cuales se les va a encontrar las variables ajustadas por la red) y el nombre del archivo donde se encuentra grabada la estructura de la red. El archivo de datos para la evaluación se muestra en la figura 41.

Como se observa, este archivo conserva la estructura que tiene el archivo de datos para el entrenamiento, la diferencia solo radica en el número de variables y el listado de datos de cada una de ellas.

Figura 41. Archivo de datos para la evaluación de la red entrenada por el software ANNBack.



```

<KDD-ENGINE version 1.0>
DIMENSIONES:
1000      2
VARIABLES:
PRESION VOLUMEN
UNIDADES:
und      und
FORMATOS:
0.00      0.00
TIEMPOS:
1         1
TIPOS:
Q         Q
DATOS:
5880      0.91796
5860      0.91861
5840      0.91926
5820      0.91992
5800      0.92059
5780      0.92125
5760      0.92192
5740      0.92260
5720      0.92327
5700      0.92395
5680      0.92464
5660      0.92533
5640      0.92602
5620      0.92672

```

Fuente: Autor

7.4. AJUSTE DE LOS PARÁMETROS DE LA ECUACIÓN DE ESTADO UTILIZANDO ALGORITMOS GENÉTICOS

Con el ánimo de completar esta investigación y en beneficio del estudio comparativo, se desarrollo un módulo numérico que permite efectuar el ajuste de los parámetros de la ecuación de estado utilizando algoritmos

genéticos. Los resultados obtenidos utilizando esta técnica se publicaron en el X Congreso Colombiano del Petróleo, efectuado en Bogotá en el presente año. El artículo que resume esta metodología se muestra en el anexo B.

7.4.1. Conceptualización.⁴⁹ En los diferentes campos de la ciencia y la ingeniería se hace necesario optimizar procesos que están regidos por funciones con gran cantidad de mínimos locales, los procedimientos matemáticos convencionales no garantizan una buena respuesta, ya que estas técnicas de búsqueda pueden quedar atrapadas en esos mínimos locales. Es por esto, que se hace necesario utilizar otro tipo de herramientas para lograr mejores soluciones a los problemas de optimización. Algunas de estas herramientas son las llamadas técnicas evolutivas. Entre estas técnicas se encuentra los Algoritmos Genéticos, que utilizan criterios propios para realizar la búsqueda del mínimo global, estos criterios no tienen nada en común con los aplicados en los procedimientos netamente matemáticos permitiendo así encontrar soluciones acertadas sin manipular la heurística del problema.

La búsqueda del mínimo global conlleva a encontrar respuestas concretas a problemas típicos de la ingeniería de petróleos. Estas aplicaciones van desde el diseño de las facilidades de superficie, análisis nodal óptimo para el sistema de levantamiento artificial de pozos productores de gas y petróleo, pasando por la solución de la ecuación de estado utilizada en el cálculo del comportamiento de fases de los fluidos de yacimientos petrolíferos.

7.4.2. Metodología de ajuste utilizando algoritmos genéticos.⁵⁰ Para resolver el problema planteado se desarrolló una nueva metodología con Algoritmos Genéticos para ajustar el coeficiente (d_{C1-C7+}) y el factor acéntrico (γ_{C7+}) con los datos del experimento a composición constante.

Para ajustar el coeficiente (d_{C1-C7+}) y el factor acéntrico (φ_{C7+}), se tomo como base la metodología utilizada en la aplicación PVTLAB⁴⁵, la cual se basa en la simulación de los experimentos PVT de gases condensados y aceites volátiles. Esta metodología utiliza la ecuación de estado y los cálculos flash de equilibrio entre la fase gaseosa y la fase líquida de un fluido composicional de yacimiento para simular pruebas experimentales de laboratorio PVT.

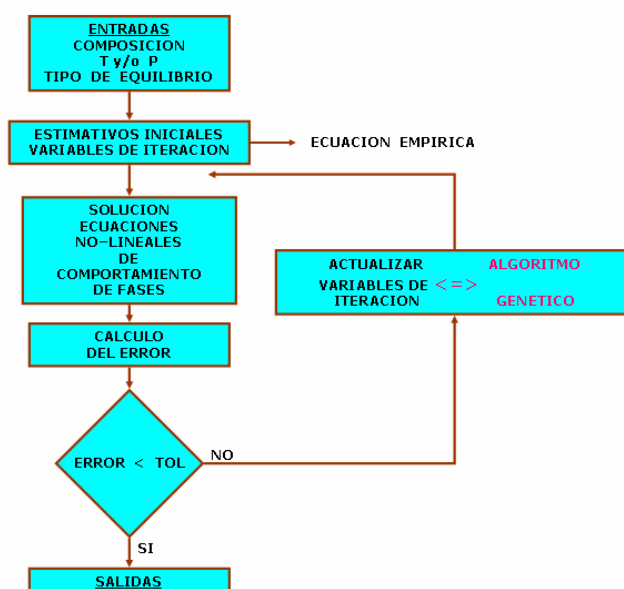
De manera general la Figura 42 muestra el desarrollo de los cálculos en la metodología planteada para trabajar con una ecuación de estado. Se muestra el cambio en la función de actualización de variables, donde se utiliza el algoritmo genético en reemplazo de la función tradicional basada en los parámetros de interacción binaria entre el metano y la fracción pesada (d_{C1-C7+}) y el factor acéntrico de la fracción pesada y de esta manera optimizar la selección de estos parámetros.

7.4.2.1. Cálculo del volumen relativo. En el algoritmo genético se utilizó como población inicial los valores de d_{C1-C7+} calculados por el método de Varotsis¹⁷ y el factor acéntrico de la fracción pesada (φ_{C7+}) calculado por medio de las correlaciones de propiedades físicas de los hidrocarburos propuesto por Whitson²⁰. De la población inicial de valores se generó cada nueva población por el cruce de los cromosomas binarios de estos datos. Para evaluar los individuos más aptos de la población se utilizó una función de aptitud definida por la comparación del error generado al evaluar los volúmenes relativos del experimento PVT dados por la ecuación de estado con los parámetros d_{C1-C7+} , φ_{C7+} obtenidos genéticamente contra la data experimental. Luego de tener cada nueva población se efectuó el mismo procedimiento hasta alcanzar la convergencia definida por el cálculo flash.

La población de datos en cada iteración se define por medio de las correlaciones de Varotsis y Whitson, las cuales son tomadas de experimentación con la fracción pesada, por lo cual representan fielmente el comportamiento de esta fracción a diferentes condiciones de temperatura y presión. Estas correlaciones experimentales aseguran al algoritmo genético tener siempre poblaciones que cumplan con los rangos de validez de los parámetros.

7.4.2.2. Cálculo de la presión de saturación. Para el cálculo de la presión de saturación el procedimiento con algoritmos genéticos es equivalente al efectuado para el volumen relativo, la función de aptitud que se define en este caso es el error entre la presión de saturación dada por la ecuación de estado y la obtenida en la experimentación PVT. A su vez se utilizan las mismas correlaciones de propiedades de Varotsis y Whitson para la fracción pesada.

Figura 42. Esquema de solución de las ecuaciones del comportamiento de fases.



Fuente: Autor

Los resultados que se derivan de esta técnica se muestran en el estudio comparativo que se efectuó para diversas muestras de fluidos.

8. ANÁLISIS DE RESULTADOS Y ESTUDIO COMPARATIVO

Con el fin de validar el programa y comprobar la consistencia de las predicciones del comportamiento de fases, se utilizaron muestras de sistemas de hidrocarburos trabajados en la literatura²⁵

Para la evaluación de los resultados se utilizó la desviación absoluta promedio definida por:

$$DAP = \frac{\sum_{i=1}^n \left| \frac{\text{valor medido} - \text{valor calculado}}{\text{valor medido}} \right|}{n}$$

$$\%Error = \frac{(P_{exp} - P_{calc})}{P_{exp}} * 100$$

8.1. CÁLCULO DE LA PRESIÓN DE SATURACIÓN

Para la presión de saturación los resultados que se lograron están consignados en la tabla 15. Se muestran en la tabla 16 las estructuras neuronales que fueron utilizadas para generar estos datos.

Tabla 15. Comparación de resultados en el cálculo de la presión de saturación para diferentes sistemas de hidrocarburos por medio de diferentes procedimientos.

MEDICIÓN	GAS 2*		GAS 2**		GAS 5	
EXPERIMENTAL	4465 (190°F)		4430 (190°F)		4857 (267°F)	
(EDE PR)	3680	17.58%	3664	17.29%	4494	7.47%
(PVTLABSIM)	3899	12.68%	3827	13.61%	4808	1.00%
(PVTLABRNA) KFW	3810	14.7%	3766	15%	3420	29.6%
(PVTLABSIM) HYSYS	4003	10.6%	4378	1.17%	3515	27.6%
(PVTLAB_AG)	4435	0.68%	4415	0.34%	4840	0.35%

MEDICIÓN	OIL 1		OIL 2		OIL 3		OIL 4		OIL 6	
EXPERIMENTAL	2534.7 (180°F)		4475 (176°F)		2612 (180°F)		2562 (250°F)		2760.7 (234°F)	
(EDE PR)	ND		3344	25.27%	2195	15.96%	2259	11.83%	2412.7	12.61%
(AHMED)	ND		4502	-0.603%	2289	12.37%	2640	-3.044%	2930.7	-6.16%
(PVTLABSIM)	2494.3	1.59%	4478.5	-0.078%	2368.5	09.32%	2647.45	-3.335%	2907	-5.29%
(PVTLABRNA) KFW	2577	-1.6%	4738	5.8%	2378	10.2%	2123	17.13%	3001	-8.70%
(PVTLABRNA) HYSYS	2680	-5.76%	4981	11.3%	2390	10%	2201	14%	3011	-9.065%
(PVTLAB_AG)	2530.8	0.15%	4470.5	0.10%	2561	1.95%	2534.6	1.07%	2780.1	-0.7%

KFW: Denota que los datos de factor acéntrico y coeficiente de interacción binaria son tomados de las tablas generalizadas de Katz-Firoozabadi-Whitson.

HYSYS: Denota que los datos de factor acéntrico y coeficiente de interacción binaria son tomados del software HYSYS.

Fuente: Autor

Tabla 16. Estructuras neuronales utilizadas en el estudio de presión de saturación.

MUESTRA	No DE CAPAS	NEURONAS DE ENTRADA	NEURONAS OCULTAS	NEURONAS DE SALIDA	TASA DE APRENDIZAJE	MOMENTO	ITERACIONES	ERROR DE ENTRENAMIENTO	W	CIB
GAS2* KFW	3	2	5	2	0.00010	0.20	4000	0.1545	0.700	0.055
GAS2* HYSYS	3	2	5	2	0.00001	0.80	4000	0.278	0.740	0.067
GAS2** KFW	3	2	5	2	0.00990	0.75	4000	1.01016E-3	0.880	0.057
GAS2** HYSYS	3	2	5	2	0.00900	0.78	1000	1.149E-6	1.020	0.089
GAS5 KFW	3	2	5	2	0.00001	0.90	10000	0.11800	0.768	0.084
GAS5 HYSYS	3	2	5	2	0.00010	0.85	1000	0.16323	0.71	0.100
OIL1 KFW	3	2	5	2	0.00001	0.20	1000	0.211	0.680	0.052
OIL1 HYSYS	3	2	5	2	0.00001	0.75	3000	0.0615	0.715	0.063
OIL2 KFW	3	2	6	2	0.00020	0.60	1000	0.2212	0.689	0.049
OIL2 HYSYS	3	2	6	2	0.00001	0.85	1000	0.3319	0.710	0.060
OIL3 KFW	3	2	5	2	0.00001	0.85	1000	0.2264	0.690	0.050
OIL3 HYSYS	3	2	5	2	0.00010	0.77	1000	0.3314	0.720	0.060
OIL4 KFW	3	2	5	2	0.00010	0.50	1000	0.1140	0.699	0.077
OIL4 HYSYS	3	2	5	2	0.00010	0.75	1000	0.1580	0.740	0.095
OIL6 KFW	3	2	5	2	0.00010	0.85	1000	0.2340	0.670	0.085
OIL6 HYSYS	3	2	5	2	0.00010	0.70	1000	0.2150	0.740	0.072

Fuente: Autor

Las tablas anteriores reflejan puntualmente algunos resultados obtenidos al entrenar diferentes redes neuronales a cada muestra y con diferentes metodologías de cálculo de los parámetros (KFW o HYSYS). Además se tuvieron en cuenta los diferentes límites de la fracción pesada: para gases se asumió que en la fracción pesada se podrían encontrar moléculas máximo de C_{30}^+ y en los aceites que se podrían encontrar moléculas de C_{45}^+ .

Como se ha podido observar, los datos para el entrenamiento de la red afectan considerablemente la predicción y el ajuste que se pueda lograr. En la tabla 17 se muestran algunos de los valores de estos parámetros utilizados en la industria.

Tabla 17. Datos físicos para el entrenamiento de la red neuronal.

Parámetro	DESKTOP	K-F-W	GPSA	ECLIPSE	HYSYS	K-F	F-A
WC_7	0.34980	0.28000	0.34900	0.28000	-	-	-
CIB_{C1-C7}	0.02500	-	-	-	0.02933	0.03498	0.05232
WC_{18}	0.76640	0.63800	0.79000	0.63800	-	-	-
CIB_{C1-C18}	0.04320	-	-	-	0.07406	0.05304	0.08609
WC_{30}	1.23250	0.94100	-	0.94100	1.2325	-	-
CIB_{C1-C30}	-	-	-	-	0.10364	0.05999	0.11725
WC_{45}	1.82130	1.04800	-	1.04800	-	-	-
CIB_{C1-C45}	-	-	-	-	-	0.06480	0.14889

Fuente: Autor

DESKTOP: Software de simulación composicional.

K-F-W: Correlación de Katz-Firoozabadi-Whitson.

GPSA: Tablas generalizadas de compuestos puros.

ECLIPSE: Software de simulación composicional.

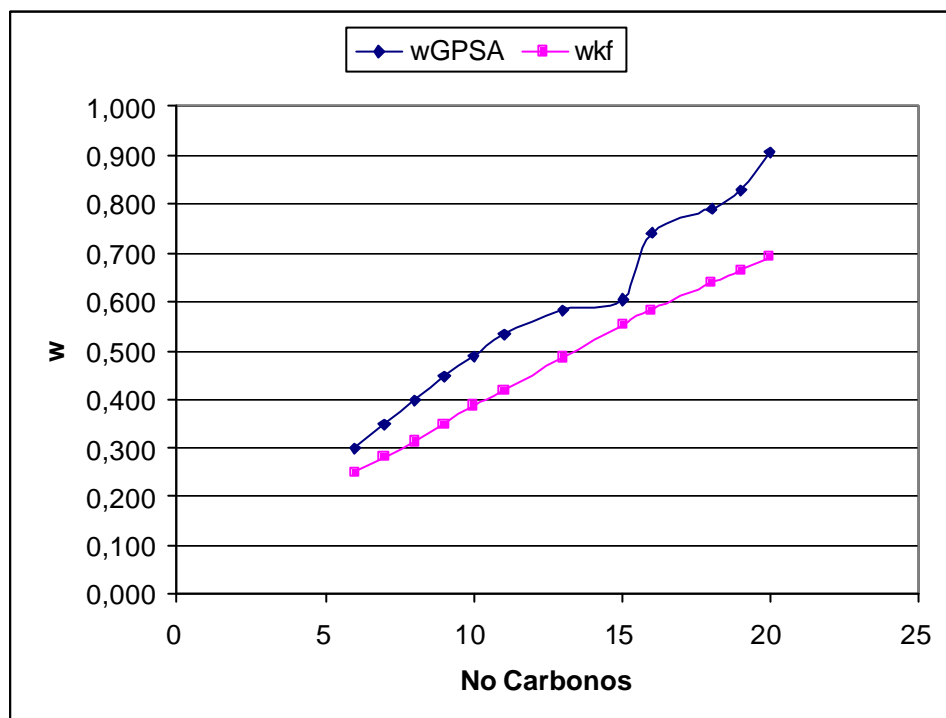
HYSYS: Software de ingeniería química.

K-F: Correlación de Katz-Firoozabadi.

F-A: Correlación de Firoozabadi-Arbabi.

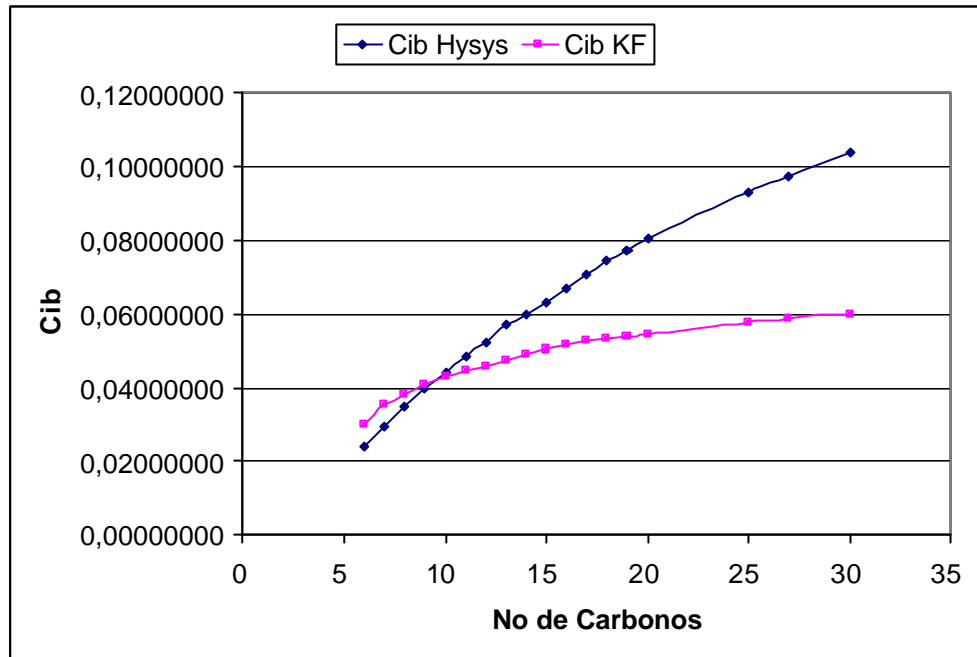
Las tendencias de los parámetros según el método de su cálculo se pueden observar en las figuras 43 y 44.

Figura 43. Comportamiento de factor acéntrico con respecto al número de carbonos.



Fuente: Autor

Figura 44. Comportamiento del Coeficiente de Interacción Binaria con respecto al número de carbonos.



Fuente: Autor

8.2. EXPANSIÓN A COMPOSICIÓN CONSTANTE

A partir de datos de presión se evaluó el comportamiento del volumen relativo, obtenido durante el desarrollo de esta prueba. En las tablas 18 y 19 se muestran los resultados de este experimento para el gas2*.

Este fluido se encuentra en yacimiento a su temperatura crítica, por ello en algunos momentos puede presentar punto de burbuja o punto de rocío. A esta temperatura la muestra presentó una presión de rocío de 4465 psia.

Tabla 18. Resultados de la Expansión a Masa Constante para el Gas2*.

PRESIÓN (PSIG)	% FRACCIÓN DE VOLUMEN				
	EXPERIMENTAL	PVTLABSIM	PVTLABRNA KFW	PVTLABRNA HYSYS	PVT_AG
4440	4,35	0	0	0	0
4420	47,38	0	0	0	33,56353
4388	50,82	0	0	0	34,53989
4339	51,64	0	0	0	35,66077
4300	51,94	0	0	0	36,63072
4180	51,95	0	0	0	39,34753
3993	51,32	0	0	0	43,05863
3780	50,07	35,75683	35.33761	41.0079	46,79117
3490	47,86	43,69655	44.16003	47.91214	48,57951
2998	42,96	45,78994	44.85577	42.65818	41,18708
2505	36,75	36,51561	35.47932	34.12666	33,54723
2000	28,88	27,32753	26.35064	25.54196	25,78090
1485	20,2	18,15569	17.3534	16.91081	17,42014
1058	13,06	11,05435	10.46563	10.22705	10,73368

Fuente: Autor

Tabla 19. Parámetros ajustados para la Expansión a Masa Constante del Gas2*.

MEDIDA	EXP.	PVTLABSIM	PVTLABRNA KFW	PVTLABRNA HYSYS	PVT_AG
d_{C1-C7+}	-	0,061441	0.057	0.067	0,10366484
$?_{C7+}$	-	0,58734834	0.7	0.74	0,53246336
PRESIÓN SAT	4465	3823,04	3810	4003	4434,63
DAP (%)	-	0,5862993	0.55083	0.59294	0,24740387
ERROR RELATIVO (%)	-	6,55397992	6.98	0.5508	1,16634421
ERROR (PSAT) (%)	-	14,3778	14.6696	10.6	0,68017086

Fuente: Autor

8.3. ANÁLISIS DE RESULTADOS

De los resultados numéricos obtenidos podemos inferir los siguientes análisis:

- La topología de red que mejor ajustó parámetros seleccionados fue de dos capas, dos neuronas de entrada y dos de salida. Esto es debido a la relación directa entre las variables: $V_r = f(P, W, CIB)$.
- La tasa de aprendizaje más adecuada para el entrenamiento de los datos de presión y volúmenes relativos en la experimentación PVT fueron valores entre 0.001 y 0.002.
- El momento utilizado en el entrenamiento osciló en valores promedio entre 0.25 hasta 0.5.
- El error permisible dentro del proceso se mantuvo aproximadamente constante para el rango de valores de entrenamiento, la variación de error generalmente fue del orden de $10E-5$ aproximadamente.
- El número de iteraciones utilizado en el entrenamiento fue del orden de 4000 a 6000. Esto permite una mejor representación del error y su comportamiento.
- El coeficiente de interacción binaria muestra tener mayor influencia sobre el desempeño y la capacidad de predicción de la red. Esto se puede observar comparando los resultados predichos por valores pequeños obtenidos por la correlación KFW contra los valores altos de HYSYS.

- Para seleccionar un par específico de parámetros (Fact. Acéntrico y CIB) estos deben ser uniformes, es decir utilizar una sola correlación para todo el rango de presiones a trabajar, con el fin de minimizar el efecto de ruido en los valores de entrenamiento de la red.
- De todos los valores presentes en la literatura al seleccionar la correlación y tabla de valores generalizados de Katz y Firozabadi se garantiza que la red será entrenada con el menor ruido posible, como se puede observar en las figuras 43 y 44.
- La regla del dedo gordo indica que se deben utilizar valores de C_7 a C_{30} para los gases y de C_7 a C_{45} para los aceites volátiles. Sin embargo al aplicar esto a una red neuronal sin una preclasificación de datos se corre el riesgo de inducir ruido y dificultando la determinación de la arquitectura óptima de la red.
- Se observó la notoria influencia que tiene el equi-espaciamiento de los datos utilizados en la predicción (datos experimentales) sobre todo aquellos con delta de presión que superan 400 psi, que contrastan con los 10 psi utilizados para entrenar la neurona.
- Se recomienda seleccionar correlaciones o tablas de datos de los parámetros a ajustar con una buena uniformidad.
- En el caso de la correlación de Katz Firozabadi se encontró que bajo-predice en el caso de los gases condensados y sobre-predice en el caso de los aceites volátiles por lo cual se recomienda hacer una preclasificación de los datos a entrenar.

- Al utilizar los datos de HYSYS se encontró que siempre sobre-predicen, ya que presenta los más altos valores de coeficiente de Interacción Binaria y su factor acéntrico es muy parecido a los valores de las otras correlaciones.

9. MANUAL DEL USUARIO

El presente manual busca mostrar al usuario del paquete PVTLabRNA 1.0 los procedimientos, cálculos y las diferentes opciones que el simulador maneja para el desarrollo y ajuste de datos PVT composicionales.

El software consta básicamente de dos componentes: una barra menú que permite establecer las acciones iniciales y los procedimientos más utilizados en un paquete software.

El segundo componente permite efectuar las pruebas PVT y esta compuesta de una serie de carpetas que van de derecha a izquierda. El usuario va llenando los datos solicitados en estas carpetas y al final podrá obtener los resultados de las pruebas en formatos gráfico, tabulado y texto.

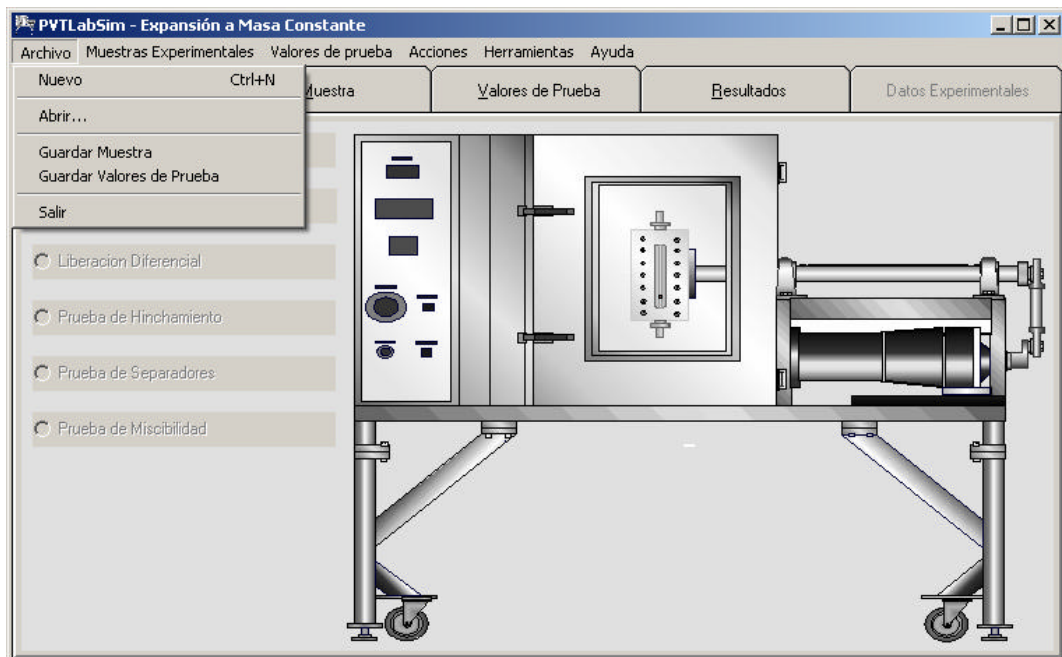
Este programa se elaboró en un lenguaje de desarrollo visual y sus requerimientos son:

- Hardware: PENTIUM II ó superior. RAM de 64Mb ó superior. Pantalla SVGA.
- Software: Windows 98 ó superior. Microsoft Word97 ó superior. Acrobat Reader 4.0.

A continuación se presentarán de manera detallada los componentes del software.

9.1. INTERFASE GENERAL

9.1.1. Barra de menú-Archivo.



Nuevo. Permite reiniciar todas las rutinas y entradas de datos.

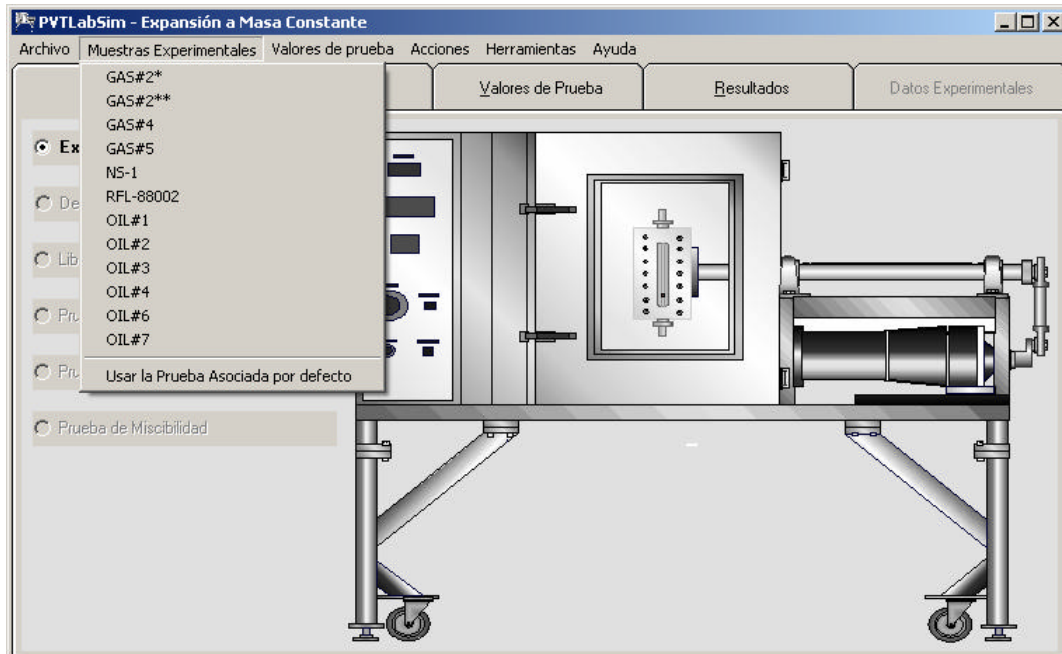
Abrir. Permite abrir los dos tipos de datos de entrada que utiliza el simulador:

- Datos de la muestra. (.PT0).
- Datos de la prueba. (.PT1).

Guardar Muestra. Genera el archivo con los datos de muestra. (.PT0).

Guardar Valores de Prueba. Genera el archivo con los valores de la prueba. (.PT1).

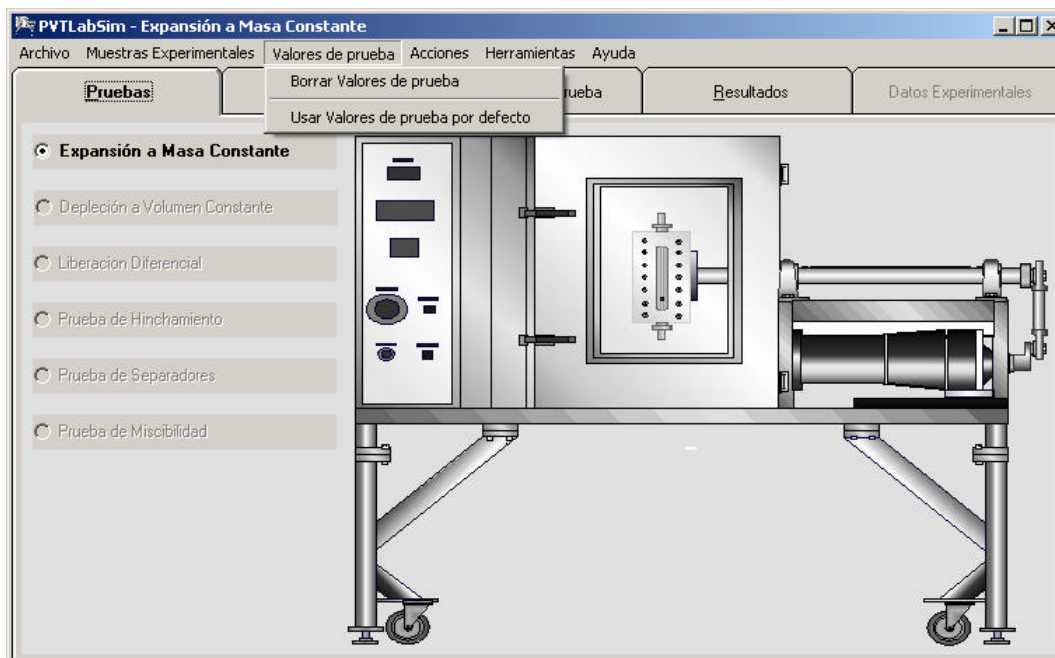
9.1.2. Barra de menú-Muestras Experimentales.



En este menú se presentan las diferentes muestras que se utilizaron para validar el software. Al elegir alguna de ellas se llenan los formularios de entrada de datos con los valores de la muestra elegida.

Con la opción *Usar la Prueba Asociada por Defecto*. Además de llenar los formularios de entrada de datos de la muestra selecciona la prueba que esta por defecto establecida para esta muestra.

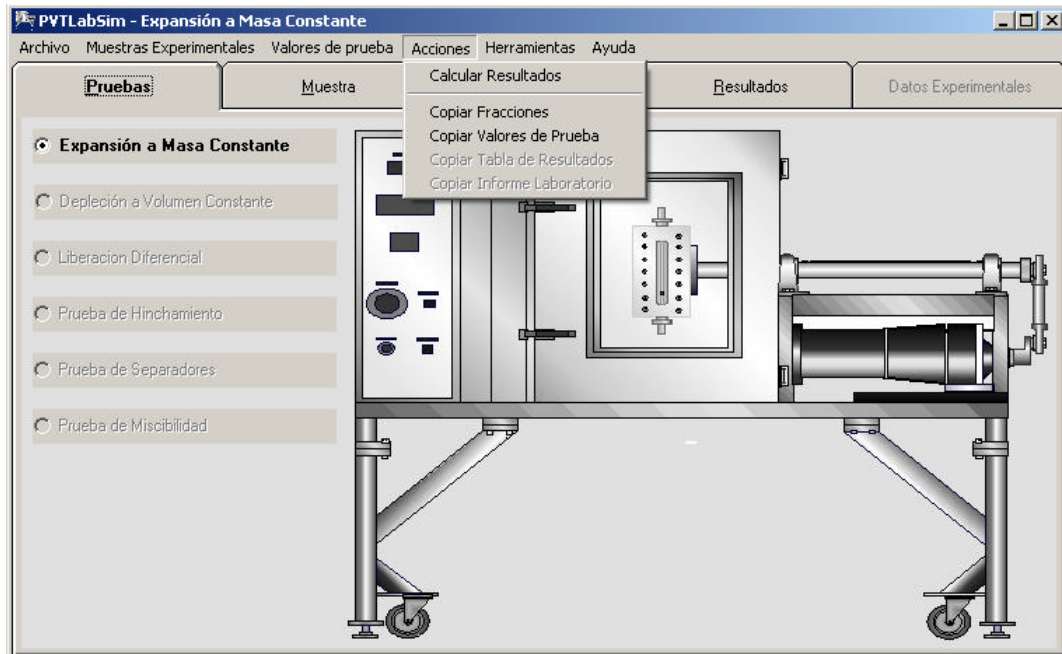
9.1.3. Barra de menú-Valores de Prueba.



La opción de *Borrar Valores de Prueba* elimina todos los valores de prueba en los formularios.

Con la opción *Usar Valores de Prueba por Defecto*. Además de llenar los formularios de entrada de datos de la muestra llena el formulario de datos de prueba con los datos de prueba que se utilizaron en la prueba seleccionada.

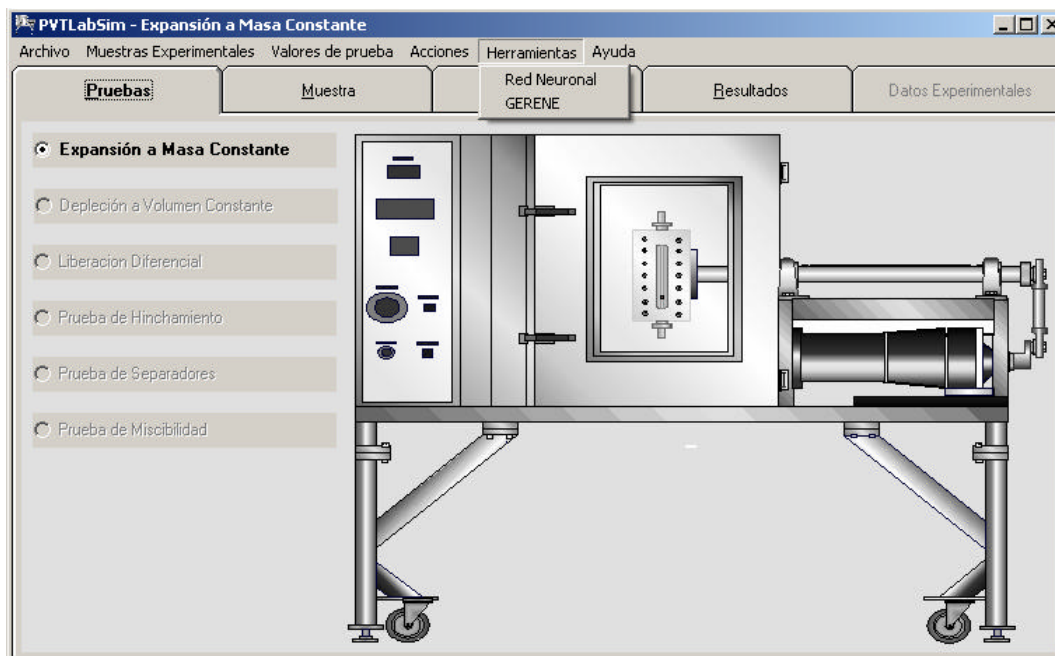
9.1.4. Barra de menú-Acciones.



Las acciones permiten:

- *Calcular Resultados.*
- Manipular los diferentes tipos de resultados y llevarlos a otros archivos utilizando el portapapeles de Windows (*Copiar Fracciones, Valores de Prueba, Tabla de Resultados, Informe de Laboratorio, Gráfica Actual*).

9.1.5. Barra de menú-Herramientas.



El software tiene asociado dos herramientas software que permiten trabajar redes neuronales: Red Neuronal (ANNBack) y GERENE.

La herramienta GERENE tiene su propio manual de manejo como un tutorial y puede ser consultado al ser ejecutado este software.

La herramienta ANNBack se divide en tres partes, que permiten cargar datos, entrenar la red y efectuar la predicción y evaluación de datos a partir de la estructura neural entrenada. A continuación se presenta una descripción breve de su funcionamiento para el usuario interesado en su utilización.

Entrada de datos y entrenamiento.

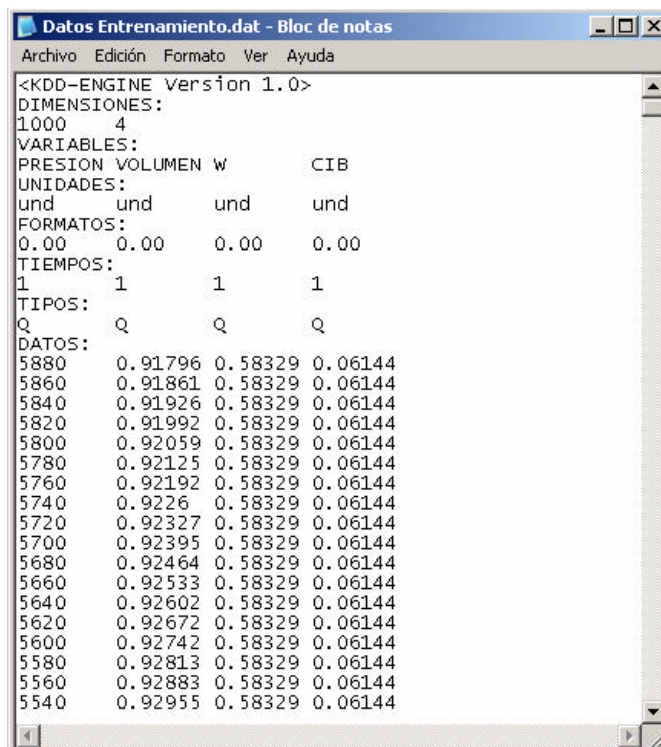
Los datos de entrada son la información de la red y los datos para su entrenamiento. La red utilizada es de tipo perceptrón multicapa y puede tener

cualquier cantidad de capas, las cuales se deben introducir como el número de neuronas por capa separadas por comas.

Ej. 5 capas: 3;4;2;2;3: 3 neuronas de entrada, tres capas ocultas con 4,2,2 neuronas respectivamente y 3 neuronas de salida.

Los datos de entrenamiento se deben introducir por medio de un archivo de datos con separación por tabulaciones como se muestra en la figura 45.

Figura 45. Archivo de entrada de datos para el software ANNBack.



Fuente: Autor

Como se puede observar, existen varios campos de entrada de datos:

- **DIMENSIONES:** Son respectivamente las dimensiones de los datos. Filas x Columnas.
- **VARIABLES:** El nombre de las variables de entrenamiento
- **UNIDADES:** Las unidades de las variables. Si las hay.
- **FORMATOS:** Son variables de control interno del software, hay que agregarlas sin modificarlas.
- **TIEMPOS:** Para una simulación dinámica se pueden modificar, en nuestro caso se agregan igual.
- **TIPOS:** El tipo de datos a entrenar. Q por defecto trabaja datos reales.
- **DATOS:** Los datos de entrenamiento separados por tabulaciones.

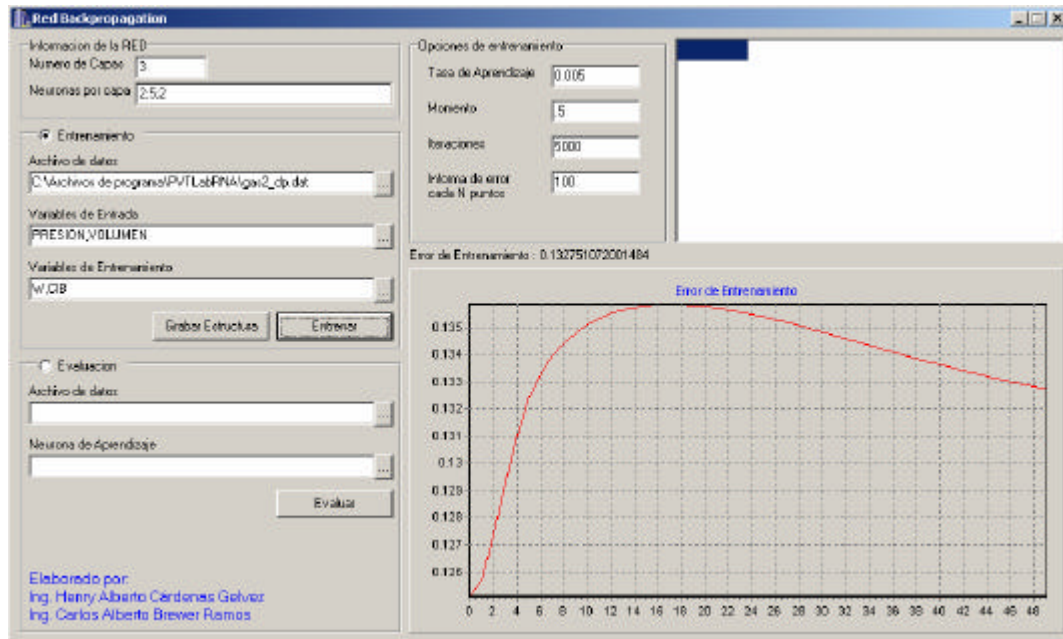
Entrenamiento de la red.

La red creada y con datos e puede entrenar al ver y modificar los parámetros como se muestra en la figura 46. Los campos que el usuario debe modificar son los siguientes:

- **VARIABLES DE ENTRADA:** El nombre de las variables de entrada (deben estar en el archivo de datos).
- **VARIABLES DE SALIDA:** El nombre de las respectivas variables de salida (guardadas en el archivo de datos).
- **TASA DE APRENDIZAJE:** Factor de entrenamiento que permite graduar la forma como la red va a aprender de los datos de entrada. Se recomienda comenzar a trabajar con 0.01 e irlo disminuyendo a la mitad.
- **MOMENTO:** Factor que permite suavizar la tendencia del error permisible. Este valor fluctúa entre 0 y 1. Se recomienda comenzar con un valor de 0.5 e irlo incrementando agregando cada vez la mitad de su valor actual.
- **ITERACIONES:** Son las iteraciones máximas que debe efectuar la red. Se recomienda un valor de 10.000 para evitar el sobreentrenamiento.

- **INFORME DE ERROR:** Este valor representa a que número de iteraciones se mostrará el valor en la curva generada (error permisible).

Figura 46. Entrenamiento de redes con el software ANNBack.

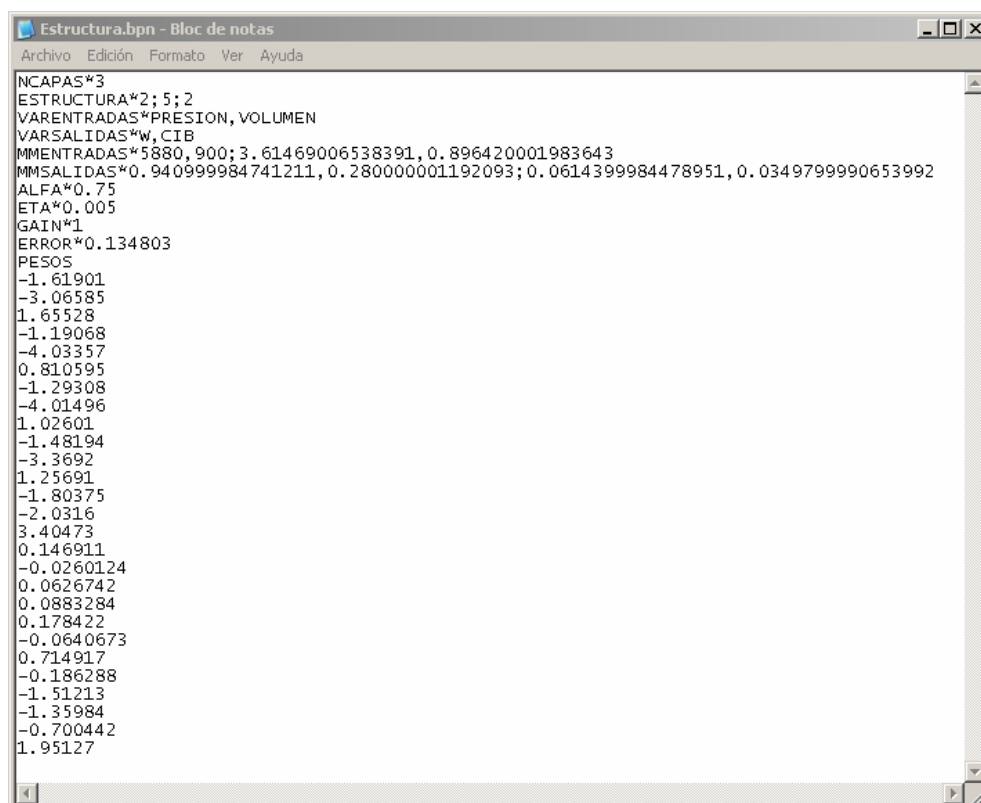


Fuente: Autor

El entrenamiento comienza cuando el usuario oprime el botón ENTRENAR. Esta etapa finaliza cuando el usuario modifica los parámetros de entrenamiento de tal manera que la curva de error permisible mostrada tenga un valor mínimo y se mantenga con tendencia decreciente a través de las iteraciones. Para grabar la estructura de la red al finalizar la sesión se debe oprimir el botón GRABAR ESTRUCTURA.

El archivo de la estructura entrenada y grabada tiene la forma mostrada en la figura 47.

Figura 47. Archivo de la estructura de la red entrenada por el software ANNBack.



```
Estructura.bpn - Bloc de notas
Archivo Edición Formato Ver Ayuda

NCAPAS*3
ESTRUCTURA*2;5;2
VARENTRADAS*PRESION,VOLUMEN
VARSAIDAS*W,CIB
MMENRADAS*5880,900;3.61469006538391,0.896420001983643
MMSALIDAS*0.940999984741211,0.280000001192093;0.0614399984478951,0.0349799990653992
ALFA*0.75
ETA*0.005
GAIN*1
ERROR*0.134803
PESOS
-1.61901
-3.06585
1.65528
-1.19068
-4.03357
0.810595
-1.29308
-4.01496
1.02601
-1.48194
-3.3692
1.25691
-1.80375
-2.0316
3.40473
0.146911
-0.0260124
0.0626742
0.0883284
0.178422
-0.0640673
0.714917
-0.186288
-1.51213
-1.35984
-0.700442
1.95127
```

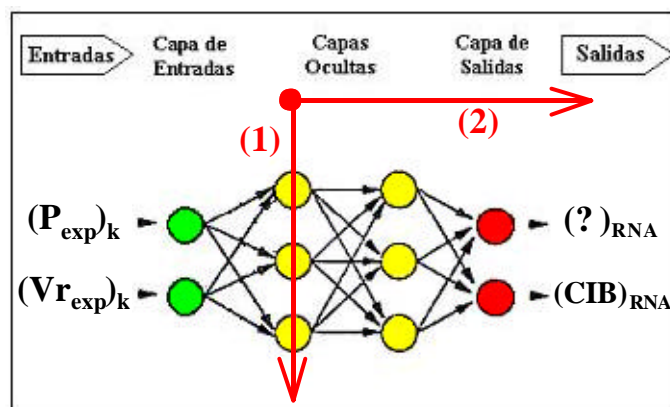
Fuente: Autor

Los campos presentes en el archivo de estructura son los siguientes:

- NCAPAS: Número de capas.
- ESTRUCTURA: Topología de la red.
- VARENTRADAS: Variables a entrenar.
- VARSAIDAS: Salidas de la red.
- MMENRADAS: Dato máximo y mínimo de cada variable de entrada para efectuar la normalización.
- MMSALIDAS: dato máximo y mínimo de cada variable de salida para efectuar la normalización.

- ALFA: Momento utilizado.
- ETA: Tasa de aprendizaje utilizada.
- GAIN: Indica la función de transferencia utilizada en la red (Sigmoidal=1).
- ERROR: Error permisible final.
- PESOS: Lista de los pesos finales de cada conexión entre neuronas. Antes de los pesos de cada neurona se coloca las *bias* obtenidas en la neurona. La forma de recorrer la red para guardar los pesos se puede ver en la figura 48.

Figura 48. Dirección del recorrido para guardar los pesos de cada conexión luego del entrenamiento.



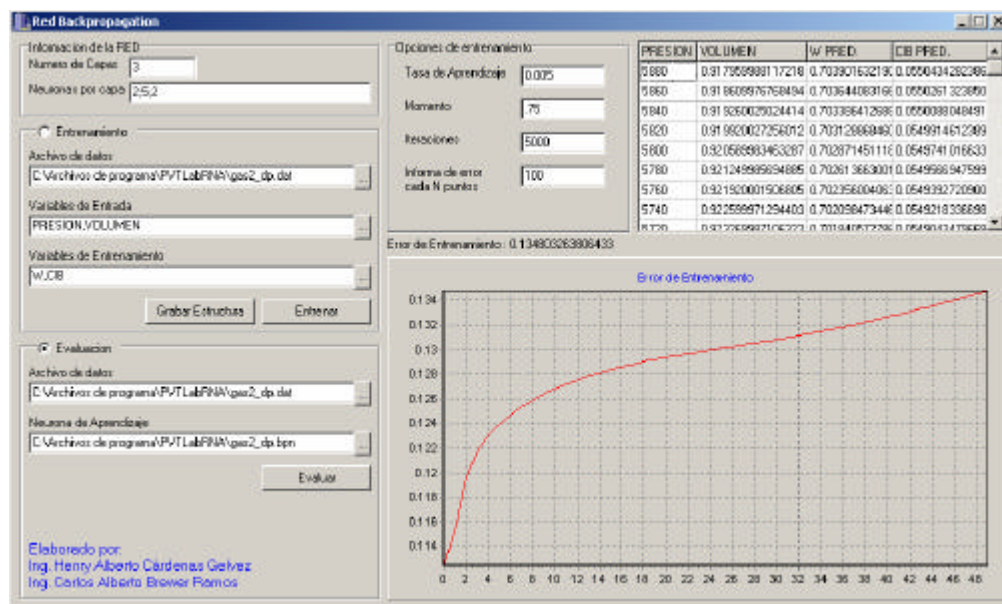
Fuente: Autor

Se puede observar que primero se recorre cada capa oculta en el sentido de la flecha (1) desde la primera (oculta) y se continúa de capa en capa hasta terminar en la capa de salida como lo indica la flecha (2).

Evaluación y datos de salida de la red.

Para efectuar la evaluación de la estructura que se entreno se utiliza la sección del software llamada EVALUACION. Ver figura 49.

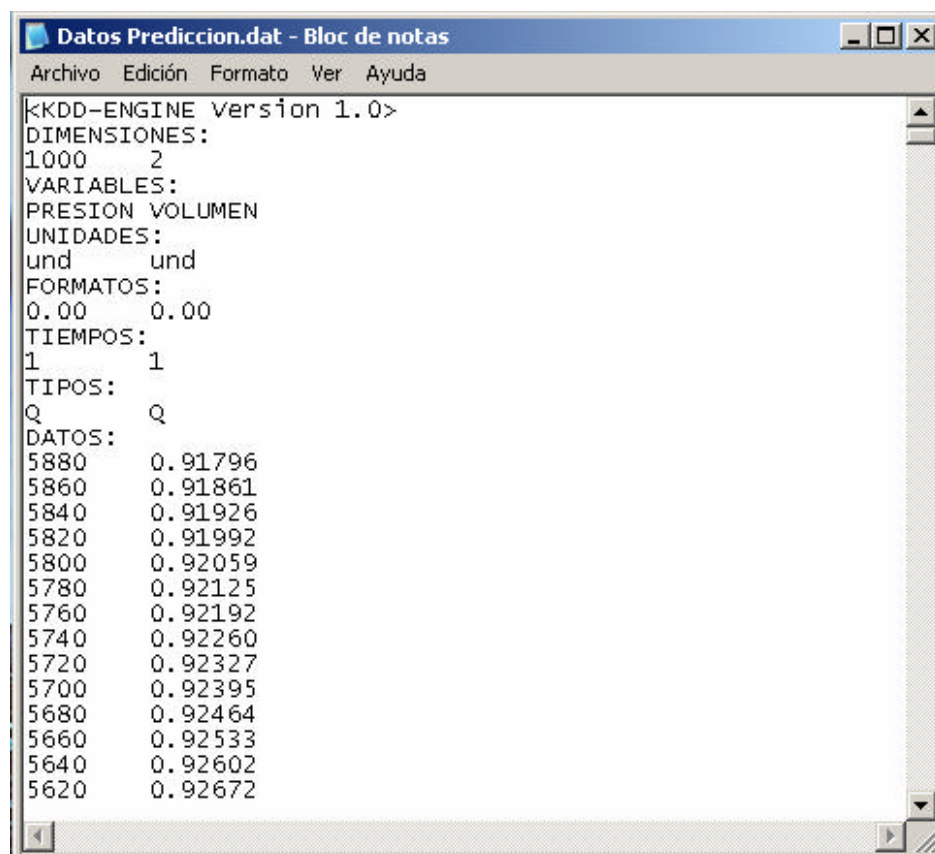
Figura 49. Evaluación de la red entrenada con el software ANNBack.



Fuente: Autor

Se introducen el nombre del archivo de datos de predicción (donde están los datos a los cuales se les va a encontrar las variables ajustadas por la red) y el nombre del archivo donde se encuentra grabada la estructura de la red. El archivo de datos para la evaluación se muestra en la figura 50.

Figura 50. Archivo de datos para la evaluación de la red entrenada por el software ANNBack.



```

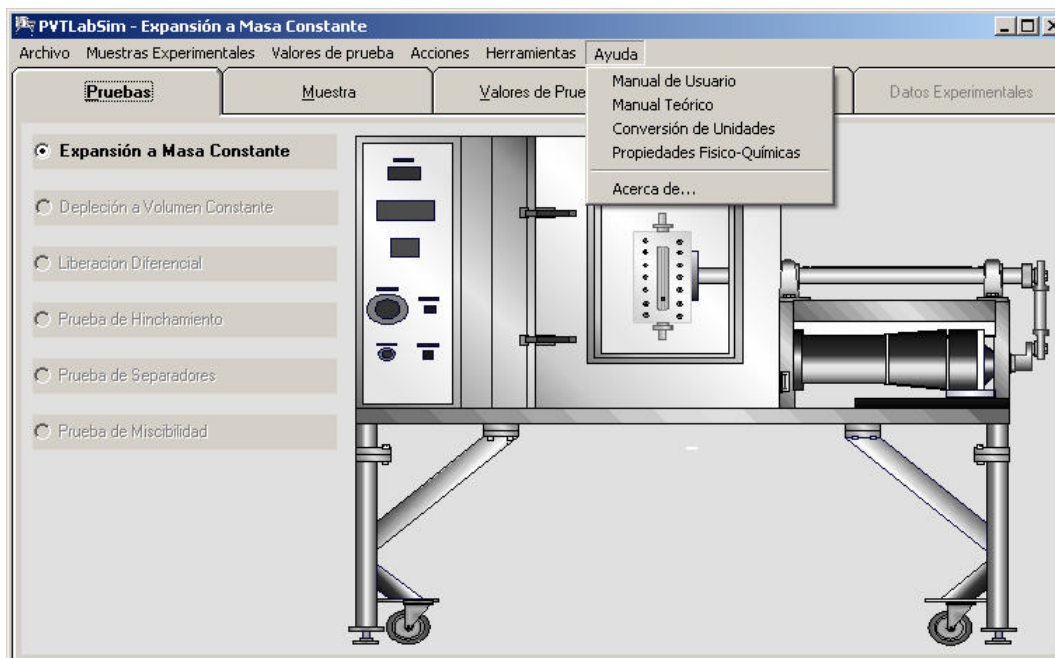
<KDD-ENGINE version 1.0>
DIMENSIONES:
1000  2
VARIABLES:
PRESION VOLUMEN
UNIDADES:
und    und
FORMATOS:
0.00   0.00
TIEMPOS:
1      1
TIPOS:
Q      Q
DATOS:
5880   0.91796
5860   0.91861
5840   0.91926
5820   0.91992
5800   0.92059
5780   0.92125
5760   0.92192
5740   0.92260
5720   0.92327
5700   0.92395
5680   0.92464
5660   0.92533
5640   0.92602
5620   0.92672

```

Fuente: Autor

Como se observa, este archivo conserva la estructura que tiene el archivo de datos para el entrenamiento, la diferencia solo radica en el número de variables y el listado de datos de cada una de ellas.

9.1.3. Barra de menú-Ayuda.

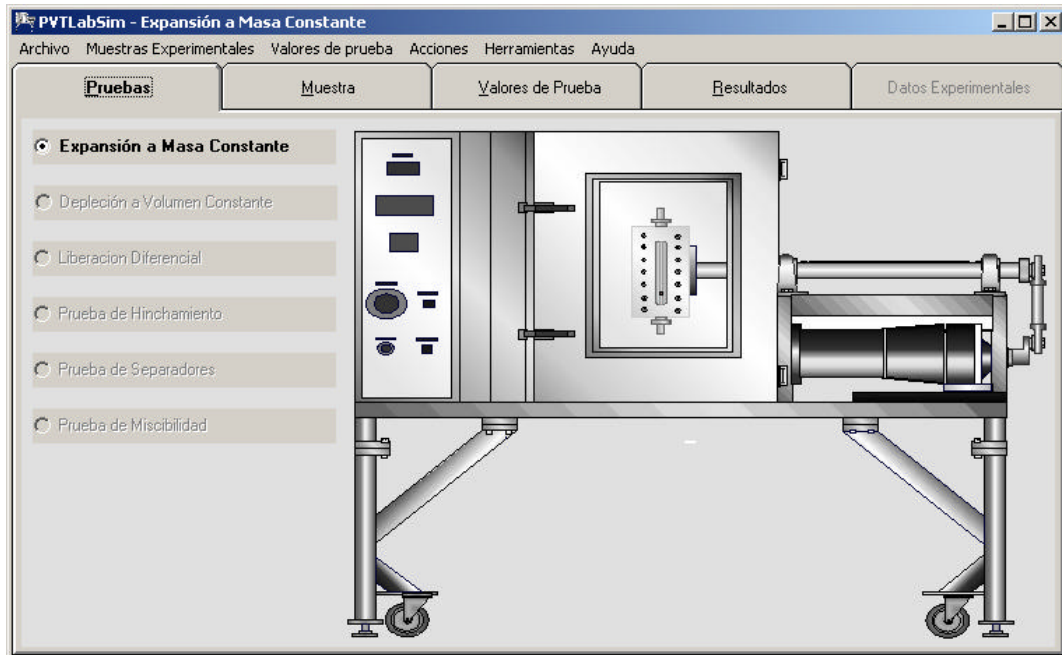


Contiene y permite visualizar los dos manuales de ayuda: el teórico y del usuario. También visualiza la propiedad intelectual y el soporte técnico del programa.

Se adiciona en este menú dos opciones muy útiles para el investigador de laboratorio PVT, un *Convertidor de Unidades* y la *Tabla de Propiedades Físico-Químicas* de los compuestos puros, estas aplicaciones le permitirán tener acceso a información básica instantáneamente.

9.2. CARPETAS PARA EL TRABAJO EXPERIMENTAL

9.2.1. Carpeta de Pruebas.



Permite seleccionar el tipo de prueba a efectuar:

- Expansión a Masa Constante.
- Depleción a Volumen Constante.
- Liberación Diferencial.
- Prueba de Hinchamiento.
- Prueba de Separadores.
- Prueba de Miscibilidad.

Para el ajuste a pruebas experimentales PVT solo se ha dejado activo el experimento de Expansión a Masa Constante, el cuál puede ser ajustado por medio de la metodología desarrollada en este proyecto.

9.2.2. Carpeta Muestra.

PVTLabSim - Expansión a Masa Constante

Archivo Muestras Experimentales Valores de prueba Acciones Herramientas Ayuda

Pruebas **Muestra** Valores de Prueba Resultados Datos Experimentales

Fracción Molar Inicial

Nombre	Fracción Molar
Nitrógeno	0
Dióxido de Carbono	0,0069
Sulfuro de Hidrógeno	0,0004
Metano	0,5832
Etano	0,1355
Propano	0,0761
Isobutano	0,02015
n-Butano	0,02015
Isopentano	0,01205
n-Pentano	0,01205
n-Hexano	0,019
Fracción Pesada C7+	0,1145
Total	1
Por asignar	0

El fluido inicial es un Gas Condensado

Peso Molecular C7+ 193

Gravedad Específica C7+ 0,8135

Temperatura Yacimiento 190 °F

COATS & SMART [SPERE(MAYO/86)] GAS#2*
* Muestra con Punto de Rocío.

☒ Introducir W y CIB de Cn

W(cn)

GEsp
CIB

En esta carpeta se manejan los datos de entrada del fluido inicial que se va a someter a la prueba PVT.

Los datos mínimos del fluido son:

- Fracciones molares de los componentes del fluido.
 - Tipo de fluido: Aceite Volátil o Gas Condensado.
 - Peso Molecular de C₇₊.
 - Gravedad Específica de C₇₊.
 - Temperatura del Yacimiento.
 - Factor acéntrico y Coeficiente de interacción Binaria del compuesto C_n.
- Esta opción permite modificar el valor de estos datos. El usuario puede introducir los valores que crea pueden mejorar las predicciones. Para el Coeficiente de Interacción Binaria se incluye la posibilidad de calcularlo por medio de la correlación de Katz-Firoozabadi ingresando solo la

gravedad Específica (GEsp) del compuesto Cn (útil si no tiene a la mano el valor puntual de este coeficiente).

Además se ofrece un espacio de texto donde se pueden consignar apuntes relevantes de la muestra.

9.2.3. Carpeta Valores de Prueba.

The screenshot shows the 'PVTLabSim - Expansión a Masa Constante' window. The 'Valores de Prueba' tab is selected, displaying a table with 21 rows and 2 columns: 'Presión [psig]'. The values range from 5000 down to 3200. To the right of the table are control options for sorting (No Ordenar, Ordenar Descendentemente, Ordenar Ascendentemente), a 'Reiniciar Valores de Prueba' button, input fields for 'Valor Inicial' (5000), 'Ancho Intervalo' (100), and 'Valor Límite' (3000), an 'Agregar a Presión [psig]' button, and a large 'Efectuar Cálculos' button.

	Presión [psig]
21	
1	5000
2	4900
3	4800
4	4700
5	4600
6	4500
7	4400
8	4300
9	4200
10	4100
11	4000
12	3900
13	3800
14	3700
15	3600
16	3500
17	3400
18	3300
19	3200

En esta carpeta se consignan los datos de las variables de laboratorio que se van a evaluar (Presiones).

Para utilizar la lista de valores es necesario dar click en la grilla en blanco y luego proceder a introducir el valor. La validación del dato se efectúa con la tecla *flecha abajo* (↓). Para insertar un nuevo espacio en blanco se debe oprimir la tecla *INSERT* y para eliminar se utiliza la tecla *SUPR*.

Se ofrece una opción para ordenar los datos de entrada:

- *No ordenar valores.*
- *Ordenar Descendentemente los valores.*
- *Ordenar Ascendentemente los valores.*

Se agregó una opción para el generar series de datos de entrada. Con esta opción se pueden generar series de datos de entrada teniendo en cuenta:

- *Valor Inicial:* Donde comienza la serie de datos.
- *Valor de cambio de la serie:* Este delta puede ser fijado en # de divisiones hasta el valor límite o un valor delta de cambio en la serie hasta el valor límite.
- *Valor Límite:* Es el último valor de la serie: Puede ser un valor puntual o un # de pasos definido.

Ejemplos:

Valor Inicial: 5000

Valor de cambio en la serie: Ancho intervalo: 500

Valor Limite: 3000

15	Presion [psig]
1	5000
2	4500
3	4000
4	3500
5	3000
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

☒ No Ordenar por valores de Presión
☐ Ordenar Descendentemente
☐ Ordenar Ascendentemente

Reiniciar Valores de Prueba

Valor Inicial: 5000
Ancho Intervalo: 500
Valor Límite: 3000

Agregar a Presion [psig]

Efectuar Calculos

Valor Inicial: 5000

Valor de cambio en la serie: Divisiones: 10

Valor Limite: 3000

The screenshot shows the 'PVTLabSim - Expansión a Masa Constante' software interface. The 'Valores de Prueba' tab is active. On the left, a table lists pressure values in psig for 15 divisions, starting at 5000 and decreasing by 200 until 3000. On the right, configuration options include sorting (No Ordenar, Descendente, Ascendente), a 'Reiniciar Valores de Prueba' button, and input fields for 'Valor Inicial' (5000), 'Divisiones' (10), and 'Valor Limite' (3000). A 'Agregar a Presion [psig]' button and an 'Efectuar Calculos' button are also present.

	Presion [psig]
15	
1	5000
2	4800
3	4600
4	4400
5	4200
6	4000
7	3800
8	3600
9	3400
10	3200
11	3000
12	
13	
14	
15	

Valor Inicial: 5000

Valor de cambio en la serie: Ancho intervalo: 500

Valor Límite: Número de Pasos: 6

The screenshot shows the 'PVTLabSim - Expansión a Masa Constante' software interface. The 'Valores de Prueba' tab is active. On the left, a table lists pressure values in psig for 15 divisions, starting at 5000 and decreasing by 500 until 2000. On the right, configuration options include sorting (No Ordenar, Descendente, Ascendente), a 'Reiniciar Valores de Prueba' button, and input fields for 'Valor Inicial' (5000), 'Ancho Intervalo' (500), and 'Número de Pasos' (6). A 'Agregar a Presion [psig]' button and an 'Efectuar Calculos' button are also present.

	Presion [psig]
15	
1	5000
2	4500
3	4000
4	3500
5	3000
6	2500
7	2000
8	
9	
10	
11	
12	
13	
14	
15	

9.2.4. Carpeta Resultados.

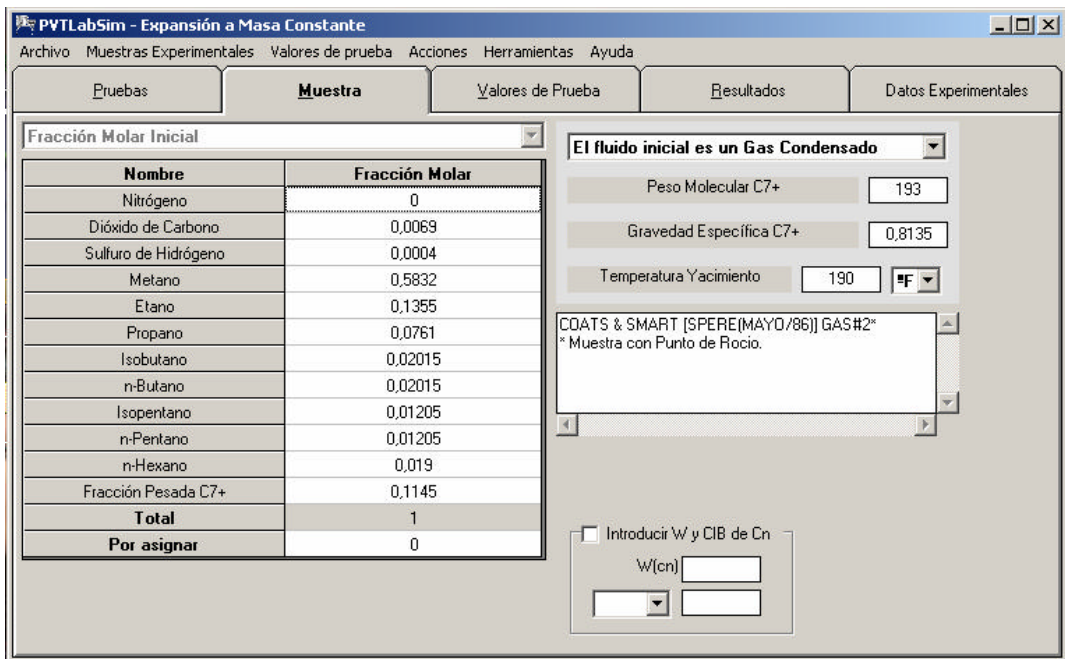
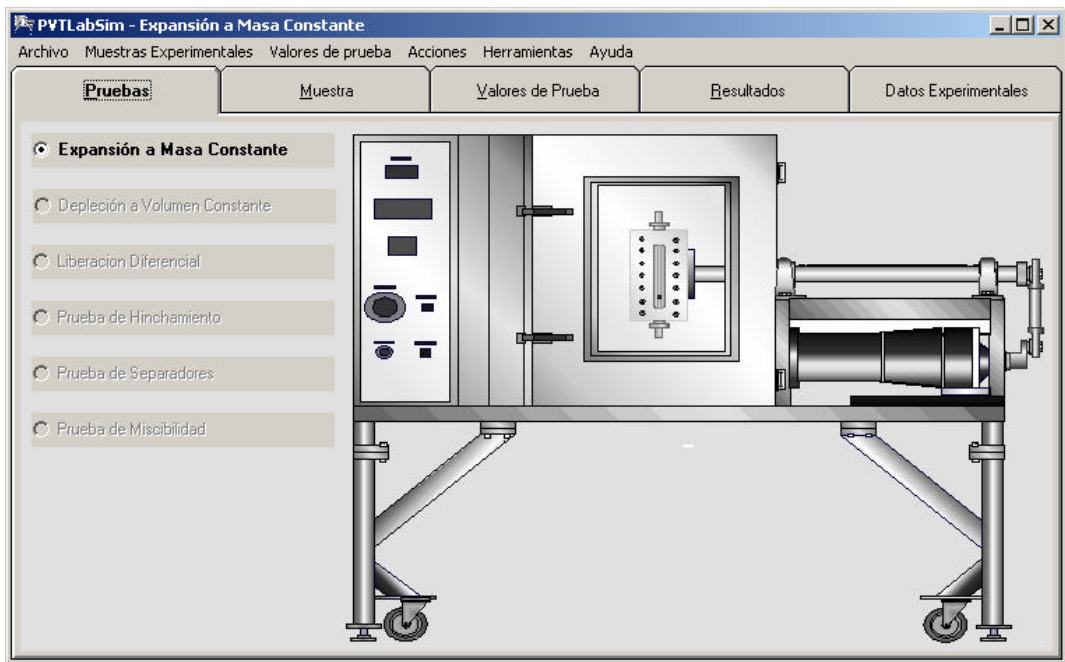
Presión [psig]	Volumen Relativo Vr = [V/Vs]	Fracción de Líquido [%]	Factor de Compresibilidad Zg
5000,0	0,9505	0,0	1,0668
4000,0	1,0012	0,0	0,8994
3000,0	1,2063	43,7009	0,8068
2000,0	1,6273	25,9158	0,8009
1500,0	2,1244	17,3654	0,8213
1000,0	3,2285	9,4837	0,8595
900,0	3,6147	8,0613	0,8694
800,0	4,1055	6,7097	0,88
500,0	6,846	3,1814	0,9168
250,0	14,4352	1,0291	0,9538

En esta carpeta se visualizan los datos de salida del simulador. Se presentan sub-carpetas con la información:

- *Tabla de Datos.* Se Presentan las tablas de los resultados para cada tipo de prueba.
- *Informe de Laboratorio.* Se muestra un informe tipo de laboratorio con los cálculos intermedios más importantes de cada experimento.
- *Reporte Gráfico.* Muestra las principales curvas que se efectúan con los resultados de cada prueba.

9.3. EXPERIMENTO DE EXPANSION A MASA CONSTANTE

Este experimento se desarrolla de la siguiente manera:



PVTLabSim - Expansión a Masa Constante

Archivo Muestras Experimentales Valores de prueba Acciones Herramientas Ayuda

Pruebas Muestra **Valores de Prueba** Resultados Datos Experimentales

	Presión (psig)
11	
1	5000
2	4000
3	3000
4	2000
5	1500
6	1000
7	900
8	800
9	500
10	250
11	

☒ No Ordenar por valores de Presión
☐ Ordenar Descendentemente
☐ Ordenar Ascendentemente

Reiniciar Valores de Prueba

Valor Inicial: 5000
 Ancho Intervalo: 100
 Valor Límite: 3000

Agregar

Efectuar Cálculos

PVTLabSim - Expansión a Masa Constante

Archivo Muestras Experimentales Valores de prueba Acciones Herramientas Ayuda

Pruebas Muestra Valores de Prueba **Resultados** Datos Experimentales

Ps = 3893.627 [psia]

Presión [psig]	Volumen Relativo Vr = [V/Vs]	Fracción de Líquido [%]	Factor de Compresibilidad Zg
5000.0	0.9505	0.0	1.0668
4000.0	1.0012	0.0	0.8994
3000.0	1.2063	43.7009	0.8068
2000.0	1.6273	25.9158	0.8009
1500.0	2.1244	17.3654	0.8213
1000.0	3.2285	9.4837	0.8595
900.0	3.6147	8.0613	0.8694
800.0	4.1055	6.7097	0.88
500.0	6.846	3.1814	0.9168
250.0	14.4352	1.0291	0.9538

Tabla de Datos Informe Laboratorio Reporte Gráfico

CONCLUSIONES

Se desarrolló una nueva metodología conjunta entre las redes neuronales y la metodología utilizada en el software de simulación de pruebas PVT - PVTLABSIM -, para el ajuste de los Parámetros de la Ecuación de Estado de Peng-Robinson.

La metodología desarrollada permite mejorar de manera Meta-Heurística la selección de los parámetros de ajuste en la ecuación de estado.

Se aplicaron las redes neuronales al experimento de expansión constante, permitiendo simular por medio de la ecuación de estado las propiedades PVT de la muestra: Volumen relativo y presión de saturación.

Se verificó la gran influencia que tienen los parámetros de interacción binaria y el factor acéntrico de la fracción pesada en el comportamiento composicional de una muestra de gas condensado y/o aceite volátil.

La nueva metodología habilita a la ecuación de estado de Peng-Robinson para efectuar cálculos volumétricos en la región bifásica, donde es bien sabido que esta ecuación presenta fallas debido a su baja definición de la fracción líquida.

Los parámetros encontrados (Factor Acéntrico y Coeficiente de Interacción Binaria) por medio de la red neuronal permiten esclarecer hasta que familias de carbonos están presentes en la fracción pesada, por ello estos valores servirán de partida en un proceso posterior de caracterización y extensión de la fracción pesada.

La nueva metodología se implementó dentro de una aplicación software llamada PVTLabRNA. Los resultados que se derivan de ésta son la disminución del error relativo total y una mejor aproximación al valor experimental de la presión de saturación en comparación con los resultados obtenidos con PVTLAB.

El software desarrollado es eficiente y abierto al usuario. Permitiendo así que los resultados dependan de los datos que el experto utilice según su criterio.

El uso de la inteligencia artificial como complemento de la simulación de procesos hidrocarburos complejos ayuda a un mejor entendimiento de los mismos. En el caso del ajuste de la data PVT permite visualizar lo que pasa con propiedades físicas de difícil determinación como el factor acéntrico y el coeficiente de interacción binaria, necesarios para la caracterización de fluidos petrolíferos.

Se encontró que el ajuste de los parámetros utilizando algoritmos genéticos es más preciso que al ser efectuado por redes neuronales, pero el trabajo con redes permite al usuario utilizar valores que por su experiencia son más representativos en el proceso. Además el tiempo de cálculo de los algoritmos genéticos es mucho mayor que la interpolación de valores que efectúa la red neuronal.

RECOMENDACIONES

Se recomienda extender esta metodología a otras pruebas PVT para verificar la consistencia de los parámetros predichos por la red neuronal.

Es necesario que el usuario establezca los límites físicos que deben tener los parámetros a ser ajustados, para que la interpolación efectuada por la red tenga este sentido físico y los valores predichos sean los de moléculas presentes en el fluido.

Se recomienda utilizar otro tipo de tipologías neuronales que permitan establecer más relaciones entre los parámetros a ajustar y por ello mejores predicciones de los mismos.

Se recomienda depurar los datos de entrada a la red para evitar un sobre-entrenamiento de la misma con sus consecuencias: Bajo-predicción y Sobre-predicción.

Se recomienda tener en cuenta el número de cifras significativas con las que se entrena la red neuronal (4 cifras significativas después del decimal) para establecer claramente el rango del error permisible.

Se encontró que el uso de diferentes metodologías (tablas, correlaciones) de los parámetros físicos a ser ajustados afecta la predicción de la red y su capacidad de encontrar un valor final acertado. Por ello se recomienda efectuar un estudio comparativo que permita observar el ajuste que se logra al utilizar cada metodología.

REFERENCIAS BIBLIOGRAFICAS

1. EYTON, D. G. P. "Practical Limitations in Obtaining PVT Data for Gas Condensate Systems". SPE 15765.
2. HOFFMANN, A. E., CRUMP, J. S. and HOCOTT, C. R. "Equilibrium Constants for a Gas Condensate System". Trans. AIME 1953.
3. CORE LABORATORIES. "Basic Phase Behavior Course". 1999.
4. SUNBURY B. P. RESEARCH CENTER. Conferencias sobre Comportamiento de Fases. Santafé de Bogotá. 1996.
5. AHMED, T. H. "Hydrocarbons Phase Behavior" Gulf Publishing Company, USA. 1989.
6. PEDERSEN, K. S., FREDENSLUND, A. and THOMASSEN, P., "Properties of Oil and Natural Gases". Gulf Publishing Company. 1989.
7. VAROTSIS, N. "A Robust Prediction Method for Rapid Phase Behavior Calculations". SPE 16943 Mayo 1989.
8. FUSELL, D. D. and YANOSIK, J. L. "An Iterative Sequence for Phase Equilibria Calculations Incorporating the Redlich-Kwong Equation of State". SPEJ May. 1989.
9. NGHIEM, L. X., AZIZ, K., LI, Y. K., "A Robust Iterative Method for Flash Calculations Using the Soave-Redlich-Kwong or the Peng-Robinson Equation of State". SPE 8285.

10. PENG, D. Y. and ROBINSON, D. B. , "A new Two-Constant Equation of State". Ind. Eng. Chem. Fund. Vol. 15 No. 1, 1976.
11. SLOT-PETERSEN, C., "A Systematic and Consistent Approach to determine Binary Interaction Coefficients for the Peng-Robinson Equation of State". SPE 16941. 1987.
12. VAROTSIS, N. , STEWART, G. , TODD, A. C. and CLANCY, M. "Phase Behavior of Systems Comprising North Sea Reservoir Fluids and Injection Gases". JPT, Nov. 1986.
13. CARROLL, J. J. , "A Generalized Correlation for the Peng-Robinson Interaction Coefficients for Paraffin-Hydrogen Sulfide Binary Systems" Fluid Phase Equilibria, 105 (1995).
14. FIROOZABADI, A. and ARBABI, S., "Near-Critical Phase Behavior of Reservoir Fluid using EOS". ADNOC / SPE 24491.
15. AHMED, T. H. "A Practical Modification of the Peng-Robinson Equation of State". SPE 18532. Nov. 1988.
16. JHAVERI, B. S., " Three - Parameter Modification of the Peng-Robinson Equation of State to Improve Volumetric Predictions". SPE 13118.
17. VAROTSIS, N. "A Robust Prediction Method for Rapid Phase Behavior Calculations". SPE 16943 Mayo 1989.
18. MARTINEZ, M. y PEREZ, R. "Ingeniería del Gas Natural: Características y Comportamiento de los Hidrocarburos". Maracaibo - Venezuela. 1994.
19. RACHFORD, H. H. "Procedure for use of Electronic Digital Computers in Calculating Flash Vaporization Hydrocarbon Equilibrium". Trans. AIME. Vol. 195. 1952.

20. WHITSON, C. H., "Characterizing Hydrocarbon Plus Fractions". SPEJ. Aug. 1983.
21. AHMED T.H. "A Generalized Correlation for Characterizing the Hydrocarbon Heavy Fractions". SPE 14266, 1985.
22. RIAZI, M. R. and DAUBERT, T. E. "Characterization Parameters for Petroleum Fractions". Ind. & Eng. Chem. Res. 1987. Vol. 26, No.4 Pag. 755.
23. PEDERSEN, K. S., FREDENSLUND, A. and THOMASSEN, P., "Properties of Oil and Natural Gases". Gulf Publishing Company. 1989.
24. AHMED T.H. "A Generalized Correlation for Characterizing the Hydrocarbon Heavy Fractions". SPE 14266, 1985.
25. COATS, K.H. and SMART, G. T. "Application of a Regression Bases EOS PVT Program to Laboratory Data ". SPERE, May. 1980.
26. AGARWAL, RAJEEV K., LI, YAU-KUN, NGHIEM, LONG, "A regression technique with dynamic parameter selection for phase-behavior matching". SPE Reservoir Engineering. (Feb. 1990). p.115 – 120.
27. DA SILVA M. B., RODRIGUEZ, F., "Automatic fitting of equation of state for phase behavior matching". SPE 23703. (1992). p.1 – 12.
28. TRUJILLO, C. W. "Ajuste automático de la ecuación de estado de Peng-Robinson a mezclas de hidrocarburos". Tesis de Maestría UNAM. México (1996).
29. SAN MIGUEL, Luis Alberto. "Métodos alternos para el ajuste automático de las ecuaciones de estado". Tesis de Maestría UNAM. México (1998).

30. MERRILL, R. C., HARTMAN, K. J., CREEK, J. L. "A comparison of equation of state tuning methods". SPE 28598. (1994). p.5.

31. KATZ, D. L., FIROOZABADI, A. "Predicting phase behavior or condensate/crude-oil systems using methane interactions coefficients". Journal of Petroleum Technology. (Aug. 1987). p.185–191.

32. TWU, CHORNG H., COON, JOHN E. Y CUNNINGHAM, JOHN R. "A new generalized alpha function for a cubic equation of state. Part 1. Peng-Robinson equation". Fluid Phase Equilibria. Vol 105. (1995). p.49-59.

33. ALONSO, G. and BECERRIL, J.L. "Introducción a la inteligencia artificial". Ed. Multimedia Ediciones S.A. Barcelona. 1993.

34. FREEMAN J. A., SKAPURA D. M. "Redes neuronales algoritmos, aplicaciones y técnicas de programación". Addison-Wesley/ Díaz de Santos. 1993.

35. WERBOS P. J. "Backpropagation past and future". IEEE-ICNN. 1988. p.343-353.

36. WERBOS P.J. "Backpropagation and neurocontrol: A review and prospectus". IEEE-ICNN. 1989. p.209-216.

37. INSTITUTO MEXICANO DEL PETRÓLEO. "Uso de Inteligencia Artificial en la Inversión de Registros Eléctricos y Electromagnéticos". Revista IMP. (May/99). P.78-95.

38. CASTIBLANCO GALVIS, IOWANY Y NIZ VELÁSQUEZ EIDER. "Ajuste de la Ecuación de Estado de Peng-Robinson a Datos PVT de Yacimientos Cercanos al Punto Crítico Mediante Regresión Múltiple". Proyecto de Grado UIS. Bucaramanga (2000).

39. CHRISTENSEN J. "*Regression of PVT Data*". J. Can. Pet. Tech. Vol. 38(1999), p.57.

40. PEDERSEN K.S.; THOMASSEN P. AND FREDENSLUND A. "On the Dangers of Tuning Equations of State Parameters". Chem. Eng. Sci. (1988). p.269-278.

41. HILERA J.R. Y MARTÍNEZ V.J. "Redes Neuronales Artificiales: Fundamentos, Modelos y Aplicaciones". Addison-Wesley Iberoamericana. NY. USA. 1995.

42. KAY JIM W. AND TITTERINGTON D. M. "Statistics and Neural Networks: Advances at the Interface". Oxford University Press. London. UK. 1999.

43. SARLE W. "Neural Networks and Statistical Methods". Proceedings of the Nineteenth Annual S.A.S. Users Group International Conference. NY. USA. 1994.

44. MARTÍN DEL BRÍO B. Y SANZ MOLINA A. "Redes Neuronales y Sistemas Borrosos". Ra-Ma Textos Universitarios. Madrid. España. 1998.

45. CALVETE G., FERNANDO E. y RANGEL C., NOHORA P. "Simulación de Experimentos PVT para Yacimientos de Gas Condensado y Aceite Volátil". Proyecto de grado UIS. Bucaramanga (1997).

46. CALVETE G., FERNANDO E. y RANGEL C., NOHORA P. "Simulación de Experimentos PVT para Yacimientos de Gas Condensado y Aceite Volátil". VIII Congreso Colombiano del Petróleo. Bogotá. Octubre 1999.

47. PEREZ I., MIGUEL F.; ROZO A., MIKE A. y ULLOA J., ROBERT. "Estudio de Yacimientos de Petróleo usando Redes Neuronales Artificiales". Proyecto de grado UIS. Bucaramanga (2002).

48. PEREZ I., MIGUEL F.; ROZO A., MIKE A. y ULLOA J., ROBERT. "Herramienta Software para Entrenar Redes Neuronales: Aplicación en Registros de Pozos". X Congreso Colombiano del Petróleo. Bogotá. Octubre 2003.

49. CALVETE G., FERNANDO E. y ROMAN V., SILVIA P. "Computación Evolutiva: Algoritmos Genéticos y Simulated Annealing: Técnicas Meta-Heurísticas de Optimización en Ingeniería de Petróleos". IX Congreso Colombiano del Petróleo. Bogotá. Octubre 2001.
50. CALVETE G., FERNANDO E. y CALDERON C., ZULY H. "Ajuste de la Ecuación de Peng-Robinson a Datos PVT Utilizando Algoritmos Genéticos". X Congreso Colombiano del Petróleo. Bogotá. Octubre 2003.
51. LABORATORIOS CORE, Reporte 88002.
52. WHITSON, C. H. and TORP, S. B. "Evaluating Constant Volume Depletion Data". JPT(March/83).

ANEXO A. ECUACIONES Y PROCEDIMIENTOS NUMÉRICOS UTILIZADOS EN EL SOFTWARE PvtLabRNA

PROCEDIMIENTO	ECUACIONES
Cálculos Volumétricos y Relaciones PVT	<p>Ecuación de estado de Peng-Robinson:</p> $P = \frac{RT}{v-b} - \frac{a(T)}{v(v+b)+b(v-b)} \quad \text{ó} \quad Z^3 - (1-B)Z^2 + (A-3B^2-2B)Z - (AB-B^2-B^3) = 0$ $a(T) = a(T_c) \alpha(T_r, w) \quad A = \frac{aP}{R^2 T^2} \quad \alpha = (1 + m(1 - T_r^{0.5}))^2$ $b(T) = b(T_c)$ $b(T_c) = 0.07780 \frac{RT_c}{P_c} \quad B = \frac{bP}{RT} \quad m = 0.379642 + 1.48503w - 0.1644w^2 + 0.016667w^3$
Cálculo del Equilibrio de Fases	<p>Para el equilibrio: $f_i^V = f_i^L$</p> $\ln\left(\frac{f}{p}\right) = (Z-1) - \ln(Z-B) - \frac{A}{2\sqrt{2}B} \ln\left(\frac{Z+2.414B}{Z-0.414B}\right)$ $\Phi_i^V = \frac{f_i^V}{y_i p} \quad \Phi_i^L = \frac{f_i^L}{x_i p}$ $\ln \Phi_i^L = \frac{b_i}{b} (Z^L - 1) - \ln(Z^L - B) - \frac{A}{2\sqrt{2}B} \left[\frac{2y_i}{y} - \frac{b_i}{b} \right] \ln \left[\frac{Z^L + 2.414B}{Z^L - 0.414B} \right]$ $y_i = \sum_j [x_j (a_i a_j a_j)^{0.5} (1 - d_{ij})] \quad y = \sum_j [x_j (a_i a_j a_j)^{0.5} (1 - d_{ij})]$ $a = \sum_i \sum_j x_i x_j a_{ij} \quad b = \sum_i x_i b_i$ $a_{ij} = (1 - d_{ij}) a_i^{1/2} a_j^{1/2}$
Cálculos Flash	<p>1. Balance total de materia: $F = L + V$ 2. Balance de materia por componente: $Z_i F = X_i (1-\beta) + Y_i \beta$ Donde $b = \frac{V}{F}$</p> <p>3. Equilibrio termodinámico: $f_i^L = f_i^V$ Definiendo la relación de equilibrio como: $K_i = \frac{Y_i}{X_i} = \frac{f_i^L}{f_i^V}$</p> <p>Reduciendo la expresión: $Y_i = \frac{K_i \times Z_i}{1 + b \times (K_i - 1)}$ y $X_i = \frac{Z_i}{1 + b \times (K_i - 1)}$</p> <p>Aplicando el balance de materia en cada fase: $\sum_{i=1}^{N_c} Z_i = \sum_{i=1}^{N_c} X_i = \sum_{i=1}^{N_c} Y_i = 1$</p> <p>De tal manera: $\sum_{i=1}^{N_c} X_i = \sum_{i=1}^{N_c} \frac{Z_i}{1 + b \times (K_i - 1)} = 1$ Y por lo mismo: $\sum_{i=1}^{N_c} Y_i = \sum_{i=1}^{N_c} \frac{K_i \times Z_i}{1 + b \times (K_i - 1)} = 1$</p>

ANEXO B. AJUSTE DE LA ECUACIÓN DE PENG-ROBINSON A DATOS PVT UTILIZANDO ALGORITMOS GENÉTICOS.



ACIPET

Ajuste de la Ecuación de Peng-Robinson a Datos PVT Utilizando Algoritmos Genéticos.

F. E. Calvete G.*; Z. H. Calderón C.**

* Ingeniero de Petróleos, ECOPETROL, Casabe. fernando.calvete@ecopetrol.com.co

** Docente, Escuela de Ingeniería de Petróleos, Universidad Industrial de Santander. calderon@uis.edu.co

Derechos de Autor 2003, ACIPET

Este artículo técnico fue preparado para presentación en el X Congreso Colombiano del Petróleo organizado por ACIPET en Bogotá D.C. Colombia, 14- 17 de Octubre de 2003.

Este artículo fue seleccionado para presentación por el comité técnico de ACIPET, basado en información contenida en un resumen enviado por el autor(es).

Resumen

Los yacimientos de fluidos composicionales además de ser muy valiosos presentan una gran complejidad e inestabilidad. En ingeniería de petróleos, se han utilizado ecuaciones de estado para predecir y estimar el comportamiento de las fases presentes en ellos, constituyéndose en la base fundamental de la simulación composicional. Sin embargo, su utilización es limitada porque pierden validez en diferentes rangos de temperatura y presión.

Por otro lado, se han utilizado algoritmos de análisis estadístico, basados en el método de mínimos cuadrados, con el fin de ajustar los datos

predictivos de las ecuaciones de estado a los datos experimentales PVT y de esta manera obtener los parámetros que permitan utilizarla a diferentes condiciones de temperatura y presión. Estas metodologías, tiene gran aplicación y una alta precisión, pero ofrecen una gran desventaja: el ajuste estadístico fuerza a cambiar los parámetros de la ecuación de estado sin tener en cuenta que estos reflejan características físico-químicas de los compuestos que hacen parte de la mezcla de hidrocarburos.

El presente trabajo tiene como fin, presentar una metodología, basada en algoritmos genéticos, que permite generar los parámetros de la ecuación de estado que reproducen con exactitud los datos PVT de laboratorio; permitiendo posteriormente predecir el comportamiento de fases de los fluidos a condiciones de yacimiento.

Esta nueva metodología ajusta los Coeficientes de Interacción Binaria, los cuales son parte fundamental de la

ecuación de estado y reflejan las interrelaciones moleculares de los compuestos químicos de los fluidos. Los algoritmos genéticos que se utilizan no permiten que los cambios de los parámetros sean forzados sino que lleguen a valores tales que continúen representando las características físicas de los compuestos y permitan predicciones del orden del 99% de precisión como se muestra en los respectivos análisis de resultados.

Introducción

En simulación composicional es requisito fundamental tener una herramienta numérica que permita predecir el comportamiento termodinámico a condiciones de yacimiento. Esta herramienta es la ecuación de estado. La data PVT es utilizada para poder configurar y calibrar las ecuaciones de estado utilizadas en los simuladores composicionales. La evaluación matemática del desarrollo de un experimento PVT se dificulta, en ocasiones debido a la inestabilidad de fases en el equilibrio y a la pérdida de validez de las ecuaciones de balance de energía en zonas de alta presión.

Uno de los problemas que hacen más sensibles los resultados de la ecuación de estado es la descripción adecuada de la fracción pesada. Para suplir esta deficiencia se ha venido trabajando en métodos que permitan la caracterización apropiada de las propiedades de la misma para su posterior utilización en cálculos del equilibrio termodinámico. Las principales desventajas de estos procedimientos son su complejidad y la gran cantidad de tiempo de cómputo

Básicamente la herramienta que se ha utilizado tradicionalmente en las diversas metodologías de ajuste es la *regresión de mínimos cuadrados* de los parámetros de la Ecuación de Estado (EdE)¹. Estas metodologías ajustan parámetros por separado a una función de error, producto de la diferencia entre los datos obtenidos experimentalmente con respecto a los arrojados por la ecuación de estado, sin

embargo en algunas metodologías^{2,3} no se tiene en cuenta que los parámetros de ajuste representan características reales de los compuestos que conforman los fluidos de los yacimientos y al variarlos la EdE pierde representatividad aún cuando reproduce fidedignamente los datos PVT experimentales. Esta falta de representatividad del fluido puede causar una mala evaluación de los fenómenos composicionales que puedan estarse presentando en el yacimiento.^{4,5}

En términos matemáticos el problema del ajuste de los parámetros de la ecuación de estado a datos experimentales PVT es una *regresión multi-paramétrica y multi-objetivo*, ya que se pretende ajustar los valores de varios parámetros de la ecuación de estado de manera que permitan reproducir todos los datos experimentales y a su vez mantener los valores de estos parámetros dentro de rangos que les permitan seguir representando las propiedades físico-químicas de las especies químicas que componen los fluidos.

Aunque existen diversas metodologías que permiten realizar el ajuste de las ecuaciones de estado^{6,7,8}, algunas de ellas de manera automática⁹, otras basados en un fuerte trabajo de regresión con respecto a los datos experimentales de laboratorio¹⁰, se presentan grandes restricciones y problemas debido a la heurística y las condiciones matemáticas de las cuales fueron derivadas –principio de los estados correspondientes–¹¹. Estas metodologías permiten efectuar un ajuste estadístico que minimiza el error entre los datos experimentales y los predichos por la ecuación de estado, modificando los parámetros de la misma sin tener en cuenta los límites de estas variables, es decir sin respetar los valores de las propiedades termodinámicas de las especies químicas que están presentes en los fluidos de los yacimientos hidrocarburíferos. Al efectuar un ajuste con esta visión netamente estadística se logra un mínimo error pero la modificación de los parámetros de la ecuación le resta representatividad, ya que estos son los

valores de las propiedades naturales de cada compuesto químico.

Ecuación de Estado

Una Ecuación de Estado (EdE) es la relación entre presión, volumen y temperatura, la cual permite estudiar el comportamiento físico de una sustancia con respecto a las variables mencionadas.

La ecuación de estado más conocida es la de los gases ideales:

$$P \cdot V = R \cdot T \dots\dots\dots (\text{Ec.1})$$

Donde:

P: Presión absoluta a la que se encuentra el gas.

V: Volumen que ocupa una mol a las condiciones de P y T.

R: Constante Universal de los gases.

T: Temperatura absoluta a la cual se encuentra el gas.

La Ec.1 tiene poca aplicación práctica pues rara vez un gas se comporta como ideal; esto hizo necesario que se buscara el desarrollo de una ecuación de estado para gases reales.

Una de las ecuaciones de estado más usadas en la industria del petróleo en el estudio del comportamiento de fases, es la ecuación de Peng-Robinson la cual presenta la siguiente forma:

$$\left(P + \frac{a(T)}{V(V+b)+b(V-b)} \right) \cdot (V-b) = RT \dots\dots\dots (\text{Ec.2})$$

Los parámetros de esta ecuación están relacionados con las propiedades de los elementos que conforman el fluido sometido a las condiciones de presión, temperatura y volumen y su interrelación entre las diferentes especies químicas presentes en la mezcla. Los parámetros de esta ecuación y las ecuaciones anexas a ésta que permiten definir el comportamiento de fases se dan en el anexo A.

Para aplicar la ecuación de Peng-Robinson a mezclas multicomponentes se

debe conocer la composición de ésta. Esta composición se da en fracción molar de los componentes más livianos y una fracción pesada o fracción C_7^+ pues generalmente en el análisis que se hace a los hidrocarburos se identifica hasta el componente hexano.

Las propiedades físicas de los compuestos puros se pueden obtener de tablas existentes en la literatura, pero la fracción pesada es una mezcla indefinida a la cual se le deben conocer sus propiedades críticas (Presión, Volumen) y el factor acéntrico (?)¹ para poder calcular los parámetros a y b de la ecuación de estado para la mezcla.

Las principales dificultades que se presentan, para aplicar una ecuación de estado a una mezcla de hidrocarburos reales son:

- Manejo de la Fracción pesada C_7^+ .
- Obtención de los coeficientes de interacción binaria (d_{ij}).

Ajuste de Ecuaciones de Estado a Datos PVT

El proceso de ajuste ha sido muy estudiado y aplicado en la industria del petróleo⁶⁻¹⁰, para mejorar las predicciones realizadas con ecuaciones de estado y lograr reproducir el comportamiento real de mezclas de hidrocarburos de yacimiento. Aunque los trabajos son diversos, la mayoría se fundamenta en la minimización de una función de error aplicando métodos numéricos y estadísticos para variar propiedades de componentes como la fracción pesada, Coeficientes de Interacción Binaria (CIB) o parámetros propios de las ecuaciones de estado.

Entre estos trabajos se encuentra el realizado por Coats y Smart¹² donde se presenta un programa de ajuste de los parámetros de la ecuación de estado $-\Omega_a$

¹ El Factor acéntrico es una medida de la desviación del Principio de los Estados Correspondientes de una sustancia dada; Es característico de cada sustancia, y aumenta con el peso molecular de ésta.

y Ω_b — de cualquier componente así como los CIB, y se busca reproducir simultáneamente los diferentes tipos de información experimental disponible para un fluido. Se utilizaron en este trabajo las EdE de Zudkevitch-Joffe¹³, Redlich-Kwong¹⁴ y la de Peng-Robinson¹⁵. En este caso la función a minimizar E dada por:

$$E(\bar{I}) = \sum_{i=1}^{ND} W_i \left| \frac{(d_i - dc_i)}{d_i} \right| \dots\dots\dots (Ec.3)$$

Donde \bar{I} es el conjunto de variables de regresión, d_i el dato experimental, dc_i el dato calculado, W_i el factor de peso que se asigna a la propiedad a la cual se está aplicando la ecuación y ND el número de datos experimentales con que se cuenta. El método de minimización utilizado es una extensión de la técnica de mínimos cuadrados, donde se definen regiones o rangos para cada variable y en cada iteración estas se acotan en subregiones cada vez más pequeñas hasta lograr que todas las variables alcancen un valor estable permaneciendo en la última subregión alcanzada. Es importante anotar que en este estudio se concluye que los parámetros más importantes y que deben estar presentes en cualquier ajuste son: Ω_a y Ω_b del metano y la fracción pesada junto con los CIB entre estos componentes.

Otro estudio realizado acerca de técnicas de regresión para representar el comportamiento de fases fue planteado por Agarwal y Nghiem³ en el cual se realiza un proceso de selección y descarte de variables, que permite utilizarlas o no como parámetros de ajuste dependiendo de la sensibilidad que a éstas presente la función objetivo — función a minimizar—, durante el proceso de ajuste. Dentro del conjunto de posibles parámetros de regresión se encuentran las propiedades críticas de los componentes de la mezcla y los CIB. La función objetivo utilizada fue:

$$E(\bar{I}) = \sum_{i=1}^{ND} \left[W_i \frac{(dc_i - d_i)}{d_i} \right]^2 \dots\dots\dots (Ec.4)$$

Para minimizar la función objetivo se utilizó una modificación del algoritmo de mínimos cuadrados de Dennis¹⁶ que generaliza conceptos de optimización no lineal para seleccionar la dirección y el incremento de las variables. La EdE utilizada para realizar los cálculos fue la de Peng-Robinson.

Por otro lado Da Silva y Rodríguez² presentan una metodología diferente a las anteriores para hacer el ajuste. Ellos dividen el proceso en dos partes, una inicial donde mediante la variación de los CIB entre el metano y la fracción pesada se reproduce la presión de saturación. Posteriormente aplican un proceso de regresión no lineal utilizando la técnica de mínimos cuadrados ajustando los valores de Ω_a y Ω_b del metano y la fracción pesada hasta lograr reproducir con el menor error posible los datos experimentales de una fase. La función objetivo utilizada fue la siguiente:

$$E(\bar{I}) = \sum_{i=1}^{ND} \left[\frac{(dc_i - d_i)}{d_i} \right]^2 \dots\dots\dots (Ec.5)$$

Los cálculos se llevan a cabo con la EdE de Peng-Robinson. Mediante esta técnica los autores lograron reproducir los datos experimentales satisfactoriamente, sin embargo se presentaron desviaciones en el cálculo de la presión de saturación.

Este inconveniente fue posteriormente superado por Trujillo⁹ que utiliza en su metodología el mismo esquema de Da Silva y Rodríguez para llevar a cabo el ajuste, pero lo hace en forma iterativa. Trujillo incluye además un conjunto alternativo de variables de regresión conformado por las propiedades críticas — P_c y T_c — y el factor acéntrico de la fracción pesada. La función objetivo trabajada por Trujillo fue la siguiente:

$$E(\bar{I}) = \sum_{i=1}^{ND} (dc_i - d_i)^2 \dots\dots\dots (Ec.6)$$

En este trabajo se utilizaron datos experimentales de las regiones monofásica y bifásica obteniéndose resultados similares. Es importante

anotar que el tiempo de cálculo empleado al utilizar datos de la región bifásica se incrementa notablemente debido a la complejidad de los cálculos de equilibrio.

En un estudio posterior, San Miguel⁷ retoma la metodología de Da Silva y Rodríguez aplicándola en forma iterativa pero utilizando técnicas de minimización diferentes a la de mínimos cuadrados, como son: mínimo valor absoluto –MVA– y mínimo valor absoluto modificado –MVAM–, para cada una de las cuales utiliza una función objetivo diferente. Para la técnica MVA se minimiza la sumatoria de las desviaciones de los valores absolutos:

$$E(\bar{I}) = \sum_{i=1}^{ND} |d_i - dc_i| \quad \dots\dots (Ec.7)$$

Mientras que, para la técnica MVAM se minimiza:

$$E(\bar{I}) = \sum_{i=1}^{ND} W_i (d_i - dc_i)^2 \quad \dots (Ec.8)$$

Donde:

$$W_i = \frac{1}{|d_i - dc_i|} \quad \dots\dots\dots (Ec.9)$$

El ajuste se lleva a cabo con datos experimentales de la región monofásica y bifásica alcanzando resultados satisfactorios con pequeñas desviaciones y reduciendo el tiempo de cálculo empleado cuando se trabaja en la zona de dos fases.

Esquemáticamente las diferentes técnicas estadísticas de ajuste se representan en la Fig.1.

Algoritmos Genéticos

Los Algoritmos Genéticos (AG) son métodos adaptativos que pueden ser utilizados para implementar búsquedas y problemas de optimización. Ellos están basados en los procesos genéticos de organismos biológicos como la selección, cruzamiento y mutación, codificando una posible solución a un problema en un "cromosoma" compuesto por una cadena de bits o caracteres¹⁷.

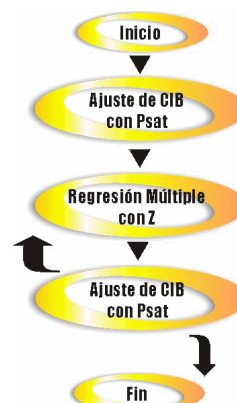


Fig. 1. Estructura jerárquica del ajuste estadístico de los parámetros de la ecuación de estado⁶.

Los Algoritmos Genéticos utilizan una analogía directa del fenómeno de evolución en la naturaleza. Trabajan con una población de individuos, cada uno representando una posible solución a un problema dado. A cada individuo se le asigna una puntuación de adaptación, dependiendo de que tan buena fue la respuesta al problema. A los más adaptados se les da la oportunidad de reproducirse mediante cruzamientos con otros individuos de la población, produciendo descendientes con características de ambos padres. Los miembros menos adaptados poseen pocas probabilidades de que sean seleccionados para la reproducción, y desaparecen.

Una nueva población de posibles soluciones es generada mediante la selección de los mejores individuos de la generación actual, emparejándolos entre ellos para producir un nuevo conjunto de individuos. Esta nueva generación contiene una proporción más alta de las características poseídas por los mejores miembros de la generación anterior. De esta forma, a lo largo de varias generaciones, las características buenas son difundidas a lo largo de la población mezclándose con otras. Favoreciendo el emparejamiento de los individuos mejor adaptados, es posible recorrer las áreas más prometedoras del espacio de búsqueda. Si el Algoritmo Genético ha sido diseñado correctamente, la población

convergerá a una solución óptima del problema.

Los dos procesos que más contribuyen a la evolución son el cruce y la adaptación basada en la selección / reproducción. La mutación también juega un papel significativo pero es de menor relevancia en el funcionamiento de un Algoritmo Genético, aunque la mutación utilizada de una manera correcta puede ser de gran ayuda para la convergencia del Algoritmo, sin embargo ella no debe ser utilizada demasiado, ya que el Algoritmo Genético se puede convertir en una búsqueda al azar, pero su utilización asegura que ningún punto en el espacio de búsqueda tiene probabilidad 0 de ser examinado.

Terminología. Para comprender el funcionamiento de los algoritmos genéticos es necesario tener en cuenta los siguientes conceptos básicos:

Gen: Es un campo que puede tomar el valor de 0 ó 1.

Cromosoma: Es un conjunto de varios genes, cuya longitud (L) depende de la precisión con que se quiera trabajar. Poseen ciertos parámetros que son característicos de cada uno como su valor, su aptitud, porcentaje de aptitud también llamado probabilidad, entre otros. Llamados también individuos (Teoría de Darwin), y es sobre los cuales se aplican los operadores, que más adelante se describirán.

Generación: Es un conjunto de cromosomas de tamaño N. También conocida como población. También posee parámetros propios como la Aptitud de la generación.

Función de Aptitud: Es la función que nos determina que tan bueno o tan malo es un cromosoma.

Función Objetivo: Es la ecuación o modelo matemático que representa el problema a resolver en forma abstracta y aproximada.

Operadores. En la utilización de los algoritmos genéticos, se utilizan operadores que permiten efectuar el proceso de optimización numérica entre los datos ó población de trabajo del problema a tratar. Estos operadores son:

Selección: Simula el proceso de selección natural, en el que el más fuerte tiene mayor capacidad de supervivencia. En los AG la capacidad de supervivencia de un individuo esta relacionado con el valor numérico de la función objetivo (*fitness of good*). Como el algoritmo genético es un algoritmo de carácter iterativo al igual que los métodos numéricos de optimización, esta selección se aplica en cada iteración, sobre una población de N individuos, para escoger los mejores N individuos de la siguiente iteración. Por consiguiente, en los N individuos escogidos para la siguiente iteración, pueden encontrarse dos o más individuos idénticos debido a que los individuos con bajo *fitness of good* tienen menos posibilidades de ser escogidos, en cambio, los que poseen altos *fitness of good* tienen más posibilidades de ser frecuentemente seleccionados.

Cruza: Una vez que los individuos fueron seleccionados se aplica a cada par de individuos escogidos, el operador de cruzamiento. En primer lugar, se escoge aleatoriamente un punto de corte que se aplicará al par de cromosomas de los individuos seleccionados. Luego, los caracteres más significativos, a partir del punto de corte, conservan sus posiciones relativas en el nuevo par de cromosomas y los caracteres restantes son intercambiados de los cromosomas progenitores a los nuevos dos cromosomas obtenidos.

Mutación: El proceso de mutación es básicamente una búsqueda aleatoria. Se selecciona aleatoriamente una posición específica dentro de un cromosoma, para cambiar luego el valor contenido en dicha posición. Como en la naturaleza, la probabilidad de que ocurra una mutación en los individuos de una especie es muy pequeña, en condiciones normales de

vida, en el AG se trata de representar lo mismo con un valor de probabilidad muy baja para la aplicación del operador de mutación.

Estructura del algoritmo genético.

Para la utilización de un algoritmo genético en un problema de optimización, se debe utilizar un esquema de trabajo como el que se da en la Fig.2. La utilización del algoritmo genético será efectiva si se escogen unos buenos datos iniciales, las reglas y criterios de cruce, mutación y estimación de la población final.

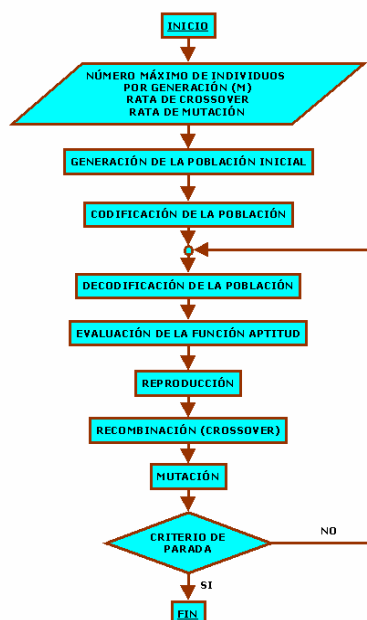


Fig. 2. Estructura jerárquica de un algoritmo genético¹⁸

Planteamiento del Problema

Las ecuaciones de estado son muy importantes en el estudio de fases, una de las más utilizadas debido a su sencillez y aproximación en los resultados producidos es la Ecuación de Peng-Robinson. Una de las propiedades físicas de la fracción pesada que más influye en el resultado de la ecuación de estado de Peng-Robinson, es el coeficiente de interacción binaria entre el Metano (C_1) y la fracción pesada (C_7^+) ($d_{C_1-C_7^+}$), por ello se debe tener especial cuidado en el cálculo de dicho parámetro; Otro

parámetro de igual importancia en el comportamiento de la ecuación de estado es el factor acéntrico de la fracción pesada ($\phi_{C_7^+}$) el cual mide la desviación de la idealidad del pseudocomponente pesado.

Como se ha visto, los diversos desarrollos estadísticos ofrecen la posibilidad de acercar notoriamente los resultados de la ecuación de estado a la data PVT. Estos permiten recalculer los coeficientes de interacción binaria y los factores acéntricos de manera efectiva. La desventaja aquí presente radica en que el esfuerzo numérico no tiene en cuenta los límites físicos que las variables deben tener, si se efectúan cambios sustanciales en los valores de los parámetros individuales, los compuestos a quienes ellos representan en la ecuación de estado dejan de ser y se transforman en otros. Otra desventaja de la metodología tradicional de minimización de cuadrados es la posibilidad de encontrar mínimos locales y dar respuestas falsas, la metodología de optimización basada en los algoritmos genéticos permite encontrar un mínimo global ya que utiliza como población la totalidad de las posibles respuestas en su desarrollo, permitiendo encontrar un mínimo global en los rangos estipulados para cada variable.

Metodología Propuesta

Para resolver el problema planteado se desarrolló una nueva metodología con Algoritmos Genéticos para ajustar el coeficiente ($d_{C_1-C_7^+}$) y el factor acéntrico ($\phi_{C_7^+}$) con los datos del experimento a composición constante.

Para ajustar el coeficiente ($d_{C_1-C_7^+}$) y el factor acéntrico ($\phi_{C_7^+}$), se tomó como base la metodología utilizada en la aplicación PVTLAB¹⁹, la cual se basa en la simulación de los experimentos PVT de gases condensados y aceites volátiles. Esta metodología utiliza la ecuación de estado y los cálculos flash de equilibrio entre la fase gaseosa y la fase líquida de un fluido composicional de yacimiento

para simular pruebas experimentales de laboratorio PVT.

De manera general la Fig.3 muestra el desarrollo de los cálculos en la metodología planteada para trabajar con una ecuación de estado. Se muestra el cambio en la función de actualización de variables, donde se utiliza el algoritmo genético en reemplazo de la función tradicional basada en los parámetros de interacción binaria entre el metano y la fracción pesada (d_{C1-C7+}) y el factor acéntrico de la fracción pesada y de esta manera optimizar la selección de estos parámetros.

Cálculo del volumen relativo. En el algoritmo genético se utilizó como población inicial los valores de d_{C1-C7+} calculados por el método de Varotsis²⁰ y el factor acéntrico de la fracción pesada ($?_{C7+}$) calculado por medio de las correlaciones de propiedades físicas de los hidrocarburos propuesto por Whitson²¹. De la población inicial de valores se generó cada nueva población por el cruce de los cromosomas binarios de estos datos. Para evaluar los individuos más aptos de la población se utilizó una función de aptitud definida por la comparación del error generado al evaluar los volúmenes relativos del experimento PVT dados por la ecuación de estado con los parámetros d_{C1-C7+} , $?_{C7+}$ obtenidos genéticamente contra la data experimental. Luego de tener cada nueva población se efectuó el mismo procedimiento hasta alcanzar la convergencia definida por el cálculo flash.

La población de datos en cada iteración se define por medio de las correlaciones de Varotsis y Whitson, las cuales son tomadas de experimentación con la fracción pesada, por lo cual representan fielmente el comportamiento de esta fracción a diferentes condiciones de temperatura y presión. Estas correlaciones experimentales aseguran al algoritmo genético tener siempre poblaciones que cumplan con los rangos de validez de los parámetros.

Cálculo de la presión de saturación.

Para el cálculo de la presión de saturación el procedimiento con algoritmos genéticos es equivalente al efectuado para el volumen relativo, la función de aptitud que se define en este caso es el error entre la presión de saturación dada por la ecuación de estado y la obtenida en la experimentación PVT. A su vez se utilizan las mismas correlaciones de propiedades de Varotsis y Whitson para la fracción pesada.

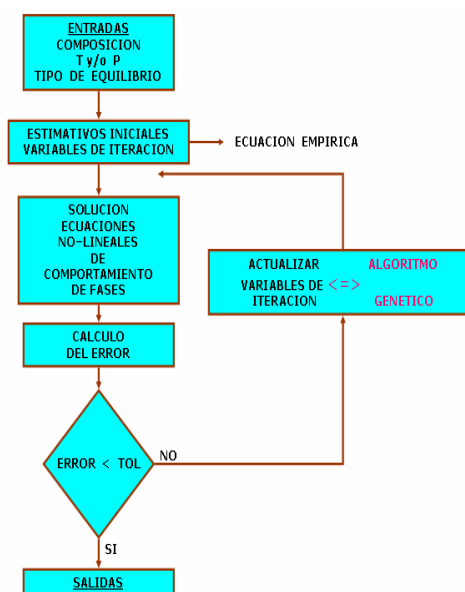


Fig.3. Cálculo del comportamiento de fases por medio de una Ecuación de Estado.¹⁹

Análisis de Resultados

Con el fin de validar la metodología y comprobar la consistencia de las predicciones del comportamiento de fases, se utilizaron las muestras Gas 2*, Gas 2**, Gas 5, Oil 1, Oil 2, Oil 3, Oil 4, Oil5, Oil 6 del estudio efectuado por Coats y Smart²², la muestra NS-1 de Whitson²³ y el fluido 88002 de Corelab²⁴.

Para la evaluación de los resultados se utilizó la desviación absoluta promedio definida por:

$$DAP = \frac{\sum_{i=1}^n \frac{|Valormedido - Valorcalculado|}{Valormedido}}{n} \quad \text{Ec.10)}$$

Para el cálculo del error porcentual se utilizó la siguiente expresión:

$$\%Error = \frac{(P_{exp} - P_{calc})}{P_{exp}} * 100 \quad \dots (Ec.11)$$

Los resultados numéricos y comparativos se muestran en el anexo B.

Conclusiones

- Se desarrolló una nueva metodología conjunta entre los Algoritmos Genéticos y la metodología utilizada en PVTLAB, para el ajuste de los Parámetros de la Ecuación de Estado de Peng-Robinson.
- Se verificó la gran influencia que tienen los parámetros de interacción binaria y el factor acéntrico de la fracción pesada en el comportamiento composicional de una muestra de gas condensado y/o aceite volátil.
- La nueva metodología habilita a la ecuación de estado de Peng-Robinson para efectuar cálculos volumétricos en la región bifásica, donde es bien sabido que está ecuación presenta fallas debido a su baja definición de la fracción líquida.
- La nueva metodología se implementó dentro de una aplicación software llamada PVT_AG. Los resultados que se derivan de esta son la disminución del error relativo total y una mejor aproximación al valor experimental de la presión de saturación en comparación con los resultados obtenidos con PVTLAB.
- La gran ventaja de la utilización del método AG radica en el nivel de exactitud y precisión mucho mayor que con la metodología PVTLAB.
- La principal desventaja del nuevo método es el tiempo utilizado en los cálculos, esta desventaja se debe a la gran cantidad de iteraciones que los AG tiene que efectuar antes de dar la

respuesta. Pero en este tipo de problemas de ajuste se requiere una mayor exactitud y precisión en la respuesta y este mayor tiempo de cálculo se compensa con la exactitud que se debe manejar en la simulación de millones de celdas en una malla composicional, disminuyendo así el tiempo total de cálculo en la simulación composicional.

- La metodología desarrollada permite mejorar de manera Meta-Heurística la selección de los parámetros de ajuste en la ecuación de estado.
- Se aplicaron los Algoritmos Genéticos al experimento de expansión constante, permitiendo simular por medio de la ecuación de estado las propiedades PVT de la muestra: Volumen relativo y presión de saturación.

Nomenclatura

$a(t)$	Coefficiente de la EDE.
b	Coefficiente de la EDE.
C_1	Metano.
C_{n+}	Fracción pesada.
d	Dato experimental.
dc	Dato calculado.
E	Función error.
P	Presión, MPa.
P_c	Presión crítica, MPa.
R	Constante universal de los gases.
T	Temperatura, °K.
T_c	Temperatura crítica, °K.
V	Volumen, m ³ .
V_c	Volumen crítico, m ³ .
W	Factor de peso.
I	Vector de variables de regresión.
W_a	Parámetro de la EDE.
W_b	Parámetro de la EDE.
d_{ij}	Coefficiente de interacción binaria.
$?$	Factor acéntrico.

Superíndices y subíndices

ND	Número de datos experimentales
i	elemento
j	elemento

1. Schlumberger (Geoquest - Intera Group). *PVTi Reference Manual. Version 99A*. Denver. Colorado. USA. 1999.
2. Da Silva M. B., Rodriguez, F. "Automatic Fitting of Equation of State for Phase Behavior Matching". SPE 23703. (1992).
3. Agarwal, Rajeev K.; Li, Yau-Kun and Nghiem, Long X. "A Regression Technique with Dynamic Parameter Selection for Phase Behavior Matching". *SPE* (Feb/90). p.115 - 120.
4. Wang P. and Pope G.A. "Proper Use of Equations of State for Compositional Reservoir Simulation". Distinguished Author Series. *JPT* (Jul/01). p.74-81.
5. Osorio Raúl. "Estrategias de Modelamiento Composicional de Fluidos en Yacimientos Colombianos de Aceite Volátil, Gas Condensado y Sistemas Cercanos al Punto Crítico". VIII Congreso Colombiano del Petróleo. Bogotá. 1999.
6. Castiblanco Galvis, Iowany y Niz Velásquez Eider. *Ajuste de la Ecuación de Estado de Peng-Robinson a Datos PVT de Yacimientos Cercanos al Punto Crítico Mediante Regresión Múltiple*. Proyecto de Grado UIS. Bucaramanga (2000).
7. San Miguel, Luis Alberto. *Métodos Alternos para el Ajuste Automático de las Ecuaciones de Estado*. Tesis de Maestría UNAM. México (1998).
8. Merrill, R.C.; Hartman K.J. and Creek J.L. "A Comparison of Equation of State Tuning Methods". SPE 28598. (1994).
9. Trujillo, C. W. *Ajuste Automático de la Ecuación de Estado de Peng-Robinson a Mezclas de Hidrocarburos*. Tesis de Maestría UNAM. México (1996).
10. Christensen J. "Regression of PVT Data". *J. Can. Pet. Tech.* Vol. **38**(1999), p.57.
11. Pedersen K.S.; Thomassen P. and Fredenslund A. "On the Dangers of Tuning Equations of State Parameters". *Chem. Eng. Sci.* (1988). p.269 - 278.
12. Coats K.H. and Smart G. T. "Application of a Regression Bases EOS PVT Program to Laboratory Data". *SPE* (May/86). p.277-299.
13. Zudkevitch D. and Joffe J. "Correlation and Prediction of Vapor - Liquid Equilibria with the Redlich-Kwong Equation of State". *AIChE Journal* (Jan/70). p.112-119.
14. Redlich O. and Kwong J. "On the Thermodynamics of Solutions. An Equation of State. Fugacities of Gaseous Solutions". *Chemical Reviews* **44**(1949). p.233-247.
15. Peng D.Y. and Robinson D.B. "A new Two-Constant Equation of State". *Ind. Eng. Chem. Fund.* **15**(1976). No.1. p.59-64.
16. Dennis Jr, J.E. Non-Linear Least Squares and Equations. In *The State of the Art in Numerical Analysis*. (A.H. Jacobs Ed.) p. 269-312. Academic Press. London. UK. 1977.
17. Goldberg, David. *Genetic Algorithms in search, optimization, and machine learning*. Adison Wesley. USA. 1989.
18. Fernando E. Calvete G.; Nicolás Santos S.; Silvia P. Román V.; Claudia P. Soto T. "Computación Evolutiva: Algoritmos Genéticos y Simulated Annealing, Técnicas Meta-Heurísticas de Optimización en Ingeniería de Petróleos". *IX Congreso Colombiano del Petróleo. ACIPET. Bogotá. Colombia. Octubre 2001*.
19. Fernando E. Calvete G.; Nohora P. Rangel. "Simulación de Experimentos PVT Para Yacimientos de Gas Condensado y Aceite Volátil". *VIII Congreso Colombiano del Petróleo. ACIPET. Bogotá. Colombia. Octubre 1999*.
20. Varotsis N.; Stewart G.; Todd A.C. and Clancy M. "Phase Behavior of Systems Comprising North Sea Reservoir Fluids and Injection Gases". *JPT* (Nov/86).

21. Whitson C.H. "Characterizing Hydrocarbon Plus Fractions". *SPEJ* (Aug/83).
22. Coats, K.H. and Smart, G.T. "Application of a Regression Bases EOS PVT Program to Laboratory Data". *SPERE* (May/80).
23. Whitson C.H. and Torp S.B. "Evaluating Constant Volume Depletion Data". *JPT* (March/83).
24. CORELAB Laboratories, Report 88002.

Tabla A-1. Ecuaciones y procedimientos numéricos utilizados en el Comportamiento de Fases presente en las mediciones PVT de laboratorio.

PROCEDIMIENTO	ECUACIONES
Cálculos Volumétricos y Relaciones PVT	<p>Ecuación de estado de Peng-Robinson:</p> $P = \frac{RT}{v-b} - \frac{a(T)}{v(v+b)+b(v-b)} \quad \text{ó} \quad Z^3 - (1-B)Z^2 + (A-3B^2-2B)Z - (AB-B^2-B^3) = 0$ $a(T) = a(T_c) \alpha(T_r, w) \quad A = \frac{aP}{R^2 T^2} \quad \alpha = (1 + m(1 - T_r^{0.5}))^2$ $b(T) = b(T_c) \quad b(T_c) = 0.07780 \frac{RT_c}{P_c} \quad B = \frac{bP}{RT} \quad m = 0.379642 + 1.48503w - 0.1644w^2 + 0.016667w^3$
Cálculo del Equilibrio de Fases	<p>Para el equilibrio: $f_i^V = f_i^L$</p> $\ln\left(\frac{f}{P}\right) = (Z-1) - \ln(Z-B) - \frac{A}{2\sqrt{2}B} \ln\left(\frac{Z+2.414B}{Z-0.414B}\right)$ $\Phi_i^V = \frac{f_i^V}{y_i P} \quad \Phi_i^L = \frac{f_i^L}{x_i P}$ $\ln \Phi_i^L = \frac{b_i}{b} (Z^L - 1) - \ln(Z^L - B) - \frac{A}{2\sqrt{2}B} \left[\frac{2y_i - b_i}{y} \right] \ln \left[\frac{Z^L + 2.414B}{Z^L - 0.414B} \right]$ $y_i = \sum_j [x_j (a_i a_j a_i a_j)^{0.5} (1 - d_{ij})] \quad y = \sum_i \sum_j [x_i x_j (a_i a_j a_i a_j)^{0.5} (1 - d_{ij})]$ $a = \sum_i \sum_j x_i x_j a_{ij} \quad b = \sum_i x_i b_i$ $a_{ij} = (1 - d_{ij}) a_i^{1/2} a_j^{1/2}$
Cálculos Flash	<p>1. Balance total de materia: $F = L + V$ 2. Balance de materia por componente: $Z_i F = X_i (1-\beta) + Y_i \beta$</p> <p>Donde $b = \frac{V}{F}$</p> <p>3. Equilibrio termodinámico: $f_i^L = f_i^V$ Definiendo la relación de equilibrio como:</p> $K_i = \frac{Y_i}{X_i} = \frac{f_i^V}{f_i^L}$ <p>Reduciendo la expresión: $Y_i = \frac{K_i \times Z_i}{1 + b \times (K_i - 1)} \quad \text{y} \quad X_i = \frac{Z_i}{1 + b \times (K_i - 1)}$</p> <p>Aplicando el balance de materia en cada fase: $\sum_{i=1}^{N_C} Z_i = \sum_{i=1}^{N_C} X_i = \sum_{i=1}^{N_C} Y_i = 1$</p> <p>De tal manera: $\sum_{i=1}^{N_C} X_i = \sum_{i=1}^{N_C} \frac{Z_i}{1 + b \times (K_i - 1)} = 1$ Y por lo mismo:</p> $\sum_{i=1}^{N_C} Y_i = \sum_{i=1}^{N_C} \frac{K_i \times Z_i}{1 + b \times (K_i - 1)} = 1$

Tabla B-1. Comparación de resultados en el cálculo de la presión de saturación para sistemas de hidrocarburos por medio de diferentes procedimientos .

MEDICION	GAS 2*		GAS 2**		GAS 5		NS-1		RFL-88002	
EXPERIMENTAL	4465 (190°F)		4430 (190°F)		4857 (267°F)		6764.7 (280°F)		4535.7 (275°F)	
CALC (PR-EOS)	3680	17.58%	3664	17.29%	4494	7.47%	ND		ND	
CALC (PVTLAB)	3899	12.68%	3827	13.61%	4808	1.00%	6201	8.33%	4047.5	10.76%
CALC (PVTLAB_AG)	4435	0.68%	4415	0.34%	4840	0.35%	6678	1.28%	4435	2.22%
MEDICION	OIL 1		OIL 2		OIL 3		OIL 4		OIL 6	
EXPERIMENTAL	2534.7 (180°F)		4475 (176°F)		2612 (180°F)		2562 (250°F)		2760.7 (234°F)	
CALC (EDE PR)	ND (*)		3344	25.27%	2195	15.96%	2259	11.83%	2412.7	12.61%
CALC (AHMED)	ND (*)		4502	-0.603%	2289	12.37%	2640	-3.044%	2930.7	-6.16%
CALC (PVTLAB)	2494.3	1.59%	4478.5	-0.078%	2368.5	09.32%	2647.45	-3.335%	2907	-5.29%
CALC (PVTLAB_AG)	2530.8	0.15%	4470.5	0.10%	2561	1.95%	2534.6	1.07%	2780.1	-0.70%

(*) ND: Dato No Disponible.

Tabla B-2. Resultados para el Gas 2.

MEDIDA	EXP	PVTLAB	PVT_AG
d_{C1-C7+}		0,061441	0,10366484
$?_{C7+}$		0,58734834	0,53246336
PRESIÓN SAT	4465	3823,04	4434,63
DAP (%)		0,5862993	0,24740387
ERROR RELATIVO (%)		6,55397992	1,16634421
ERROR (PSAT) (%)		14,3778	0,68017086

Tabla B-3. Resultados de la fracción volumétrica para el Gas 2.

PRESIÓN	% FRACCIÓN DE VOLUMEN		
(PSIG)	EXPERIMENTAL	PVTLAB	PVT_AG
4440	4,35	0	0
4420	47,38	0	33,56353
4388	50,82	0	34,53989
4339	51,64	0	35,66077
4300	51,94	0	36,63072
4180	51,95	0	39,34753
3993	51,32	0	43,05863
3780	50,07	35,75683	46,79117
3490	47,86	43,69655	48,57951
2998	42,96	45,78994	41,18708
2505	36,75	36,51561	33,54723
2000	28,88	27,32753	25,7809
1485	20,2	18,15569	17,42014
1058	13,06	11,05435	10,73368

Tabla B-4. Interfase de la aplicación PVT_AG.

Ajuste a la Ecuación de Peng-Robinson

Expansión a Compresión Constante para un Gas a una Temperatura de 190 F y Presión de Saturación de 4465 psig

Iniciar Ver Resultados Composición del Gas

Resultados del Ajuste

Datos Experimentales	PVTLAB	Datos PVT_AG
4.35	0	0
47.38	0	33.56353
50.82	0	34.53989
51.64	0	35.66077
51.84	0	36.63072
51.95	0	38.34753
51.32	0	43.05853
59.07	35.75583	45.75117
47.86	43.03855	48.37315
42.96	45.70994	41.18708
36.75	36.51561	33.74523
28.88	27.32751	25.7809
20.2	18.15563	17.42014
13.06	11.05435	10.73368

Tabla B-5. Comparación grafica de resultados de la aplicación PVT_AG.

