

**METODOLOGÍA PARA LA PREDICCIÓN DE CURVAS DE DECLINACIÓN
DE POZOS DE PETRÓLEO, APLICANDO REDES NEURONALES
ARTIFICIALES**

**EMILIO JUSTINIANO CARCAMO TROCONIS
GUIDO JOSÉ POLO NULE**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO MECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA
2007**

**METODOLOGÍA PARA LA PREDICCIÓN DE CURVAS DE DECLINACIÓN
DE POZOS DE PETRÓLEO, APLICANDO REDES NEURONALES
ARTIFICIALES**

**EMILIO JUSTINIANO CARCAMO TROCONIS
GUIDO JOSÉ POLO NULE**

**Trabajo de grado presentado para optar el título de
Ingeniero de Sistemas.**

**Director:
ING. JUAN CARLOS REYES FIGUEROA
Profesor Escuela de Ingeniería de Sistemas, UIS**

**Codirector:
M.E. FERNANDO RUIZ DÍAZ
Profesor Escuela de Ingeniería de Sistemas, UIS**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO MECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA
2007**

A Dios
A mis padres, Beatriz y Justo
A mis hermanos, Eugenia y Ernesto

EMILIO CARCAMO

A Fanny , mi mamá
A Guido, mi papá en lo alto
Dayana y María J, sístole y diástole de mi corazón

GUIDO

AGRADECIMIENTOS

Los autores expresan sus agradecimientos a:

Todas aquellas personas que hicieron posible la realización de este proyecto.

MSc FERNANDO CALVETE, por sus consejos y aportes al proyecto.

Dayana Sarmiento por su colaboración

A todos nuestros amigos, que nos soportaron durante el paso por la universidad, y que incondicionalmente brindaron su confianza, apoyo y honestidad.

A NUESTRAS FAMILIAS, por su paciencia y apoyo incondicional.

TABLA DE CONTENIDO

	PÁG.
CAPITULO 1	20
INTRODUCCIÓN	20
<i>1.1 CONTEXTO GENERAL</i>	22
1.1.1 DEFINICIÓN DEL PROBLEMA	22
1.1.2 OBJETIVOS	24
1.1.3 JUSTIFICACIÓN	25
CAPITULO 2	26
MARCO TEÓRICO	26
<i>2.1 REDES NEURONALES ARTIFICIALES</i>	26
2.1.1 Panorama Histórico	27
2.1.2 Ventajas de las RNA	28
2.1.3 Fundamentos de las RNA	29
2.1.4 Algoritmo Backpropagation	37
<i>2.2 CURVAS DE DECLINACIÓN</i>	42
2.2.1 Introducción	42
2.2.2 Estado del Arte	44
2.2.3 Fundamento Matemático	51
<i>2.3 MECANISMOS DE PRODUCCIÓN DEL PETRÓLEO</i>	56
2.3.1 Bombeo Mecánico	57
2.3.2 Bombeo neumático	57
2.3.3 Bombeo Hidráulico	58
2.3.4 Bombeo Electro-Centrífugo o Electro-Sumergido	59

CAPITULO 3	61
PROCESO DE IDENTIFICACIÓN DEL MODELO	61
3.1 <i>SELECCIÓN DE LAS VARIABLES</i>	62
3.1.1 Línea mensual de tiempo	62
3.1.2 Producción de petróleo	63
3.1.3 Días de producción de cada mes	65
3.1.4 Tipo de bombeo	66
3.2 <i>PREPROCESAMIENTO DE LA INFORMACIÓN</i>	68
3.2.1 Transformación de Variables Nominales a Numéricas	70
3.2.2 Filtrado	70
3.2.3 Normalización	71
3.3 <i>ARQUITECTURA DE RED</i>	73
3.3.1 Redes de Base Radial (RBF)	73
3.3.2 Sistemas Neuro-Borrosos	74
3.3.3 Perceptrón Multicapa	76
3.4 <i>DIMENSIONAMIENTO DE LA RED</i>	77
3.5 <i>FUNCION DE ACTIVACIÓN</i>	83
3.6 <i>ALGORITMO DE ENTRENAMIENTO</i>	84
3.7 <i>INICIALIZACIÓN DE PESOS</i>	86
3.8 <i>RESULTADOS DEL ENTRENAMIENTO</i>	88
CAPITULO 4	91
METODOLOGÍA DE DESARROLLO	91
4.1 <i>RECOPILACIÓN DE NECESIDADES</i>	91
4.2 <i>ANÁLISIS</i>	92
4.3 <i>DISEÑO</i>	92
4.4 <i>DESARROLLO</i>	93

4.5 HERRAMIENTA DE DESARROLLO	93
4.6 APLICACIÓN DE LA METODOLOGÍA	95
4.6.1 Prototipo 1	95
4.6.2 Prototipo 2	97
4.6.3 Prototipo 3	98
4.7 DIAGRAMAS UML	99
4.7.1 Diagrama de Casos de Uso	99
4.7.2 Diagramas de Clases	104
4.7.3 Diagrama de Secuencias	109
CAPITULO 5	113
ANÁLISIS DE RESULTADOS	113
5.1 POZO 1	114
5.2 POZO 2	117
5.3 POZO 3	119
5.4 POZO 4	122
5.5 POZO 5	124
5.6 POZO 6	127
5.7 POZO 7	129
5.9 OBSERVACIONES	132
5.10 CASO DE ESTUDIO, POZO 8	132
5.10.1 Curvas De Declinación	132
5.10.2 Producción Acumulada	136
5.10.3 Tasa de declinación	136
5.10.4 Límite Económico	139
5.11 VERIFICACION DE LA HERRAMIENTA	141
6. CONCLUSIONES	145

7. RECOMENDACIONES	147
8. BIBLIOGRAFIA	148
9. ANEXO A: MANUAL DE USUARIO	152
9.1 INTRODUCCION	152
9.2 Carga de datos.	152
9.3 Entrenamiento.	155
9.4 Simulación	157
9.5 Presentación de resultados	158
9.5.1 Curvas De Declinación	160
9.5.2 Producción Acumulada	161
9.5.3 Tasa De Declinación	161
9.5.4 Limite Económico	162
ANEXO B: DATOS DE PRODUCCIÓN DEL POZO DE PRUEBA	164

LISTA DE FIGURAS

	PÁG.
Figura 1. Del modelo biológico al modelo artificial	30
Figura 2 Estructura de una Red Neuronal	35
Figura 3 Clasificación de los tipos de redes	36
Figura 4. Definición de tasa de declinación - gráfico lineal	52
Figura 5. Bombeo mecánico	57
Figura 6 Bombeo neumático	58
Figura 7. Bombeo Hidráulico	58
Figura 8. Bombeo Electro-Centrífugo	59
Figura 9. Simulación del Modelo 1	64
Figura 10. Simulación del Modelo 2	66
Figura 11. Simulación del Modelo 3	68
Figura 12. Esquema del preprocesamiento de la información	69
Figura 13. Diagrama de flujo del filtrado de datos	72
Figura 14. Simulación y validación utilizando RBF	75
Figura 15. Simulación y validación utilizando ANFIS	76
Figura 16. Modelo de red con las variables de entrada y salida	83
Figura 17. Resultados del entrenamiento	89
Figura 18. Diagrama de casos de uso	104
Figura 19. Diagrama de Clases	104
Figura 20. Diagrama de secuencias: Cargar datos, entrenar y simular la RNA	109
Figura 21. Diagrama de secuencias: Curva de declinación	110
Figura 22. Diagrama de secuencias: Producción acumulada	111
Figura 23. Diagrama de secuencias: Tasa de declinación	111
Figura 24. Diagrama de secuencias: Análisis Económico	112
Figura 25. Pozo 1: Comparación entre las técnicas clásicas y la producción original	115

Figura 26. Pozo 1: Comparación entre DECLINET y la producción original	115
Figura 27. Pozo 1: Comparación entre las técnicas clásicas y la producción acumulada original	116
Figura 28. Pozo 1: Comparación entre DECLINET y la producción acumulada original	116
Figura 29. Pozo 2: Comparación entre las técnicas clásicas y la producción original	117
Figura 30. Pozo 2: Comparación entre DECLINET y la producción original	118
Figura 31. Pozo 2: Comparación entre las técnicas clásicas y la producción acumulada original	118
Figura 32. Pozo 2: Comparación entre DECLINET y la producción acumulada original	119
Figura 33. Pozo 3: Comparación entre las técnicas clásicas y la producción original	120
Figura 34. Pozo 3: Comparación entre DECLINET y la producción original	120
Figura 35. Pozo 3: Comparación entre las técnicas clásicas y la producción acumulada original	121
Figura 36. Pozo 3: Comparación entre DECLINET y la producción acumulada original	121
Figura 37. Pozo 4: Comparación entre las técnicas clásicas y la producción original	122
Figura 38. Pozo 4: Comparación entre DECLINET y la producción original	123
Figura 39. Pozo 4: Comparación entre las técnicas clásicas y la producción acumulada original	123
Figura 40. Pozo 4: Comparación entre DECLINET y la producción acumulada original	124
Figura 41. Pozo 5: Comparación entre las técnicas clásicas y la producción original	125
Figura 42. Pozo 5: Comparación entre DECLINET y la producción original	125
Figura 43. Pozo 5: Comparación entre las técnicas clásicas y la producción acumulada original	126

Figura 44. Pozo 5: Comparación entre DECLINET y la producción acumulada original	126
Figura 45. Pozo 6: Comparación entre las técnicas clásicas y la producción original	127
Figura 46. Pozo 6: Comparación entre DECLINET y la producción original	128
Figura 47. Pozo 6: Comparación entre las técnicas clásicas y la producción acumulada original	128
Figura 48. Pozo 6: Comparación entre DECLINET y la producción acumulada original	129
Figura 49. Pozo 7: Comparación entre las técnicas clásicas y la producción original	130
Figura 50. Pozo 7: Comparación entre DECLINET y la producción original	130
Figura 51. Pozo 7: Comparación entre las técnicas clásicas y la producción acumulada original	131
Figura 52. Pozo 7: Comparación entre DECLINET y la producción acumulada original	131
Figura 53. Pozo ejemplo: Producción original	133
Figura 54. Pozo ejemplo: Comparación entre DECLINET y la producción original en escala mensual	135
Figura 55. Pozo ejemplo: Comparación entre DECLINET y la producción original en escala diaria	135
Figura 56. Pozo ejemplo: Comparación entre DECLINET y la producción acumulada original	136
Figura 57. Tasa de declinación generada por DECLINET, para una periodicidad de 24 meses	138
Figura 58. Tasa de declinación generada por DECLINET, para una periodicidad de 60 meses	138
Figura 59. Pozo ejemplo: calculo del límite económico, caso 1	140
Figura 60. Pozo ejemplo: calculo del límite económico, caso 2	141
Figura 61. Distribución de pesos y ganancias en el modelo de red propuesto	141

Figura 62. Simulación de la RNA en matlab, dado un conjunto de pesos iniciales	143
Figura 63. Simulación de la RNA en DECLINET, dado un conjunto de pesos iniciales	144
Figura 64. Formato de archivo en Excel	153
Figura 65. DECLINET 1.0: Formas de seleccionar el pozo a estudiar	153
Figura 66. DECLINET 1.0: Abrir archivos de alimentación	154
Figura 67. DECLINET 1.0: Datos del pozo escogido	154
Figura 68. DECLINET 1.0: Selección de parámetros de entrenamiento	155
Figura 69. DECLINET 1.0: Resultados del entrenamiento	156
Figura 70. DECLINET 1.0: Selección de los parámetros de simulación	157
Figura 71. DECLINET 1.0: Resultados de la simulación	158
Figura 72. DECLINET 1.0: Presentación grafica de resultados	159
Figura 73. DECLINET 1.0: Comparación gráfica entre la curva de producción real y la simulada	160
Figura 74. DECLINET 1.0: Comparación grafica entre la producción acumulada real y la simulada	161
Figura 75. DECLINET 1.0: Gráfica de la tasa de declinación con periodicidad de 24 meses	161
Figura 76. DECLINET 1.0: Gráfica de la tasa de declinación con periodicidad de 60 meses	162
Figura 77. DECLINET 1.0: Presentación grafica y detallada de el calculo del limite económico	163

LISTA DE TABLAS

	PÁG.
Tabla 1 Principales Funciones de Transferencia.	33
Tabla 2. Errores de entrenamiento y validación para el Modelo 1	63
Tabla 3. Entradas y salidas de la red neuronal al Modelo 1	64
Tabla 4. Errores de entrenamiento y validación para el Modelo 2	65
Tabla 5. Entradas y salidas de la red neuronal Modelo 2	65
Tabla 6. Errores de entrenamiento y validación para el Modelo 3	67
Tabla 7. Entradas y salidas de la red neuronal Modelo 3	67
Tabla 8. Transformación de variables nominales a numéricas	70
Tabla 9. Errores de entrenamiento y validación para la red RBF	74
Tabla 10. Errores de entrenamiento y validación para ANFIS	75
Tabla 11. Errores de entrenamiento y tiempo consumido para redes de 1 capa oculta	79
Tabla 12. Errores de validación para redes de 1 capa oculta	79
Tabla 13. Errores de entrenamiento y tiempo consumido para redes de 2 capas ocultas	80
Tabla 14. Errores de validación para redes de 2 capas ocultas	81
Tabla 15. Tiempo consumido en pruebas para configuraciones de 1 y 2 capas ocultas	82
Tabla 16. Prueba para número de capas según error cuadrático medio	82
Tabla 17. Parámetros más importantes de las redes entrenadas	83
Tabla 18. Tiempo requerido en pruebas para configuraciones de 1 y 2 capas ocultas	85
Tabla 19. Parámetros de las redes para las pruebas de desempeño	86
Tabla 20. Errores de entrenamiento para la configuración definitiva de la red	90
Tabla 21. Casos de uso: Cargar información	99

Tabla 22. Casos de uso: Entrenar RNA	100
Tabla 23. Casos de uso: Simular RNA	101
Tabla 24. Casos de uso: Producción acumulada	102
Tabla 25. Casos de uso: Tasa de declinación	103
Tabla 26. Casos de uso: Análisis económico	103
Tabla 27. Miembros principales de la clase cCapa	105
Tabla 28. Miembros principales de la clase cDatos	106
Tabla 29. Métodos principales de la clase cDatos	106
Tabla 30. Miembros principales de la clase cRed	107
Tabla 31. Métodos principales de la clase cRed	107
Tabla 32. Miembros principales de la clase cGráfica	108
Tabla 33. Métodos principales de la clase cGraficas	109
Tabla 34. Error absoluto promedio y tasa de ajuste de cada método para el pozo 1	114
Tabla 35. Error absoluto promedio y tasa de ajuste de cada método para el pozo 2	117
Tabla 36. Error absoluto promedio y tasa de ajuste de cada método para el pozo 3	119
Tabla 37. Error absoluto promedio y tasa de ajuste de cada método para el pozo 4	122
Tabla 38. Error absoluto promedio y tasa de ajuste de cada método para el pozo 5	124
Tabla 39. Error absoluto promedio y tasa de ajuste de cada método para el pozo 6	127
Tabla 40. Error absoluto promedio y tasa de ajuste de cada método para el pozo 7	129
Tabla 41. Codificación de los tipos de bombeo	152
Tabla 42. Datos de producción del pozo de prueba	164

TITULO: METODOLOGÍA PARA LA PREDICCIÓN DE CURVAS DE DECLINACIÓN DE POZOS DE PETRÓLEO, APLICANDO REDES NEURONALES ARTIFICIALES*

AUTOR(ES): CARCAMO TROCONIS, EMILIO JUSTINIANO**
GUIDO JOSÉ POLO NULE**

PALABRAS CLAVES: REDES NEURONALES ARTIFICIALES, CURVAS DE DECLINACIÓN, PREDICCIÓN DE LA PRODUCCIÓN DE PETRÓLEO.

DESCRIPCION

En este trabajo, se desarrolla una herramienta software que aborda el análisis de curvas de declinación de pozos de petróleo, basándose en redes neuronales, con el objetivo de mejorar la precisión en los resultados, respecto a otros métodos.

Para la construcción del modelo neuronal se utilizaron los registros históricos de producción de cada pozo, para alimentarlo durante las fases de entrenamiento y simulación. Para este propósito se empleó una configuración de red neuronal tipo perceptrón multicapa, con backpropagation como algoritmo de aprendizaje. Los resultados obtenidos son comparados contra los modelos representativos para la predicción de la producción: exponencial, armónico e hiperbólico, que son los más utilizados por los ingenieros de producción.

La herramienta desarrollada permite hacer una predicción más cercana a los datos reales, incluso en aquellos casos donde se realizan operaciones de mantenimiento o mejoramiento de la producción, como los cambios de bombeo, y surge como alternativa de solución ante los métodos tradicionales.

En este trabajo se evidenció que de un buen análisis, del correcto planteamiento y de eficientes técnicas de preprocesamiento de datos, depende el éxito del modelo neuronal, aunque muchas veces contar con un buen modelo no siempre es garantía para obtener buenos resultados en la predicción de la producción; pues las condiciones de operación de pozos depende en gran medida de las decisiones tomadas por los ingenieros de producción; es por esto que la aplicación no podría predecir y ajustarse a malos procedimientos que conlleven a un bajo rendimiento del pozo o a un daño irreparable.

* Trabajo de Grado

** Facultad de Ingeniería Físico - Mecánicas, Escuela de Ingeniería de Sistemas e Informática, ing Juan Carlos Reyes Figueroa

TITLE: METHODOLOGY FOR THE PREDICTION OF DECLINE CURVES USING ARTIFICIAL NEURAL NETWORKS *

AUTHOR(S): CARCAMO TROCONIS, EMILIO JUSTINIANO**
POLO NULE, GUIDO JOSÉ

KEY WORDS: ARTIFICIAL NEURAL NETWORKS, DECLINE CURVES ANALYSIS, PREDICTION OF THE OIL PRODUCTION.

ABSTRACT

This work concerns the development of a software package for the decline curves analysis in oil wells. The software is based in neural network algorithms for they provide a larger accuracy when compared with other available methods.

The model used to train the neural network was built from a historical compilation of the oil well production reports. The neural network was configured with a multilayer perceptron. Backpropagation was used as the learning algorithm. In order to compare the quality of the predictions obtained with the proposed model the predicted results from classical production models, such as exponential, harmonic and hyperbolic ones, were considered.

The software developed allows a more accurate prediction of real data, being capable to provide reasonable estimates in cases were maintenance operations, or production improvements such as pumping changes occur. Therefore, the software constitutes a real improvement for this field of knowledge respecting classical prediction methods.

In this work was evident that a good analysis, the correct approach and efficient techniques of data processing, depends on the success of neural model, although often have a good model is not always collateral to obtain good results in the prediction of production; because the operation conditions of wells depends largely on the decisions taken by the production's engineers; is why the application could not predict and conform to ill procedures that leading to a low performance of the pit or to irreparable harm.

* Work of degree.

** Faculty of Physical-Mechanics Engineering's, Systems and Informatic Engineering, Ing. Juan Carlos Reyes Figueroa

CAPITULO 1

INTRODUCCIÓN

La importancia de la industria del petróleo en términos económicos, sociales y ambientales; la diversidad y complejidad de sus procesos; las acciones y políticas gubernamentales, directamente, relacionadas con la producción de crudo; constituyen elementos fundamentales que dan lugar a un interesante campo para la investigación y al desarrollo tecnológico.

Debido a la naturaleza de los yacimientos petroleros, donde la materia prima se extrae de las profundidades, el grado de incertidumbre presente en las mediciones, los procesos y las actividades, es elevado. Por ello, con frecuencia se acude a técnicas basadas tanto en modelos estadísticos como en modelos empíricos; con el fin de obtener mayor nivel precisión en la representación de los fenómenos.

El factor más importante en el sector petrolero y que se presume, como el objeto de la mayor parte de las investigaciones, es la producción de crudo, que en últimas representa el objetivo principal de este tipo de industria; dicha producción presenta una tendencia descendente a lo largo de la vida del pozo, motivo que obliga el conocimiento de las razones que inciden en la tasa de disminución, con el fin de obtener un estimativo de la producción futura y establecer en que punto se hace necesaria la implementación de técnicas de recobro para mejorar el volumen de producción.

Una de las técnicas utilizadas hoy para la predicción de la producción de petróleo es el análisis de curvas de declinación pero, debido a su carácter estático, distan notoriamente del comportamiento real. Debido a lo anterior, el

propósito de este proyecto es plantear una alternativa que brinde resultados más precisos, mediante la implementación de una herramienta computacional, basada en Redes Neuronales Artificiales.

1.1 CONTEXTO GENERAL

1.1.1 DEFINICIÓN DEL PROBLEMA

El petróleo, desde su descubrimiento en 1859 como fuente de energía, se ha convertido en una pieza fundamental para el desarrollo de la humanidad. Según la Teoría del pico de Gubert, para el 2007 se ha predicho un declive en la producción petrolera, por esta razón se hace necesaria la optimización de los procesos de perforación, extracción, transporte y refinamiento. Para lograr este propósito se integraron múltiples disciplinas, en las que se encuentra La Ingeniería de Sistemas, que colabora con la elaboración de software, entre los que sobresalen los de simulación que permiten un mayor aprovechamiento de los recursos, basados en las propiedades de los yacimientos.

La mayoría de los campos maduros de petróleo en la actualidad, que han producido crudo durante varios años, comenzaron su producción en un tiempo donde la caracterización de yacimientos¹ no era una prioridad, de ahí la carencia de datos del mismo. Para revitalizar esos campos, en un tiempo donde el precio de los hidrocarburos es alto, era necesaria la implementación de nuevas técnicas, que ayudaran a mejorar la producción del campo o encontrar nuevos lugares con una potencial producción económica.

Es entonces cuando surge el análisis de las curvas de declinación como el primer paso de un proceso denominado "producción inteligente de análisis de datos" (IPDA)², donde se analiza La tasa de producción de crudo en función del tiempo, con el fin de evaluar el comportamiento predictivo de pozos y yacimientos. Cuando no existen cambios mayores en los procedimientos de operación del campo, las curvas de declinación muestran una tendencia

¹ Caracterización de yacimientos: Es el proceso de medición de los parámetros de un yacimiento, como la extensión de área, porosidad, permeabilidad, entre otros.

² S.D. Mohaghegh, R. Gaskari. And Jalali, West Virginia U: "A New Method for Production Data Analysis To Identify New Opportunities in Mature Fields, Methodology and Application", Paper SPE 98010, (September / 2005)

descendente por períodos extendidos; esto permite la extrapolación de la curva, con el fin de efectuar pronósticos de producción e identificar en que punto es alcanzado el límite económico.

Varios métodos pueden ser usados para estimar las reservas y predecir la producción en los yacimientos, las técnicas de simulación numérica facilitan dicha predicción, pero también pueden fallar por la gran incertidumbre y complejidad de los datos característicos de los yacimientos.

Otro método utilizado, frecuentemente, es el análisis de curvas de declinación, sin embargo, los modelos existentes de dicho método son heurísticos y se basan en las ecuaciones empíricas de Arps: exponencial, hiperbólica y armónica, estas ecuaciones presentan inconvenientes como: la curva de declinación exponencial tiende a subestimar las reservas y las tasas de producción y la armónica tiene la tendencia de sobreestimar el rendimiento del yacimiento.

Lo anterior incentivó la búsqueda de soluciones que implementaran nuevas áreas de conocimiento como la Inteligencia Artificial, que permite generar diagnósticos y predicciones con resultados cercanos a la realidad ante situaciones con alto grado de incertidumbre, comunes en la industria petrolera.

1.1.2 OBJETIVOS

1.1.2.1 Objetivo General

Desarrollar una herramienta computacional basada en Redes Neuronales Artificiales para el análisis de curvas de declinación de pozos de petróleo.

1.1.2.2 Objetivos Específicos

- i. Realizar el estado del arte sobre los modelos utilizados para la predicción de curvas de declinación de pozos petroleros.
- ii. Diseñar, construir y documentar los componentes del software con la notación estándar de UML y la metodología de prototipado.
- iii. Desarrollar una herramienta computacional, basada en Redes Neuronales Artificiales, que permita:
 - Construir la curva de declinación de un pozo petrolero.
 - Determinar la tasa de declinación de un pozo petrolero.
 - Predecir el comportamiento futuro de la producción del pozo.
- iv. Validar la aplicación desarrollada contra técnicas clásicas utilizadas como: la declinación armónica, hiperbólica, y exponencial, empleando datos de campo presentes en la literatura.

1.1.3 JUSTIFICACIÓN

Los resultados del proyecto podrán servir de soporte en la determinación de decisiones en cuanto al tiempo óptimo, en los meses venideros de producción del pozo, donde se deben implementar estrategias que contribuyan al incremento de la producción. Esto, junto a la necesidad de optimizar los procesos de extracción, transporte y procesamiento del crudo se convierte en una ayuda eficaz al progreso de países, que como Colombia, dependen en parte de la exportación de petróleo.

La industria petrolera, para disminuir gastos, mejorar sus productos y optimizar sus procesos, está necesitando cada vez más la participación de las tecnologías existentes. Por medio de este trabajo, se contribuye, en el ámbito académico local, con alternativas, proporcionando información valiosa a las próximas investigaciones.

Las Redes Neuronales Artificiales presentan un gran potencial en la generación de análisis y resultados de bases de datos históricas, y por su facilidad de implementación en la tecnología actual, están muy difundidas como parte de sistemas de análisis y de predicción. La tendencia parece ser la integración de técnicas emergentes, como la inteligencia artificial, con herramientas convencionales, como análisis estadísticos, para diseñar sistemas que puedan resolver problemas complejos.

CAPITULO 2

MARCO TEÓRICO

El fundamento teórico incluye la conceptualización necesaria para el desarrollo del presente trabajo de investigación. Incluye conceptos básicos de las Redes Neuronales Artificiales y curvas de declinación de petróleo.

2.1 REDES NEURONALES ARTIFICIALES

Los fundamentos y principios básicos de las redes neuronales artificiales (RNA) así como sus distintos modelos, que se han ido desarrollando a lo largo de su historia, surgen a partir del estudio del cerebro humano y del deseo de imitar su estructura y su comportamiento a través de la computación. Lo que se trata de conseguir es simular, a través de un ordenador, las facultades de pensamiento y raciocinio del cerebro en la resolución de problemas habituales relacionados con el reconocimiento de formas o patrones, predicción, clasificación y optimización, entre otros.

Sus múltiples aplicaciones y su probada eficacia para el tratamiento de grandes cantidades de información, han llevado a las RNA a convertirse en objeto de estudio de diversas disciplinas. Su capacidad de aprendizaje y generalización a partir de la experiencia asimilada a lo largo de su proceso de entrenamiento, les permite resolver problemas complicados para otras técnicas más convencionales, bien porque dichos problemas no poseen un algoritmo específico para su solución, o los existentes son complejos. Su habilidad para detectar y extraer patrones de comportamiento al presentar a la RNA determinadas series de datos, las convierte en una herramienta de predicción aplicable a cualquier campo.

2.1.1 Panorama Histórico

La investigación de las RNA se remonta a los trabajos realizados por McCulloch y Pitts³, quienes en 1943, desarrollaron el primer modelo artificial, que consistía en un conjunto de neuronas que poseían dos estados discretos; tenían la limitación de que sólo permitían computar funciones booleanas.

En 1957, Rosenblatt⁴, publica los resultados de un ambicioso proyecto de investigación, denominado el Perceptrón, probó que dada clases linealmente separables, un perceptrón podría, en un número finito de ensayos de entrenamiento, desarrollar un modelo de clasificación. Aproximadamente para la misma época, Widrow⁵ desarrolló una red similar, llamada adaline (Adaptative Linear Neuron), éste modelo se usó en la vida cotidiana como filtros adaptativos para eliminar ecos en las líneas telefónicas con resultados extraordinarios.

En 1969, Minsky y Paper⁶ señalaron que el teorema del Perceptrón se aplica a problemas de naturaleza lineal, pero que los cálculos elementales de problemas como un XOR, no pueden ser resueltos por capas simples de perceptrones.

Rosenblatt también estudió estructuras con más capas y creyó que podrían superar las limitaciones de los perceptrones simples, Sin embargo, ningún algoritmo de aprendizaje conocido podía determinar los pesos necesarios para implementar un cálculo dado; Minsky y Papert dudaron que se pudiera encontrar uno y recomendaron que otros modelos de inteligencia artificial

³ McCulloch, WS and Pitts, W: "A Logical Calculus of Ideas Immanent in Nervous Activity." , Bull. Mathematical Biophysics (1943) 5,15

⁴ Rosenblatt, F: "The Perceptrón: Probabilistic Model for Information Storage an Organization in the brain," Psychol. Rev. (1958) 65, 386

⁵ Widrow, B.: "Generalization and Information Storage in Networks of Adeline Neurons," Self-Organizing Systems, M.C. Yovitz, Jacobi, and G.D. Goldstein (eds.), Spartan Books, Washintong, DC

⁶ Minsky, M.L. and Papert, S.A.: Perceptrons, MIT Press, Cambridge, Massachusetts (1969)

deberían ser propuestos. Por estas investigaciones, el paradigma de las redes neuronales fue abandonado por 20 años hasta que alrededor de 1980, renació el interés por el campo conexionista gracias al trabajo del grupo Parallel Distributed Processing (PDP) y las aportaciones de Hopfield⁷. Sus esfuerzos coincidieron con el desarrollo de nuevos algoritmos de aprendizaje, como el backpropagation. El crecimiento de la investigación de las redes neuronales y sus aplicaciones ha sido fenomenal desde esta investigación.

2.1.2 Ventajas de las RNA

El modelo de una red neuronal artificial, por ser una imitación del proceso de una neurona biológica, presenta un gran número de características semejantes a las del cerebro y ofrece numerosas ventajas, entre ellas:

- *Aprendizaje adaptativo*: Las redes neuronales tienen la capacidad de adaptarse a los cambios en las condiciones de su entorno, debido a que pueden aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial.
- *No linealidad*: Las redes neuronales pueden ser lineales o no lineales, dicha característica les permite deducir relaciones complejas entre datos de entrada y salida.
- *Tolerancia a fallos*: Una red puede soportar daños parciales en su estructura, sin afectar de forma considerable su desempeño general. algunas capacidades de la red se pueden retener, incluso sufriendo un gran daño.
- *Paralelismo*: Las redes neuronales procesan en paralelo la información suministrada, operando en tiempo real; esto hace de ellas sistemas rápidos, robustos y confiables.

⁷ Hopfield, J.J.: "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," Proc., Natl. Academy of science (1982) 79, 2554

- *Capacidad de generalización:* Las redes neuronales son capaces de producir salidas acertadas ante entradas no suministradas durante el entrenamiento.

2.1.3 Fundamentos de las RNA

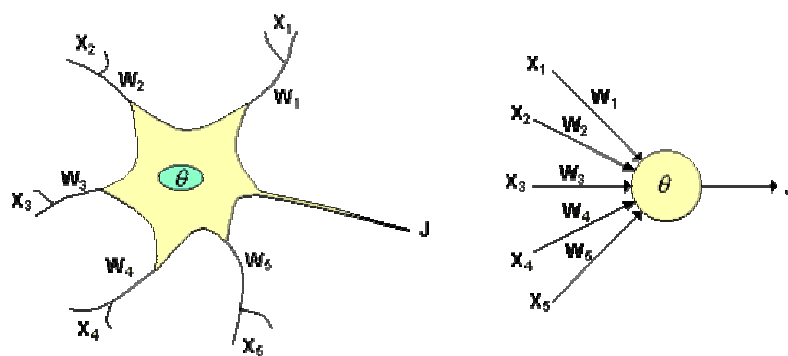
Los modelos aparecen en la biología, como una representación de la característica natural de acumulación de información visual y aprendizaje que tienen los seres vivos. Sin embargo, el proceso por el cual a partir de esta información se genera una respuesta en un sujeto es complicado y hasta cierto punto, es aún desconocido. En general, este proceso puede describirse como uno en el cual, un número determinado de células (neuronas), se organizan a través de una red e interactúan hasta obtener una respuesta a determinados estímulos. Por ejemplo, una persona recibe una serie de información que es transmitida a través de impulsos a sus células cerebrales, en el cerebro, estos impulsos se entremezclan y generan una respuesta.

El cerebro humano consta de un gran número (aproximadamente 10^{11}) de elementos altamente interconectados (aproximadamente 10^4 conexiones por elemento), llamados neuronas. Estas neuronas tienen tres componentes principales, las dendritas, el cuerpo de la célula o soma, y el axón. Las dendritas, son el árbol receptor de la red, son como fibras nerviosas que cargan de señales eléctricas el cuerpo de la célula; el cuerpo de la célula, realiza la suma de esas señales de entrada y el axón es una fibra larga que lleva la señal desde el cuerpo de la célula hacia otras neuronas. El punto de contacto entre un axón de una célula y una dendrita de otra célula es llamado sinápsis, la intensidad de la sinápsis es determinada por la complejidad del proceso químico que estabiliza la función de la red neuronal.

2.1.3.1 Elementos de una RNA. Los modelos neuronales asumen muchas simplificaciones del modelo biológico con el fin de facilitar su desarrollo matemático. En la Figura 1, se compara una neurona biológica con el modelo matemático equivalente. Dentro de estas similitudes se pueden destacar:

- Las entradas X_i representan las señales que provienen de otras neuronas y que son capturadas por las dendritas.
- Los pesos W_i son la intensidad de la sinápsis que conecta dos neuronas; tanto X_i como W_i son valores reales.
- θ es la función umbral que la neurona debe sobrepasar para activarse; este proceso ocurre biológicamente en el cuerpo de la célula.

Figura 1. Del modelo biológico al modelo artificial⁸



Las señales de entrada a una neurona artificial X_1, X_2, \dots, X_n , pasan a través de una ganancia o peso, llamado peso sináptico o fortaleza de la conexión (W_i), cuya función es análoga a la de la función sináptica de la neurona biológica; el nodo sumatorio acumula todas las señales de entradas multiplicadas por los pesos, este resultado es evaluado por medio de una función de activación que a su vez pasa su resultado a la salida, a través de una función umbral o función de transferencia (θ).

⁸ Fuente: Universidad Tecnológica de Pereira, Tutorial Redes Neuronales, 2007. Available from internet: <URL: <http://ohm.utp.edu.co/neuronales/main.htm> >

2.1.3.2 Estado de activación. Todas las neuronas que componen la RNA se encuentran en cierto estado de activación: reposo o excitado, y a cada uno de ellos, se le asigna un valor que pueden ser continuo o discreto. Si son discretos, suelen tomar valores binarios o un conjunto pequeño, en cambio si son continuos, se considera un conjunto continuo de estados de activación, en lugar de solo dos estados, en cuyo caso se le asigna un valor entre $[0,1]$ o en el intervalo $[-1,1]$, generalmente, siguiendo la función sigmoideal.

En notación binaria, un estado activo se indicaría por 1, y se caracteriza por la emisión de un impulso por parte de la neurona (potencial de acción), mientras que un estado pasivo se indicaría por 0, y significaría que la neurona está en reposo.

2.1.3.3 Función de Salida o de Transferencia. Entre las unidades o neuronas que forman una RNA existe un conjunto de conexiones que unen unas con otras. Cada unidad transmite señales a aquellas que estén conectadas con su salida. Asociada con cada unidad U_i hay una función de salida $f_i(a_i(t))$, que transforma el estado actual de activación $a_i(t)$ en una señal de salida $y_i(t)$; es decir:

$$y_i(t) = f_i(a_i(t)) \quad (2.1)$$

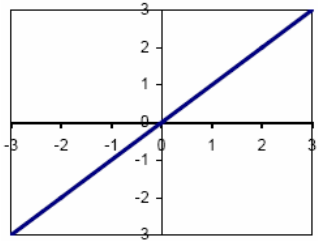
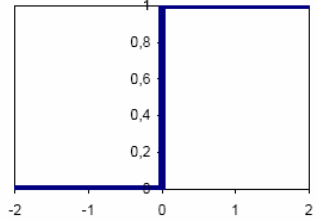
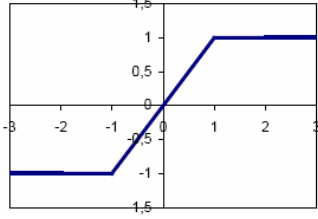
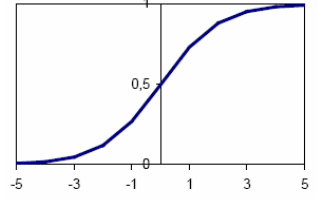
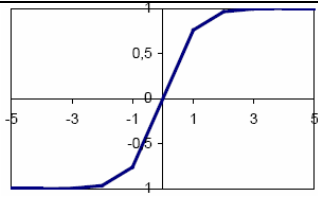
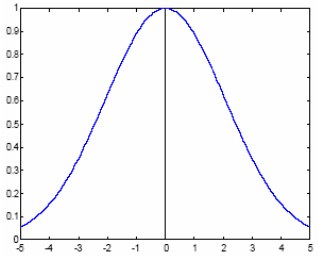
Existen cuatro funciones de transferencia típicas que determinan los distintos tipos de neuronas:

- *Función escalón o paso:* Se asocia a neuronas binarias en las cuales cuando la suma de las entradas es mayor o igual que el umbral de la neurona, la activación es 1, si es menor, la activación es 0 (ó -1).
- *Función lineal y mixta:* Corresponde a la función $F(x) = x$. En las neuronas con función mixta si la suma de las señales de entrada es

menor que un límite inferior, la activación se define como 0 (ó -1). Si dicha suma es mayor o igual que el límite superior, entonces la activación es 1. Si la suma de entrada está comprendida entre ambos límites, la activación se define como una función lineal de suma de las señales de entrada. La función de activación mixta se puede representar tomando el límite superior de la suma de todas las entradas de activación que afectan a la neurona durante el ciclo de operación con el valor de c y el límite inferior como $-c$, y es la salida de activación de la neurona.

- *Función continua:* Corresponde a cualquier función definida en un intervalo de posibles valores de entrada, con un incremento monotónico y que tengan ambos límites superiores e inferiores. Generalmente se escoge la función sigmoideal y arco tangente.
- *Función gaussiana:* Los centros y el ancho de estas funciones pueden ser adaptados, hecho que las hace más adaptativas que las funciones sigmoideales

Tabla 1 Principales Funciones de Transferencia.

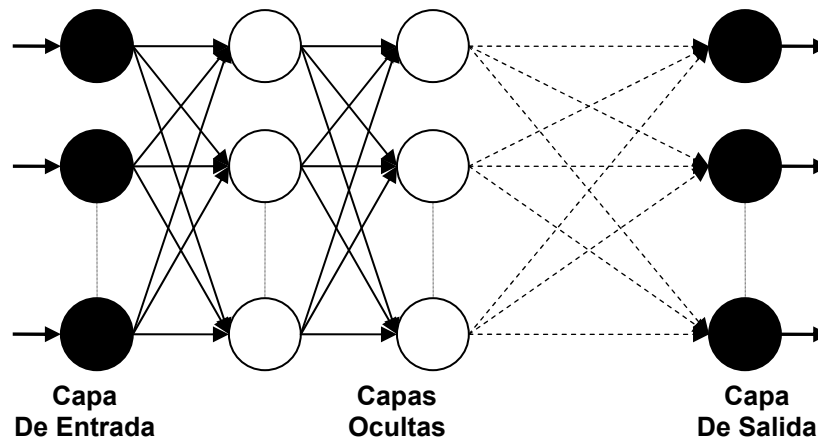
Nombre	Función	Grafica
Lineal	$f(x) = x$	
Escalón o Paso	$F(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$	
Mixta o Rampa	$f(x) = \begin{cases} -1 & \text{si } x < -c \\ 1 & \text{si } x > c \\ a.x & \text{en otro caso} \end{cases}$	
Logística	$F(x) = \frac{1}{1 + e^{-x}}$	
Hiperbólica	$F(x) = \text{Tanh}(x)$	
Gaussiana	$F(x) = e^{-x^2 / \sigma^2}$	

2.1.3.4 Arquitectura de Red. El término arquitectura de red se refiere a la forma o estructura de una RNA, y está compuesta por las capas de entrada, ocultas y de salida; así mismo, por la interconexión y dirección de la red, que se refiere a la forma en que se interrelacionan las neuronas.

- **Capa de Entrada:** Es la capa que recibe directamente la información proveniente de las fuentes externas de la RNA.
- **Capas ocultas:** También se les denomina capas intermedias, por encontrarse entre la entrada y la salida de la red. Sus elementos pueden tener diferentes conexiones y son estas las que determinan las diferentes topologías de la red.
- **Capa de salida:** Recibe la información de la capa oculta y transmite la respuesta al medio externo.
- **Interconexión:** Las interconexiones entre capas de las redes pueden clasificarse como:
 - ✓ Totalmente conectadas: si todas las salidas de un nivel llegan a todos y cada uno de los nodos del nivel siguiente.
 - ✓ Localmente conectadas: la salida de una neurona del un nivel es la entrada a una región de neuronas del nivel siguiente.
- **Dirección:** La dirección de la información de las redes pueden clasificarse en:
 - ✓ Redes de *propagación hacia delante*: las salidas de las neuronas de una capa sólo se propagan a las neuronas de la capa siguiente. Es decir, la información fluye solamente de la entrada a la salida.
 - ✓ Redes retroalimentadas: las salidas de las neuronas de una capa pueden ser entradas de las neuronas de las capas anteriores.
 - ✓ Redes de alimentación lateral: las salidas de las neuronas pueden ser entradas de neuronas de la misma capa.

- ✓ Redes recurrentes: Son aquellas redes retroalimentadas que poseen lazos cerrados, es decir, existen conexiones de una neurona con ella misma, conexiones entre neuronas de una misma capa o conexiones de las neuronas de una capa a la capa anterior.

Figura 2 Estructura de una Red Neuronal



2.1.3.5 Mecanismos de Aprendizaje. El entrenamiento de una red se lleva a cabo mediante una regla o algoritmo de aprendizaje. Este algoritmo es un procedimiento para modificar los pesos de una red y su propósito es entrenar la red para ejecutar alguna tarea⁹.

Según el algoritmo de aprendizaje, las RNA se dividen en redes de entrenamiento supervisado y no supervisado (Fig. 3):

- Supervisados: La misión de este algoritmo es ajustar los pesos de la red de manera tal, que dado un conjunto de entradas, las salidas proporcionadas por la red deberán coincidir lo más posible con las salidas especificadas en

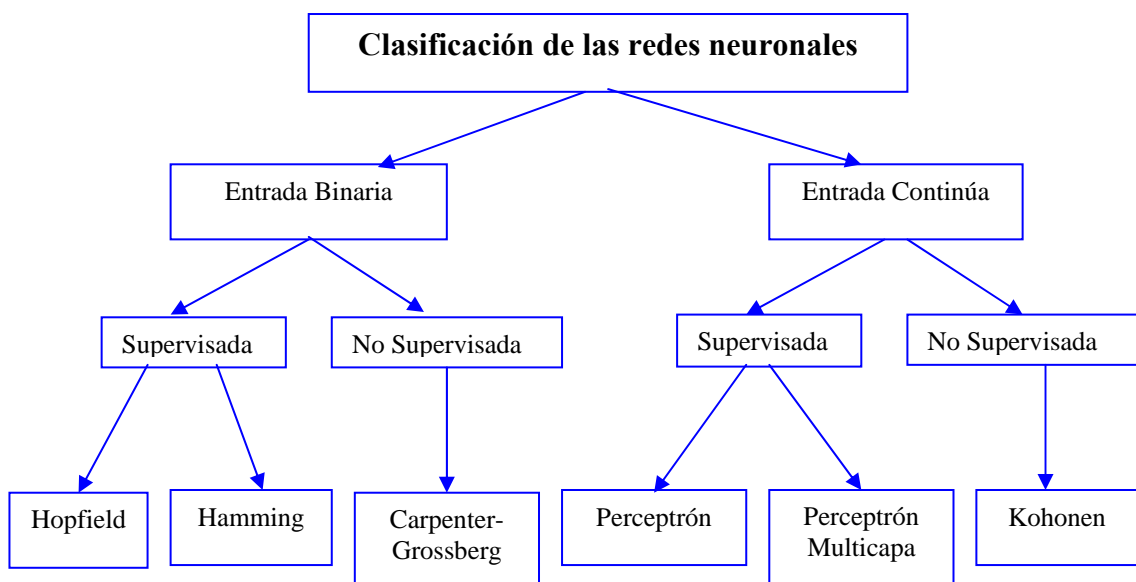
⁹ Hagan, Dermuth y Beale: Neural network design. Boston, USA: PWS Publishing Company, 1996.

el patrón de entrenamiento¹⁰. A este tipo de algoritmo pertenecen las redes Perceptrón, Hamming, Hopfield y perceptrón multicapa.

- No Supervisados: los pesos son modificados en respuesta a las entradas de la RNA, es decir no están disponibles las salidas. La mayoría de estos algoritmos tienen como misión realizar algún tipo de agrupamiento, los cuales, aprenden a categorizar los patrones de entrada en un número finito de clases. A este tipo de algoritmo pertenecen las redes Carpenter/Grossberg y Kohonen.

El aprendizaje supervisado, se destaca por poseer diferentes maneras de realizarse, entre ellas, el aprendizaje por corrección de error, que consiste en ajustar los pesos de las conexiones de La RNA en función del error cometido en la salida. Dentro de los algoritmos que utilizan el aprendizaje por corrección del error, se encuentran: la regla de aprendizaje del perceptrón, regla del mínimo error cuadrado o regla de Widrow-Hoff y la regla Delta generalizada o algoritmo de retro-propagación del error (error backpropagation), este último, nuestro caso de estudio.

Figura 3 Clasificación de los tipos de redes



¹⁰ Colina, E. y Rivas, F: Introducción a la inteligencia artificial. Mérida, Venezuela: Universidad de Los Andes.

2.1.4 Algoritmo Backpropagation

En 1986, Rumelhart, Hinton y Williams, formalizaron un método para que una RNA aprendiera la asociación que existe entre los patrones de entrada a la misma y las clases correspondientes, utilizando varios niveles de neuronas. Este método, conocido como Backpropagation (retropropagación del error) o método del gradiente descendente, está basado en la generalización de la regla delta, que actualiza los pesos de forma proporcional a la diferencia entre la salida deseada y la obtenida.

El algoritmo de backpropagation emplea un ciclo propagación – adaptación de dos fases: una vez que se ha aplicado un patrón a la entrada de la red como estímulo, este se propaga desde la primera capa a través de las capas superiores de la red, hasta generar una salida. La señal de salida se compara con la salida deseada y se calcula una señal de error para cada una de las salidas, luego las salidas de error se propagan hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa oculta que contribuyen directamente a la salida.

La importancia de este proceso consiste en que, a medida que se entrena la red, las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada.

2.1.4.1 Funcionamiento del Algoritmo. El esquema del algoritmo backpropagation puede resumirse en los siguientes pasos:

1. Inicializar los pesos de la red con valores aleatorios, o de acuerdo a alguna función de inicialización.
2. Presentar un patrón de entrada $X_p: (x_{p1}, x_{p2}, \dots, x_{pN})$ y especificar la salida deseada que debe generar la red: (d_1, d_2, \dots, d_M) .

3. Calcular la salida actual de la red, para ello, para la entrada presentada, se van obteniendo los valores de las respuestas que presenta cada capa hasta llegar a la capa de salida. Para ello se sigue el siguiente procedimiento:

- Se calculan las entradas netas (*net*) para las neuronas ocultas procedentes de las neuronas de entrada. Para una neurona oculta j (O_j):

$$net_{pj}^h = \sum_{i=0}^N w_{ji}^h x_{pi} + \theta_j^h \quad (2.2)$$

Donde el subíndice p corresponde al p -ésimo vector de entrenamiento, j a la j -ésima neurona oculta, w_{ji} es el peso de la conexión entre E_i y O_j y el término θ_j corresponde a un término de umbral mínimo a alcanzar por la neurona para su activación. A partir de estas entradas se calculan las salidas (y) de las neuronas ocultas, utilizando una función de activación f .

$$y_{pj}^h = f_j^h \left(net_{pj}^h \right) \quad (2.3)$$

Se realizan los mismos cálculos para obtener los valores de resultado de cada neurona k de la capa de salida.

$$net_{pk}^o = \sum_{j=1}^L w_{kj}^o y_{pj}^h + \theta_k^o \quad (2.4)$$

4. Después que todas las neuronas de la red tienen un valor de activación asociado para un patrón de entrada dado, el algoritmo continúa encontrando el error que se presenta para cada neurona, excepto para las de la capa de entrada. Para la neurona k de la capa de salida, si la respuesta es (y_1, y_2, \dots, y_M), dicho error (δ) se puede escribir como:

$$\delta_{pk} = (d_{pk} - y_{pk}) f'_k \left(net_{pk} \right) \quad (2.5)$$

5. Para la actualización de los pesos utilizamos el algoritmo recursivo, comenzando por las neuronas de salida y trabajando hacia atrás hasta llegar a la capa de entrada, ajustando los pesos de la forma siguiente:

- Para los pesos de las neuronas de la capa de salida:

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj}^o(t+1) \quad (2.6)$$

$$\Delta w_{kj}^o(t+1) = \alpha \delta_{pk} y_{pj}^h \quad (2.7)$$

- Para los pesos de las neuronas de la capa oculta:

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}^h(t+1) \quad (2.8)$$

$$\Delta w_{ji}^h(t+1) = \alpha \delta_{pj}^h y_{pi} \quad (2.9)$$

En ambos casos, para acelerar el proceso de aprendizaje, se incluye una tasa de aprendizaje que varía entre 0 y 1, dependiendo del problema a solucionar.

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.

6. El proceso se repite hasta que el error medio cuadrático E_p resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

$$E_p = \frac{1}{2} \sum_{k=1}^M (\delta_{pk})^2 \quad (2.10)$$

Este estadístico se utiliza para la evaluación del desempeño de la red.

2.1.4.2 Algoritmo de Levenberg-Marquardt. El algoritmo de Levenberg-Marquardt (LM) es un algoritmo iterativo de optimización en el que el método de iteración presenta una ligera modificación sobre la técnica tradicional de Newton, que trata de minimizar la distancia euclidiana entre los datos de campo y los teóricos:

$$F(W^k) = \sum_{k=1}^x (E^k)^2 \quad (2.11)$$

Donde:

F(W^k) = Función a optimizar.

k = Identifica el número de salida de la red que se relaciona con el proceso iterativo.

X= Número de elementos del conjunto de entrenamiento.

E= [e₁, e₂, e₃,..., e_n]^T= Corresponde al vector de errores de entrenamiento, entre la salida deseada y la obtenida.

W^k= Vector de parámetros de la red , definido por los pesos y ganancias.

La regla de actualización está definida por:

$$W^{k+1} = W^k + \Delta W^k \quad (2.12)$$

Δ W^k=actualiza los parámetros indicando hacia adonde debe disminuir el error, y está dado por la siguiente ecuación:

$$\Delta W^k = -[J(W^k)^T J(W^k) + \beta^k I]^{-1} J(W^k)^T E^k \quad (2.13)$$

I = matriz identidad unitaria.

$J(W^k)$ = Matriz jacobiana, expresada como:

$$J(W^k) = \begin{bmatrix} \frac{\partial e_1}{\partial w_1} & \frac{\partial e_1}{\partial w_2} & \dots & \frac{\partial e_1}{\partial w_n} \\ \frac{\partial e_2}{\partial w_1} & \frac{\partial e_2}{\partial w_2} & \dots & \frac{\partial e_2}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_n}{\partial w_1} & \frac{\partial e_n}{\partial w_2} & \dots & \frac{\partial e_n}{\partial w_n} \end{bmatrix} \quad (2.14)$$

β = es el factor de ajuste, en redes neuronales se conoce como factor de entrenamiento.

El valor de β es actualizado por la rata de decaimiento λ ($0 < \lambda < 1$). dependiendo del resultado. En particular, β es multiplicada por λ cuando $F(W^k)$ disminuye, y se divide por λ cuando en una nueva iteración el valor de $F(W^k)$ aumenta.

Para acelerar el entrenamiento del Levenber- Marquardt, se utilizó una rata de decaimiento variable¹¹ definida por:

$$\lambda = \frac{\text{Log}(F) - \text{Log}(F_{\min})}{\text{Log}(F_0) - \text{Log}(F_{\min})} \times 0.8 + 0.1 \quad (2.15)$$

Donde F es la suma reducida del cuadrado de los errores, F_0 es el primer valor calculado de $F(w)$ basado en los pesos iniciales y F es el valor requerido de aproximación.

¹¹ Chen, Tai-cong , Han, Da-jian: "Acceleration of Levenberg-Marquardt Training of Neural Networks with Variable Decay Rate", IEEE Trans. on Neural Ner., vol. 9, no. 3, pp.1873-1878, 2003.

Con esta variación, el tiempo requerido para el entrenamiento de redes neuronales disminuye considerablemente respecto al algoritmo estándar L-M.

En general, la red con el algoritmo L-M converge en menos iteraciones que la mayoría de las variaciones del backpropagation, por supuesto requiere mucha más computación por iteración, debido a que implica el cálculo de matrices inversas de gran tamaño. A pesar de su gran esfuerzo computacional sigue siendo el algoritmo de entrenamiento más rápido para redes neuronales cuando se trabaja con un moderado número de parámetros en la red, si el número de parámetros es muy grande utilizarlo resulta poco práctico.

2.2 CURVAS DE DECLINACIÓN

2.2.1 Introducción

Los dos problemas básicos presentes en la evaluación de un pozo petrolero, son la determinación de su vida futura y la estimación de su producción. Algunas veces uno o ambos problemas pueden ser resueltos mediante cálculos hechos sobre las variables inherentes al yacimiento, tales como porosidad, presión entre otros; sin embargo, no siempre se tienen disponibles los datos suficientes para eliminar todas las incógnitas. En estos casos, la posibilidad de extrapolar la tendencia de alguna de las variables características del pozo productor puede ser de considerable ayuda. La variable característica, más simple y de mayor disponibilidad es la tasa de producción, y una forma lógica de encontrar una respuesta a los dos problemas mencionados, es mediante extrapolación, graficando esta tasa de producción, bien sea contra el tiempo o contra la producción acumulada, y luego extender la curva obtenida hasta el límite económico, esto es lo que llamamos una curva de declinación. El punto de intersección entre la curva extrapolada y el límite económico, indica la posible vida futura del pozo ó la recuperación estimada de crudo. Las bases que respaldan esta estimación, es que se asume que el comportamiento futuro del pozo está gobernado por alguna tendencia o relación matemática que depende altamente de su rendimiento pasado. Esta suposición hace que el

método de extrapolación conste de bases estrictamente empíricas, y esto se hace evidente cuando los resultados obtenidos por este método, algunas veces son inferiores comparados con aquellos que arrojan los métodos volumétricos que son mas precisos.

El análisis de curvas de declinación es uno de los métodos más usados para predecir la producción futura de un campo de petróleo o de gas, desde que Arps formalizó la técnica en 1945. Esta técnica tiene pocos fundamentos teóricos pero la simplicidad y el éxito en sus predicciones son responsables de su uso y aceptación general. Hay tres tipos de curvas de declinación: exponencial, armónica e hiperbólica; en realidad las de tipo exponencial y armónica son casos particulares de la hiperbólica, por lo tanto realmente existe un solo tipo general de curva de declinación, el cual involucra tres parámetros [q_i , d_i y n], que deben ser determinados a partir de la información medida. Cuando $n = 0$, se dice que la declinación es exponencial; cuando $n = 1$, la declinación es armónica. A pesar de que n usualmente toma valores entre 0 y 1, se han reportado casos en los que $n > 1$.

Existen técnicas gráficas simples para determinar los parámetros q_i y d_i en los casos especiales donde la declinación es armónica o exponencial; tales parámetros pueden obtenerse a partir de la pendiente, m y del intercepto, b , de las gráficas del logaritmo de la producción, q_t , contra el tiempo, t , o de la producción acumulada, Q_p , contra el tiempo, y son más precisos los resultados si se utiliza regresión lineal. En el caso de la declinación hiperbólica, los parámetros a estimar son tres, y es aquí donde radica su dificultad; sin embargo se han propuesto varias técnicas con el propósito de solventar esta situación.

El análisis de curvas de declinación sólo puede ser aplicado mientras las condiciones mecánicas y el drenaje permanezcan constantes en un pozo y que este produzca a su capacidad total. Estas limitaciones conducen muchas veces

a un uso indebido de esta técnica; sin embargo, Fetkovich¹² con la combinación de muchas aproximaciones muestra algunos ejemplos de curvas de declinación en donde cambian tanto la presión como el drenaje en el pozo. En la ingeniería de yacimientos es necesario estar dispuesto a usar todas las técnicas disponibles con tal de hacer predicciones acertadas sobre el yacimiento; por esto es que, en lugar de aislarlas y hacerlas mutuamente excluyentes, el ingeniero debe ser capaz de integrarlas para extraer la mejor información de cada una.

Cómo se mencionó, normalmente el análisis de las curvas de declinación consiste en graficar la producción del pozo contra el tiempo; esta se hace en un papel semilog y luego se trata de ajustar esta gráfica a una línea recta, la cual será extrapolada hacia el futuro. Las reservas se calculan con base en la rata promedio de producción anual hallada.

2.2.2 Estado del Arte

Arnold y Anderson¹³ en 1908 presentan un estudio sobre el comportamiento de la producción del campo Caolinga (USA), donde se dan cuenta de que la tasa de producción en algún momento es una fracción constante de las tasas precedentes, es decir, que la tasa de producción durante períodos de tiempo iguales forma una serie geométrica, lo que también implica que la caída de producción sobre un intervalo constante dado es una fracción fija o un porcentaje de la tasa de producción precedente. Esta caída de producción como fracción, expresada usualmente en porcentaje por mes, la denominan **declinación**. En sus trabajos de evaluación encuentran que la mayoría de curvas de declinación son de tipo exponencial, al menos sobre períodos de tiempo limitados.

¹² Fetkovich, M.J.: "Decline Curve Analysis Using Type Curves," JPT (June 1980) 1065-1077.

¹³ Arnold, R. and Anderson R.: "Preliminary Report on Coalinga Oil District," U.S. *Geological Survey Bull.* (1908) 357,79

En 1924, **W. W. Cutler**¹⁴ mostró después de una intensiva investigación acerca de las curvas de declinación de varios pozos, que la suposición del porcentaje de declinación constante y la extrapolación de una línea recta arrojaba resultados que se corroboraban en el futuro, es decir, el proceso de extrapolación era válido. Sin embargo, notó que las curvas de declinación de tipo hiperbólico representaban mejor el comportamiento futuro de la producción que las curvas de tipo exponencial o geométrico para la gran mayoría de los casos.

R. H. Johnson y A. L. Bollens¹⁵ en 1927 introducen un nuevo método estadístico para la extrapolación de curvas de declinación de pozos de petróleo llamado "loss ratio".

R. E. Allen¹⁶ en 1931 menciona cuatro tipos de declinación y las clasifica de acuerdo a una relación matemática simple. Los tipos de declinación fueron:

- Aritmética o declinación con decrementos constantes,
- Geométrica o declinación exponencial,
- Armónica
- Básica o declinación de potencia fraccional, que es de tipo hiperbólico.

S. J. Pirson¹⁷ en 1935 publicó una investigación sobre las bases matemáticas del método de la "Loss ratio" y encontró la relación tasa - tiempo para las curvas de declinación que tienen una "Loss ratio" constante. Las del primer tipo son idénticas a las de la declinación exponencial, las del segundo tipo fueron declinaciones de tipo hiperbólico y las del tercer tipo poseían relaciones matemáticas tan complejas que no eran útiles para propósitos prácticos.

¹⁴ Cutler, W. W. Jr.: "Estimation of Underground Oil Reserves by Well Production Curves," U.S. Bureau of Mines Bull. (1924) 228.

¹⁵ Johnson, R. H. and Bollens, A. L. : "The loss ratio method of extrapolating oil well decline curves". Trans AIME. Vol 27, p 771, 1927

¹⁶ Allen, R. E.: "Control of California oil curtailment". Trans AIME. Vol 47, p92, 1931

¹⁷ Pirson, S. J.: "Production Decline Curve of Oil Wells May Be Extrapolated by Loss Ratio Method," Oil & Gas J. (1935) 34, N° 46,34.

P. J. Jones¹⁸ en 1942 presenta una aproximación para los pozos que declinan a una tasa variable donde la relación tasa - tiempo sigue una línea recta en papel log-log y corresponde a la ecuación:

$$\text{Log}(D) = \text{Log}(D_0) - m \text{Log}(t) \quad (2.16)$$

Donde D_0 representa la declinación inicial y m es una constante positiva que corresponde a la pendiente de la línea recta.

En 1945, **J. J. Arps**¹⁹ publica un documento donde clasifica las curvas de declinación en cuatro tipos generales: (1) Exponencial, (2) Hiperbólica, (3) Armónica y (4) Razón de declinación. Las tres primeras son muy familiares, la cuarta rara vez se usa. En este documento presenta además las formulas matemáticas para los tres primeros tipos de declinación tanto para las relaciones tasa tiempo como para las relaciones tasa-producción acumulada.

En 1956, **Arps** publica su segundo documento²⁰ que no contiene ninguna nueva información sobre el análisis de curvas de declinación, pero presentó las fórmulas matemáticas de las relaciones tasa-tiempo y tasa-producción acumulada en una forma mucho más simple.

H. N. Mead²¹ en 1956 presenta una lista de-valores del exponente de declinación b para varios tipos de mecanismos de empuje. Además incluye las ecuaciones para el análisis de las curvas de declinación, las cuales están basadas en la premisa de que, la tasa de cambio del recíproco de la declinación para intervalos de tiempo sucesivos es constante cuando el yacimiento está produciendo bajo unas condiciones fijas.

¹⁸ Jones, P. J.: "Estimating oil reserves from production decline rates oil and gas journal". P 43, 1942

¹⁹ Arps, J. J.: "Analysis of Decline Curves." Trans., AIME (1945) 160,228

²⁰ Arps, J. J.: "Estimation of Primary Oil Reserves," Trans., AIME (1956) 207,182

²¹ Mead, H.: "Modifications of decline curve analysis". Journal of petroleum technology, pp 1065-1076. Jun 1980.

H. C. Slider²² en 1968 propuso un método mejorado para el análisis de curvas de declinación hiperbólicas por medio de curvas tipo, elaboradas por él.

En 1972, **R. W. Gentry**²³ presenta un análisis de curvas de declinación, en el cual utilizó las ecuaciones tasa-tiempo y tasa-producción acumulada presentadas por Arps para obtener dos ecuaciones adimensionales para cada tipo de declinación. Estas ecuaciones fueron usadas para construir dos gráficos o curvas tipo que pueden ser utilizados para extrapolar rápidamente las curvas de declinación hiperbólicas y armónicas, y obtener los parámetros de declinación b y D_i .

M. J. Fetkovitch²⁴ intentando encontrar las relaciones entre las constantes de declinación (q_i , D_i y b) y las características del yacimiento y los fluidos, presenta en 1973, un estudio en el cual utilizando ecuaciones de balance de materiales y ecuaciones de tasas de producción obtiene un grupo de correlaciones que le permiten construir un grupo de curvas tipo en términos de variables de declinación adimensionales.

R. W. Gentry y A. W. McCray²⁵ publican en 1978, un estudio donde investigan los efectos que tienen las propiedades de la formación y de los fluidos sobre la historia de producción de un pozo o un yacimiento. Obtienen expresiones tasa - tiempo y tasa – producción acumulada que describen el comportamiento de la declinación mejor que las relaciones de Arps.

²² Slider, H. C., "A Simplified Method of Hyperbolic Decline Curve Analysis," JPT (march/1968) 235.

²³ Gentry, R. W. "Decline Curve Analysis," JPT (January/1972) 38.

²⁴ Fetkovitch, M.: "Decline Curve Analysis using type curves". Journal of petroleum technology, pp 1065-1076. Jun 1980.

²⁵ Gentry, R. W. and McCray, A. W. "The Effect of Reservoir and Fluid Properties on Production Decline Curves," JPT (September/1978) 1327.

R. E. Lohec²⁶ en 1984 analiza la declinación de la producción para yacimientos agotados mediante un desplazamiento frontal. Desarrolla ecuaciones que relacionan la tasa de producción, el tiempo y las reservas con la geometría del yacimiento y la productividad.

En 1988, **D. R. Long y M. J. Davis**²⁷, presentan una aproximación eficiente y versátil para el análisis de curvas de declinación hiperbólicas mediante el uso de curvas tipo.

En 1993, **F. Rodriguez y H. L. Cinco**²⁸ [16] desarrollaron un modelo para propósitos de interpretación y pronóstico de la producción. En este trabajo estudiaron el rendimiento de la producción para yacimientos con múltiples pozos que producían a una presión constante y arbitraria.

M. J. Fetkovitch y E. J. Fetkovitch²⁹ en 1996, presentaron algunas pautas y conceptos fundamentales, a los ingenieros responsables de hacer los pronósticos y determinar reservas para un gran número de pozos activos e inactivos, con el fin de permitir que ellos hagan estos trabajos más rápida y acertadamente

En 1998 **R.G Agarwal y D.C Gardner**³⁰ introducen un método, que combina los conceptos de curvas de declinación y curvas tipo, para la estimación de reservas y otros parámetros del yacimiento, para pozos de gas y de aceite, usando los datos de producción.

²⁶ Lohec, R.: "Analytic approach evaluates frontal displacement mechanism". Oil and gas journal. Pp 83-89, Sep 1984.

²⁷ Long, D. R and Davis, M.J.: "A New Aproach to the Hyperbolic Curve," JPT (July/1988) 909.

²⁸ Rodriguez, F. and Cinco-Ley, H.: "A New Model for Production Decline," Paper SPE, (March / 1993)

²⁹ Fetkovich, M.J., Fetkovich, E.J., and Fetkovich, M.D.: "Useful Concepts for Decline-Curve Forecasting, Reserve Estimation, and Analysis," Paper *SPE* (February/ 1996) 13-22.

³⁰ Agarwal, R.G., Gardner, D.C., Kleinstieber, S.W. and Fussell, D.D.: "analyzing well production data using combined type-curve and decline-curve analysis concepts," paper SPE 49222, 1998

L.O. Masoner³¹ realizó para 1998, un trabajo donde deduce la ecuación de declinación hiperbólica de Arps a partir de los principios de permeabilidad relativa. Los resultados indican que la declinación hiperbólica aplica rigurosamente en los procesos de recuperación de aceite secundarios y terciarios, dominados por la permeabilidad relativa. Además comenta que las declinaciones exponenciales y armónicas de Arps aplican solamente bajo condiciones muy especiales

J. B Aminzadeh y N.B. Toomarian³² en 1999, demostraron la efectividad de las redes neuronales para proveer una estimación acertada de los parámetros petrofísicos que describen las propiedades del yacimiento, así como las relaciones funcionales, entre las perturbaciones de las propiedades de roca subterráneas y la reacción sísmica, de la formación. Para la validación, utilizaron el método K-fold cross, que estimaba la aproximación y el intervalo de confianza para todas las redes utilizadas en el estudio.

Para el 2000, **H.H. Nguyen y C.W. Chan³³** presentaron una aplicación, basada en el modelo de redes neuronales artificiales múltiples (MNN), para estimar el rendimiento de la producción futura de pozos de petróleo. El estudio mostró que los datos de la geociencia juegan un papel poco significativo en la predicción de producción mensual de petróleo.

F. Abdelhafidh y T. Djebbar³⁴ en el 2001, proponen una ecuación universal de ajuste, que combina las ecuaciones empíricas de Arps y la solución analítica de Fetkovich, como alternativa al ajuste manual con curvas tipo.

³¹ Masoner, L.O.: "Decline Analysis' Relationship to Relative Permeability in Secondary and Tertiary Recovery," paper SPE 39928, (April / 1998)

³² F. Aminzadeh, J. Barhen, and N.B. Toomarian, "Estimation of Reservoir Parameter using a Hybrid Neural Network", *Journal of Petroleum Science and Engineering*, vol. 24, no. 1, 1999, pp. 49-56.

³³ H.H. Nguyen and C.W. Chan, "Petroleum Production Prediction: A Neural Network Approach", *Proceedings of 5th International Joint Conference on Engineering Design and Automation 2001 (EDA 2001)*, 5-8 August 2001, Las Vegas, USA, pp.85-90, 2000

³⁴ Abdelhafidh, F. and Djebbar, T. : "Application of Decline-Curve Analysis Technique in Oil Reservoir Using A Universal Fitting Equation" , paper SPE 70036, (May /2001)

Li and Horne³⁵ proponen en el 2003 un modelo de declinación mecanicista, basado en estudios teóricos y experimentales, que revela una relación lineal entre la tasa de producción de petróleo y el recíproco de la producción de aceite acumulada

En el 2004, **H.H. Nguyen y C.W. Chan**³⁶ presentan una comparación de diferentes estrategias de preprocesamiento de datos, para el desarrollo de modelos de redes neuronales, usados para la predicción de la tasa de producción de petróleo. Las estrategias de preprocesamiento consideradas fueron: (1) aproximación secuencial, en la cual los datos de los pozos son ubicados secuencialmente para formar un conjunto de datos, (2) aproximación promedio, donde el conjunto de datos es formado a partir del promedio de todos los pozos, y por último (3) la aproximación individual que construye un modelo por cada pozo del yacimiento.

H.H. Nguyen y C.W. Chan³⁷ desarrollan para el mismo año, un sistema de soporte de decisiones que ayuda al usuario en la predicción de la producción de petróleo. El sistema muestra los mejores y peores escenarios, basado en las diferentes curvas de producción, con el propósito de que un experto puede revisar las gráficas del pronóstico de la tasa de producción de cada pozo y pueda determinar qué modelo da el mejor ajuste. En su estudio señalaron, que las curvas generadas por las RNA's, deben tener un conjunto de entrenamiento mayor a 30 meses, para que no sufran una tendencia descendiente brusca y anormal al final de la vida de producción.

³⁵ Li, K. and Horne, R.N.: "A Decline Curve Analysis Model Based on Fluid Flow Mechanisms," paper SPE 83470, (May /2003).

³⁶ H.H. Nguyen and C.W. Chan, "A Comparison of Data Preprocessing Strategies for Neural Network Modeling of Oil Production Prediction", Proceedings of the Third IEEE International Conference on Cognitive Informatics, (ICCI 2004)

³⁷ H.H. Nguyen and C.W. Chan, "A Decision Support System For Oil Production Prediction", Proceedings of Electrical and Computer Engineering. Canadian Conference, (CCECE 2004)

R. Camacho, G. Fuentes y M. Vásquez³⁸ publican en el 2006 un estudio donde investigan el comportamiento de la declinación de la producción en yacimientos naturalmente fracturados, que presentan porosidad simple y doble, aplicando geometría fractal.

2.2.3 Fundamento Matemático

Los diferentes métodos de análisis de curvas de declinación están basados en la manera en que la rata de declinación varía, es por esto que es importante que se defina cuidadosamente. Entonces la rata de declinación puede definirse como el cambio fraccional de la producción en el tiempo:

$$a = -(\Delta q / q) / \Delta t \quad (2.17)$$

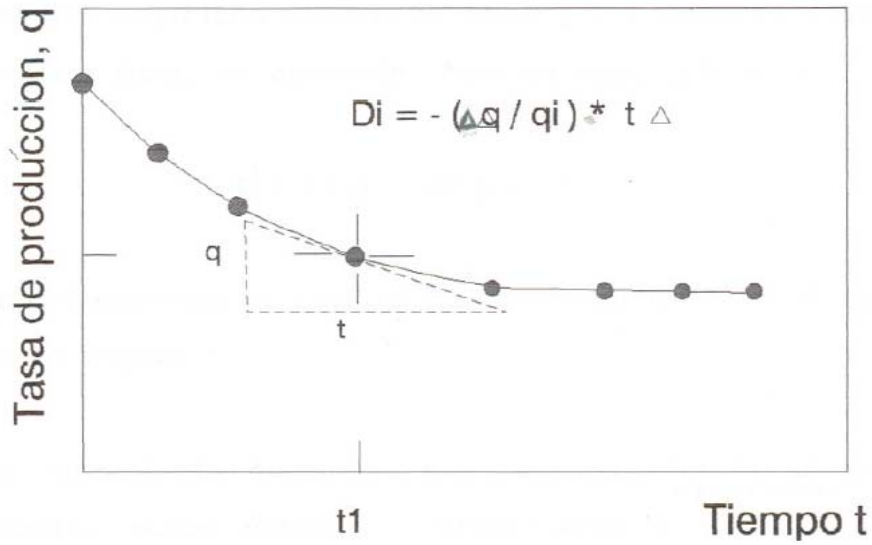
Consecuentemente, la rata de declinación en cualquier tiempo particular, puede determinarse gráficamente calculando la pendiente de la gráfica de la producción contra tiempo, en el punto de interés, y dividiendo por la producción en ese tiempo particular. Para que la rata de declinación a sea constante, la pendiente debe declinar a la misma rata que lo hace la producción (q).

La expresión (2.17) puede reescribirse de la siguiente forma:

$$a = -\Delta \ln q / \Delta t \quad (2.18)$$

³⁸ Camacho, R. , Fuentes, G, Vásquez, M.: "Decline Curve Analysis of Fractured Reservoirs with Fractal Geometry", paper SPE 104009, (september/ 2006)

Figura 4. Definición de tasa de declinación - gráfico lineal



Entonces, la pendiente de la gráfica del logaritmo natural de la producción contra el tiempo es la tasa de declinación a . Sin embargo el papel en escala del logaritmo natural no es de uso común, así que es más conveniente, para poder graficar, establecer la tasa de declinación como una función del logaritmo en base 10. El logaritmo natural es 2.3 veces el logaritmo en base 10 (\log), luego la ecuación (2.18) se convierte en:

$$a = -2.3 \Delta \log q / \Delta t \quad (2.19)$$

Los tres métodos de declinación descritos se derivan de la ecuación diferencial para la tasa de declinación normalizada d :

$$d = - \frac{dq/dt}{q} = Kq^n \quad (2.20)$$

Donde K es una constante de proporcionalidad; en las siguientes tres ecuaciones de declinación, d_i = tasa declinación inicial.

1. Declinación exponencial, $n = 0$.

$$q_t = q_i e^{-d_i t} \quad (2.21)$$

$$Q_p = \frac{q_i - q_t}{d_i} \quad (2.22)$$

2. Declinación hiperbólica, n diferente a 0 ó 1.

$$q_t = \frac{q_i}{(1 + n d_i t)^{1/n}} \quad (2.23)$$

$$Q_p = \frac{q_i^n}{(1-n)d_i} (q_i^{1-n} - q_t^{1-n}) \quad (2.24)$$

3. Declinación armónica, $n = 1$.

$$q_t = \frac{q_i}{(1 + d_i t)} \quad (2.25)$$

$$Q_p = \frac{q_i}{d_i} \ln \left(\frac{q_i}{q_t} \right) \quad (2.26)$$

2.2.3.1 Estimación de los parámetros para declinación exponencial. Si $n = 0$, la declinación es exponencial; de los tres métodos este es el de mas amplio uso, primeramente por la facilidad para determinar sus parámetros y además porque da un estimativo más conservador de la producción; si bien esta última afirmación es cierta, también lo es que la declinación hiperbólica arroja resultados mas realistas y por ello se recomienda su uso cuando las condiciones sean las apropiadas.

Si en una declinación dada se observa un fuerte carácter hiperbólico, es fácil ajustar una curva exponencial a unos pocos puntos finales, tal que mediante este ajuste se pueda proyectar el comportamiento de la producción a futuro. Ahora si se asume que $n = 0$, sólo hay que determinar dos parámetros: q_i y d_i ; tomando el logaritmo natural a ambos lados de la relación producción/tiempo de la declinación exponencial descrita en la ecuación (2.21):

$$\ln q_t = \ln q_i - d_i t \quad (2.27)$$

Entonces, la gráfica de $\ln q_t$ vs. t , es una línea recta con $m = -d_i$ y $b = \ln q_i$. Como el papel semilog es más común, entonces se estima que es conveniente graficar $\log q_t$ vs. t , donde $m = -d_i \log e$ y $b = \log q_i$.

2.2.3.2 Estimación de los parámetros para la declinación armónica. Si $n = 1$, la declinación es armónica; este es el menos común de los métodos, pero no por ello el más simple. Puede aplicarse, siempre y cuando la gráfica de $\log q_t$ vs. Q_p presente un comportamiento lineal. Esto se deduce de la ecuación (2.26), que puede reescribirse como:

$$\ln q_t = \ln q_i - d_i Q_p / q_i \quad (2.28)$$

Por tanto, una gráfica de $\ln q_t$ vs. Q_p da una línea recta donde $m = -d_i / q_i$ y $b = \ln q_i$. Si se quiere utilizar un papel semilog, entonces debe graficarse $\log q_t$ vs. Q_p , donde $m = -d_i / q_i \log e$ y $b = \log q_i$.

Así como es posible ajustar los últimos valores de producción con una declinación exponencial, del mismo modo es posible ajustar los datos finales con una declinación armónica debido a que pocos puntos serán lineales en una gráfica de $\log q_t$ vs. Q_p ; Sin embargo este es un procedimiento erróneo pues conduce a resultados muy optimistas si el verdadero valor de n es menor a 1.

2.2.3.3 Determinación de los parámetros para la declinación hiperbólica.

Realmente casi todos los pozos presentan una declinación hiperbólica; las declinaciones, armónica y exponencial son simples casos particulares de la hiperbólica, y habitualmente se tratan por separado debido a que la integración de la ecuación (2.20), no se refleja en la ecuación (2.23) cuando $n = 0$, y la integración de la ecuación (2.23) no resulta en la ecuación (2.25) cuando $n = 1$. Además estos casos particulares son más simples de utilizar porque sólo se necesita determinar dos parámetros. Sin embargo los tres parámetros referentes a las curvas de declinación, deberían calcularse para todos los pozos porque estos muestran que casi no existen pozos que tengan declinaciones con valores de n de 0 ó 1 exactamente.

Si bien es más difícil determinar tres parámetros (q_i , d_i , y n), en la literatura se pueden encontrar muchas aproximaciones para su cálculo, algunos basados en mecanismos de prueba y error, otros en la superposición de curvas tipo, en métodos gráficos y algunos otros en análisis matemáticos y estadísticos.

Explicaremos aquí sólo uno de los métodos para obtener el valor de los parámetros para la declinación hiperbólica, por ser este uno de los más utilizados y por ser el escogido para efectuar las comparaciones en este trabajo. Los pasos se describen a continuación:

1. Asumir un valor para nd_i .
2. Graficar $\log q_i$ vs. $\log(1 + nd_it)$.
3. Si no es una línea recta o si el coeficiente de regresión no es máximo, volver al paso 1 y repetir.
4. si la línea es recta y el coeficiente de regresión es máximo, $m = -1/n$ y $b = \log q_i$, entonces d_i puede obtenerse dividiendo por n el valor de nd_i asumido en el paso 1.

2.3 MECANISMOS DE PRODUCCIÓN DEL PETRÓLEO

La producción es la actividad de la industria petrolera que se ocupa de realizar todos los procesos necesarios para llevar el petróleo desde el yacimiento hasta el pozo, y de éste a la superficie, tratando de extraer la mayor cantidad de fluido³⁹.

En la producción de un yacimiento no siempre se obtienen las tasas de flujo deseadas. A raíz de esto, a medida que se desarrolla el campo se piensa en usar métodos de extracción artificial que permitan obtener tasas de flujo óptimas para hacer rentables las operaciones.

Si el yacimiento tiene suficiente presión, el petróleo fluirá a superficie por su propia energía, proceso que se denomina “**flujo natural**”. En este caso se instala en la cabeza del pozo un conjunto de válvulas para regular el paso del crudo.

Si la presión no es suficiente para que el petróleo fluya a la superficie, se emplean otros métodos de extracción conocidos como “Métodos de Levantamiento Artificial”. El más común es el Bombeo Mecánico (se utiliza en el 85% de los casos), el cual mediante un permanente balanceo, acciona una bomba colocada en el fondo del pozo la cual impulsa el petróleo hacia la superficie. El objetivo de estos es crear una presión en la tubería de producción tal, que permita al pozo fluir con un volumen determinado una vez que la energía natural del mismo ha disminuido.

Entre los diferentes sistemas de extracción artificial del petróleo se encuentran el bombeo mecánico, neumático, hidráulico y Electro-Centrífugo.

³⁹ Asociación colombiana de ingenieros de petróleo: “Memorias del seminario: Ingeniería de petróleos para no Petroleros”, Bogotá D.C., septiembre 11 al 15 de 2006

2.3.1 Bombeo Mecánico

Es uno de los métodos de levantamiento artificial más utilizados. Una bomba colocada en el fondo del pozo, es accionada por medio de una sarta de varillas de acero movidas por un machín o balancín que en su balanceo permanente impulsa el petróleo a la superficie. El diagrama de este bombeo, se puede apreciar en la Figura 4.

Figura 5. Bombeo mecánico⁴⁰

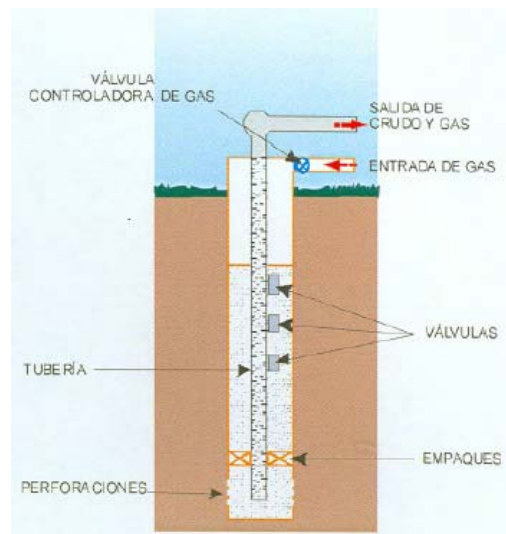


2.3.2 Bombeo neumático

Consiste en la inyección de gas, a niveles prefijados y a través de unas válvulas colocadas en la misma tubería de producción, con el fin de aligerar el peso de la columna de petróleo, permitiendo que la misma alcance la superficie con la energía propia del yacimiento.

⁴⁰ Fuente: Asociación colombiana de ingenieros de petróleo: "Memorias del seminario: Ingeniería de petróleos para no Petroleros", Bogotá D.C., septiembre 11 al 15 de 2006

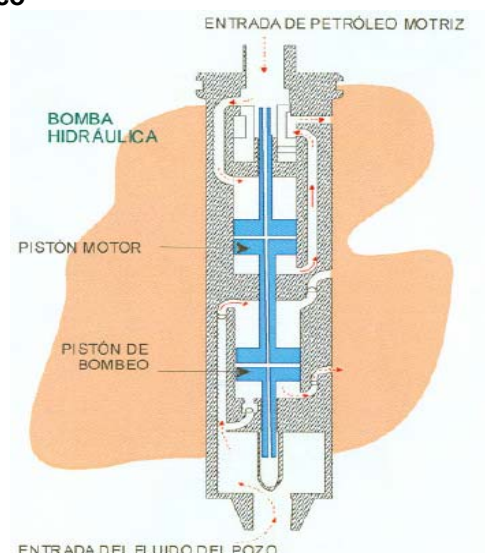
Figura 6 Bombeo neumático⁴¹



2.3.3 Bombeo Hidráulico

El principio de este sistema de levantamiento es la inyección de un fluido de potencia a altas presiones que al mezclarse con el fluido del yacimiento dentro del equipo de subsuelo, le transmite la energía necesaria para que este llegue a la superficie ya sea a través de una tubería alterna o por el anular.

Figura 7. Bombeo Hidráulico⁴²



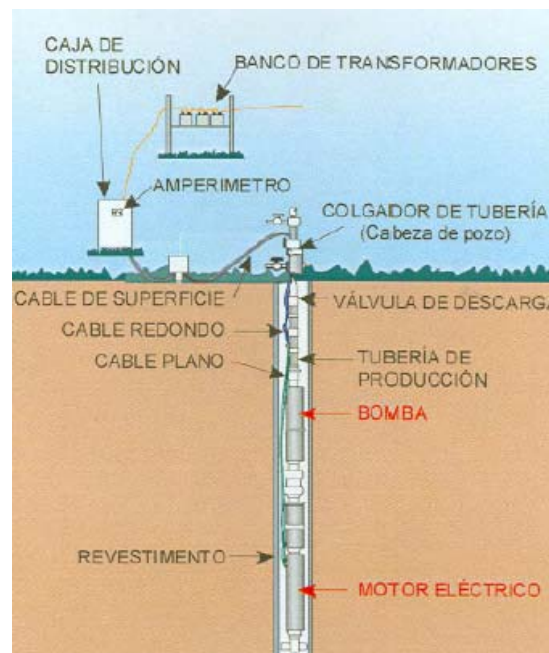
⁴¹ Fuente: Asociación colombiana de ingenieros de petróleo: "Memorias del seminario: Ingeniería de petróleos para no Petroleros", Bogotá D.C., septiembre 11 al 15 de 2006.

⁴² Fuente: Asociación colombiana de ingenieros de petróleo: "Memorias del seminario: Ingeniería de petróleos para no Petroleros", Bogotá D.C., septiembre 11 al 15 de 2006

2.3.4 Bombeo Electro-Centrífugo o Electro-Sumergido

El crudo de la formación es impulsado hasta la superficie por una bomba centrífuga colocada al mismo nivel de la formación productora y accionada por un motor eléctrico.

Figura 8. Bombeo Electro-Centrífugo⁴³



De acuerdo a las características del yacimiento y del fluido que se quiere producir se puede pensar en usar uno u otro sistema de levantamiento artificial. El grado de uso que tiene cada uno se debe a las características de los fluidos, tales como viscosidad, densidad, relación agua - aceite, contenido de gas, sólidos, temperaturas, etc. Y adicionalmente a la experiencia y conocimiento que se posea del particular, por lo cual sobresale el bombeo mecánico.

En general, los factores que deben tenerse en cuenta al decidir cual técnica de producción es mejor para un pozo se definen a continuación:

⁴³ Fuente: Asociación colombiana de ingenieros de petróleo: "Memorias del seminario: Ingeniería de petróleos para no Petroleros", Bogotá D.C., septiembre 11 al 15 de 2006

1. Profundidad
2. Caudal (presente y futuro)
3. Presión
4. GOR(relación gas – aceite)
5. Arena y parafinas
6. Viscosidad
7. Diámetro del pozo
8. Desviación del pozo
9. Disponibilidad de gas
10. Facilidades económicas
11. Facilidades (espacio para reparaciones)
12. IPR (Comportamiento de producción del pozo)
13. Planes de recuperación
14. WOR (relación agua – aceite)

Una correcta selección de un método de Levantamiento Artificial es importante para el beneficio a largo plazo de la mayoría de los pozos productores de petróleo; Una selección pobre puede reducir la selección e incrementar los costos operativos; Cambiar el método de levantamiento cuesta dinero e implica que la selección del método anterior fue errónea.

CAPITULO 3

PROCESO DE IDENTIFICACIÓN DEL MODELO

Estudia el pasado si quieres pronosticar el futuro

CONFUCIO

La parte más difícil, y de mayor importancia, en la construcción de un buen modelo de red neuronal corresponde a la selección y recolección de la información que lo alimentará en la fase de entrenamiento; identificar correctamente las variables verdaderamente relevantes dentro del proceso modelado requiere de intensas labores de observación, análisis e investigación, que han sido sistemáticamente desarrolladas a lo largo de este trabajo. En este punto es necesario expresar que la base del entrenamiento consiste únicamente de la información de los históricos de producción de pozos petroleros, con el objetivo de simular su comportamiento a futuro. Los estudios sobre la predicción de la producción realizados por Nguyen y Chan⁴⁴, mostraron que los datos de la geociencia⁴⁵ juegan un papel poco significativo en la aplicación de un modelo neuronal, y por ello se ha optado por no incluir estos factores dentro del estudio; adicionalmente cabe resaltar que una fuerte limitante cuando se trabaja con datos históricos es que sólo se cuenta con la información que está registrada, y esto implica adecuar totalmente el espacio de variables consideradas a aquellas que estén disponibles en archivo.

El modelo final al que se llegó en este proyecto de grado, es producto de una natural evolución, pues parte de uno básico que lentamente fue madurando en

⁴⁴ H.H. Nguyen and C.W. Chan, "Petroleum Production Prediction: A Neural Network Approach", Proceedings of 5th International Joint Conference on Engineering Design and Automation 2001 (EDA 2001), 5-8 August 2001, Las Vegas, USA, pp.85-90, 2000

⁴⁵ Geociencia: conjunto de disciplinas que estudian la estructura interna, la morfología superficial y la evolución de la tierra, entre ellas, la geología y la geografía.

la medida en que el tema de estudio era comprendido, hasta alcanzar el punto de estabilidad en términos de resultados satisfactorios. Como en todo modelo neuronal, este cuenta con variables que constituyen las entradas de la red y otras que corresponden con las salidas deseadas; cada tipo de variable tiene una función específica dentro del modelo: las variables de entrada tienen una participación activa durante los procesos que realiza internamente la red para tratar de identificar el patrón propuesto, mientras que las variables de salida juegan su principal rol en el proceso de aprendizaje puesto que, al ser el objetivo buscado, permiten medir la validez de los resultados arrojados por la red neuronal.

3.1 SELECCIÓN DE LAS VARIABLES

En el planteamiento de un primer modelo se utilizaron sólo dos variables para la configuración de la red neuronal:

3.1.1 Línea mensual de tiempo

Esta entrada corresponde a una secuencia de números, que indican los meses de producción del pozo. Su inclusión dentro del modelo es muy importante, porque es necesario conocer, durante la etapa de entrenamiento y de validación, en que parte de la línea de tiempo está la red. Es decir, que si para un pozo petrolero se tienen datos de 3 años de producción con registros mensuales, la línea de tiempo correspondiente a los valores históricos sería: [1, 2, 3,..., 36], donde 1 correspondería al primer mes de producción y 36 al último. Los valores de la serie mayores a 36, equivaldrían a los meses de producción venideros, y son usados para indicar que se debe realizar la predicción. Se puede decir que esta es una variable de alto impacto y que cuyo seguimiento es fundamental dentro del proceso de entrenamiento puesto que, por la naturaleza altamente variable de la producción, es necesario poder distinguir claramente entre dos periodos de tiempo de características similares, porque en nuestro estudio, la producción del mes de enero de un año

específico varía considerablemente de la de enero del año siguiente y del mismo modo de la de enero del año inmediatamente anterior.

3.1.2 Producción de petróleo

Esta variable de salida Indica la cantidad de barriles de petróleo obtenidos de un pozo durante un mes. La producción precisamente corresponde con la información histórica almacenada y va ligada en relación uno a uno a la línea de tiempo.

A medida que se fueron estableciendo las entradas y las salidas, paralelamente y de forma complementaria, fueron consideradas las diferentes características relacionadas con el diseño de la red neuronal, entre ellas, el algoritmo de entrenamiento, la topología y la función de transferencia más adecuada para la solución del problema. De esta forma se plantearon diferentes tipos de redes, en un principio de forma heurística, siempre con la política de encontrar los mejores resultados con el menor número de neuronas posibles; sin embargo los resultados obtenidos con estas dos únicas variables, no fueron satisfactorios. Los errores obtenidos durante el entrenamiento y la simulación se pueden apreciar a continuación.

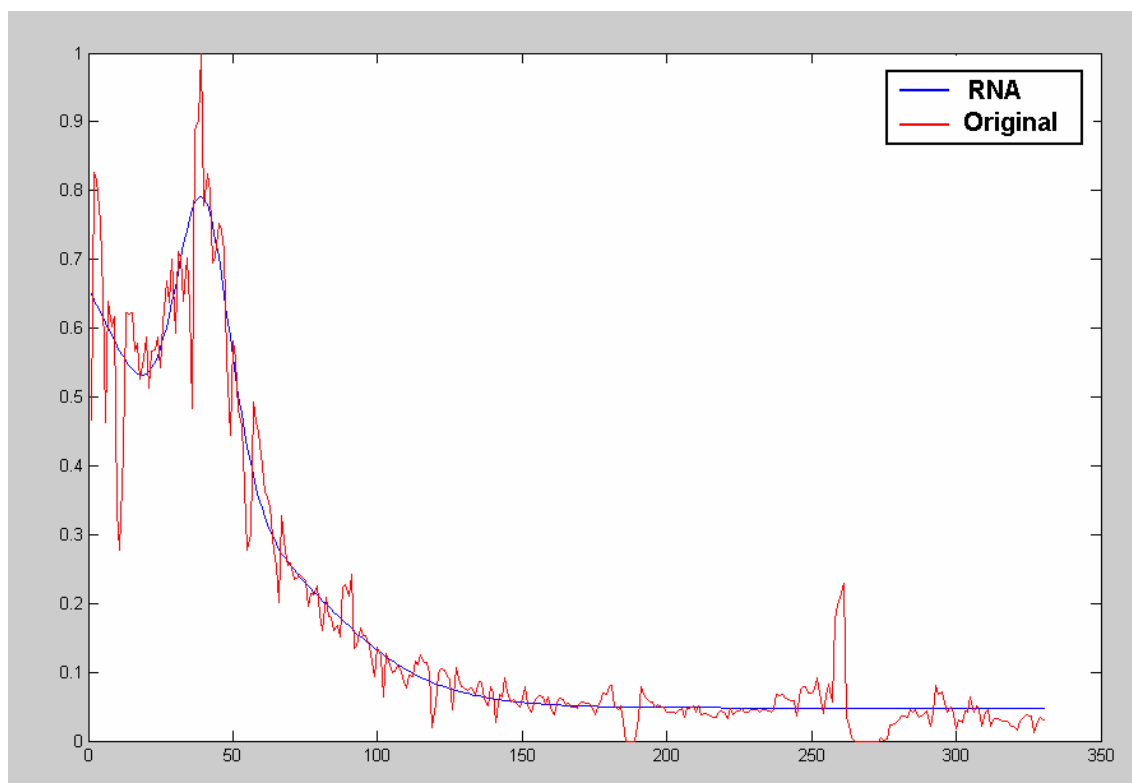
Tabla 2. Errores de entrenamiento y validación para el Modelo 1

	Error Cuadrático Medio	Error Absoluto Medio
Entrenamiento	0.7180	0.0526
Validación	0.0434	0.0414

Tabla 3. Entradas y salidas de la red neuronal al Modelo 1

Entrada	Salida
Mes	Producción de petróleo
1	P1
2	P2
3	P3
4	P4
...	...
N	Pn

Figura 9. Simulación del Modelo 1



El segundo modelo propuesto surge como complemento del primero, en aras de obtener una mayor precisión en los resultados. Se introdujo entonces para este fin una segunda variable de entrada.

3.1.3 Días de producción de cada mes

A medida que se fue desarrollando la investigación se observó que la producción de algunos meses difería notablemente de la tendencia esperada; esto surge como consecuencia de que el pozo no estuvo activo durante todo el tiempo sino que su producción fue suspendida, bien sea por labores de mantenimiento ó por fallas técnicas. A raíz de esta observación se mejoró el primer modelo considerando esta variable que corresponde exactamente con los días que realmente el pozo produjo durante un mes.

Luego de los cambios realizados al primer modelo, la predicción de la producción se ajustó de una mejor forma a los datos presentados durante el entrenamiento; sin embargo siempre se busca que los resultados obtenidos mediante la simulación sean cada vez más cercanos a los reales. Esto se puede apreciar en la tabla 4, que contiene los errores obtenidos durante el entrenamiento y la simulación del modelo 2.

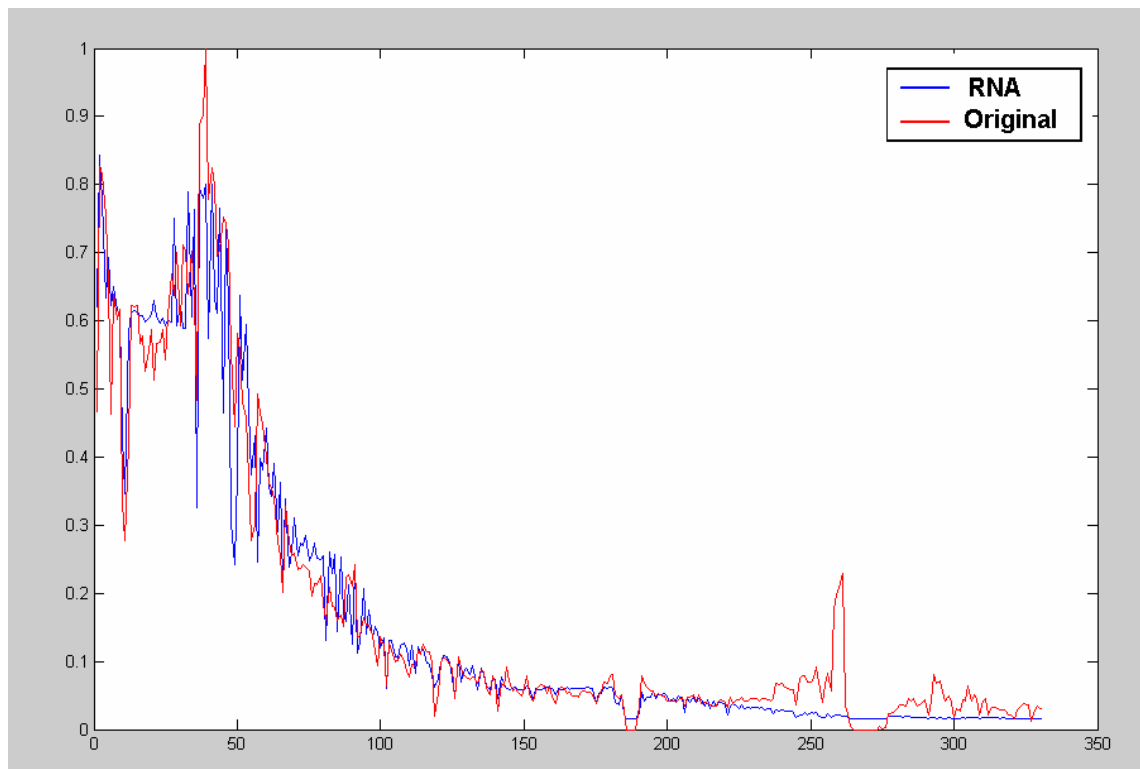
Tabla 4. Errores de entrenamiento y validación para el Modelo 2

	Error Cuadrático Medio	Error Absoluto Medio
Entrenamiento	0.0163	0.0362
Validación	0.0089	0.0278

Tabla 5. Entradas y salidas de la red neuronal Modelo 2

Entradas		Salida
Mes	Días Producción	Producción de petróleo
1	31	P1
2	22	P2
3	30	P3
4	12	P4
...
N	# Días	P _n

Figura 10. Simulación del Modelo 2



Conforme avanzaba la investigación, crecía también la necesidad de encontrar un modelo que arrojara resultados más precisos; por ello se ahondó en los estudios acerca de los factores que tienen un efecto notable en la producción de crudo por parte de un pozo petrolero. Se incluyó entonces la variable de entrada correspondiente al tipo de bombeo.

3.1.4 Tipo de bombeo

Esta variable de entrada indica la clase de bombeo que se está usando para extraer el petróleo del pozo, y es de suma importancia, por la información inherente que contiene, pues cuando cambian las condiciones de operación del pozo como presión y caudal, se hace necesario modificar el mecanismo de levantamiento; un cambio acertado en este punto conduce a un mejoramiento de la producción. Con la introducción de esta variable al modelo, le estamos indicando que si se presentan variaciones de la producción por fuera de la

tendencia esperada, se debe a que tales cambios no corresponden al comportamiento natural del pozo, sino que son el producto de procedimientos técnicos, de tal manera que cuando esta situación vuelva a presentarse, el modelo logre adaptarse lo mejor posible al cambio en la tendencia. Las mejoras obtenidas, utilizando esta nueva variable, se puede apreciar en la tabla 6.

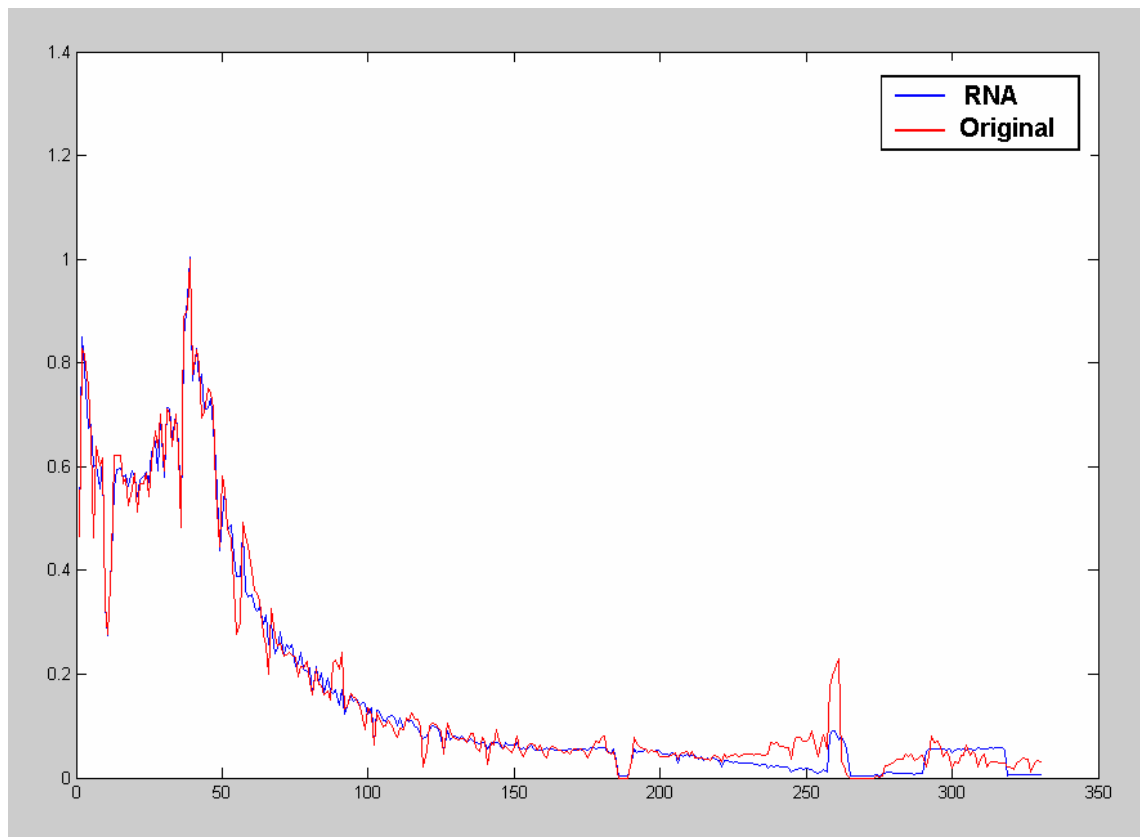
Tabla 6. Errores de entrenamiento y validación para el Modelo 3

	Error Cuadrático Medio	Error Absoluto Medio
Entrenamiento	0.0015	0.0190
Validación	0.0024	0.0236

Tabla 7. Entradas y salidas de la red neuronal Modelo 3

Entradas			Salida
Mes	Días Producción	Tipo Bombeo	Producción de petróleo
1	31	F	P1
2	22	F	P2
3	30	BM	P3
4	12	BM	P4
...
N	#Días	BE	Pn

Figura 11. Simulación del Modelo 3



3.2 PREPROCESAMIENTO DE LA INFORMACIÓN

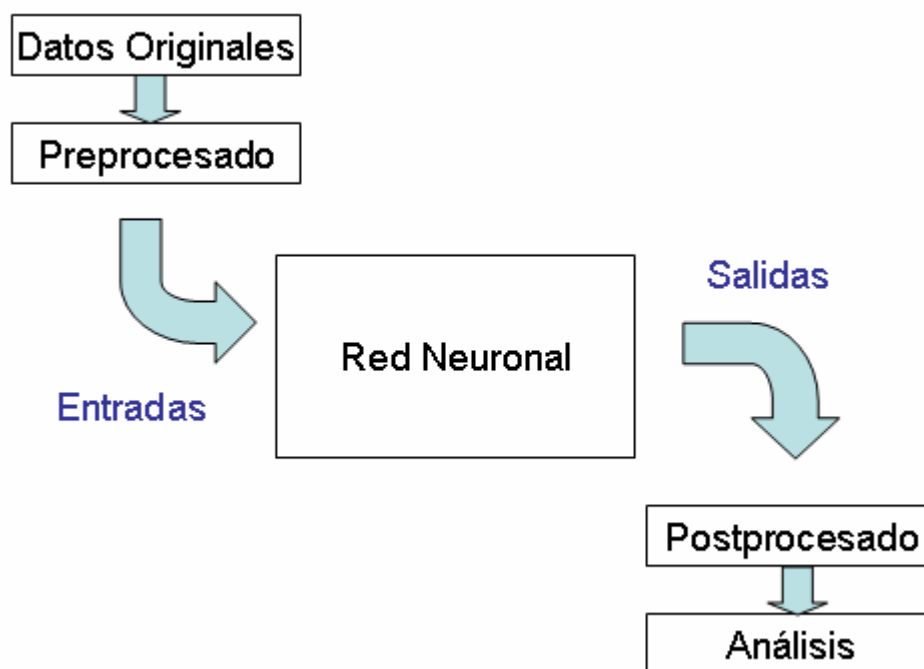
Uno de los aspectos fundamentales para el éxito de cualquier modelo neuronal, es la calidad de los datos que se le presenten en la fase de entrenamiento; es por esto que se ha hecho necesario implementar técnicas para preprocesar la información con el fin de aumentar la capacidad de aprendizaje de la red.

El preprocesamiento de la información consiste de la preparación previa que se le aplican a los datos antes de ser empleados por la red, tal que puedan mostrar de forma clara y precisa la información relevante, y es este paso tan importante dentro del desarrollo de un modelo neuronal, que es posible que el mejor de los modelos fracase si se entrena con datos de baja calidad.

Existen diversos métodos y estrategias para lograr un buen preprocesamiento, entre ellas:

- Identificar las variables de naturaleza binaria y asignarle valores de 0 ó 1.
- Evitar el uso de valores continuos para conceptos simbólicos ó meramente discretos.
- Evitar mezclar escalas (Ej: Toneladas con Kilogramos).
- Evitar utilizar variables cuyos valores contengan altas variaciones (máximos y mínimos).
- Tratar de estandarizar patrones para reducir el número de entradas
- Normalizar las variables para llevarlas a un rango entre 0 y 1.
- Transformar las variables de un valor categórico a un valor numérico.

Figura 12. Esquema del preprocesamiento de la información



En este trabajo no se emplearon todas las técnicas arriba descritas pues todas no se aplican al modelo planteado, pero el preprocesamiento que se hizo surtió

un buen efecto sobre los resultados obtenidos, mejorándolos sustancialmente. Las técnicas de preprocesamiento de la información utilizadas fueron:

3.2.1 Transformación de Variables Nominales a Numéricas

El tipo de bombeo viene expresado en los datos históricos de producción, por valores alfabéticos, este formato no es compatible con los valores de entrada de la red, pues no hay forma de que ésta los utilice en operaciones aritméticas. Por ello se le asignan valores discretos entre 0 y 1 para representar dicha categorización. En la tabla 8 se resume este procedimiento.

Tabla 8. Transformación de variables nominales a numéricas

Tipo de Bombeo	Valor Alfabético	Categorización
Flujo Natural	F	0
Mecánico	B	0.25
Electro-sumergible	BE	0.5
Hidráulico	BH	0.75
Otros	Otros	1

3.2.2 Filtrado

Consiste en repasar la serie de los datos pertenecientes a la producción buscando posibles datos atípicos. Se ha identificado, en los diferentes archivos de pozos estudiados, datos que no están dentro de la tendencia esperada, bien sea por un aumento casual y puntual en algunos meses de producción, ó por su decremento, que muchas veces puede interpretarse como un error en el registro de la información, y otras veces como un cambio repentino producto de algún procedimiento técnico provisional.

El proceso de identificación de tales datos atípicos, se hizo comparando inicialmente la diferencia entre el registro actual de producción y el siguiente, de tal manera que si este valor supera uno de dos umbrales, superior o inferior, definidos con valores del 50% por encima y 40% por debajo de la producción, respectivamente, el dato pasa a ser un candidato atípico. El registro candidato es comparado ahora contra la media aritmética de los 5 registros siguientes, con el fin de evaluar si éste estaba acorde con la tendencia, para ello se utilizaron los umbrales definidos, tal que si la diferencia los superaba, el dato se considera definitivamente atípico. Antes de hacer la identificación, los datos deben dividirse por el número de días de producción asociados para calcular la rata de producción diaria, pues es mas precisa que la mensual, porque guarda una relación directa con los días en que el pozo estuvo activo durante el periodo registrado. De esta forma se evita la interpretación de falsos datos atípicos basados en registros donde la producción del mes fue baja debida a pocos días de actividad.

Una vez identificado el dato como atípico, se procede a reemplazarlo por el promedio de las medias de los 5 registros siguientes y los 5 registros anteriores a él. El proceso de filtrado de datos se puede apreciar en la fig. 12.

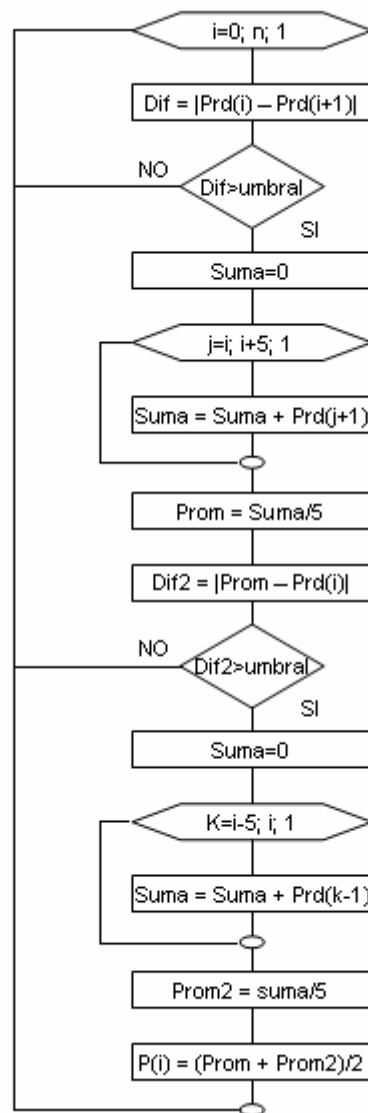
3.2.3 Normalización

Se cambió la escala de las variables Producción y Días de producción a valores entre 0 y 1, la primera de ellas luego de pasar por el proceso de filtrado, con el fin de lograr una mayor generalización de la red en el proceso de aprendizaje. Esto se logra dividiendo el conjunto de datos de cada variable por el mayor valor encontrado en cada uno.

$$Valor_Normalizado = \frac{Valor}{Valor_Máximo} \quad (3.1)$$

Una vez conocido el número de entradas y salidas a tener en cuenta, se pasa a considerar las diferentes características relacionadas con el diseño de la red, entre ellas, el algoritmo de entrenamiento y la topología más adecuada para la solución del problema.

Figura 13. Diagrama de flujo del filtrado de datos



Dentro de la topología de red se define el tipo de red, número de capas, número de neuronas por capa y las funciones de transferencia, que mejor se adapten a las características inherentes de los datos con que se cuenta.

3.3 ARQUITECTURA DE RED

A pesar de la amplia variedad de modelos y variantes de redes neuronales que existen, pocos son los que habitualmente se utilizan a la hora de desarrollar aplicaciones. Desde el punto de vista práctico, los más utilizados son el perceptrón multicapa o MLP (Multilayer Perceptron), dentro de los modelos de aprendizaje supervisados, y los mapas auto-organizativos o mapas de Kohonen (Self-organizing Feature Maps), red de Hopfield o las más recientes funciones de base radial o RBF (Radial Basis Function), englobados dentro de los modelos no supervisados⁴⁶. También se utilizan modelos híbridos de redes neuronales con lógica borrosa, entre los más usados está el ANFIS (Adaptive Neuro Fuzzy Inference System).

3.3.1 Redes de Base Radial (RBF)

En las redes de base radial, añadir o eliminar unas pocas neuronas ocultas podría influir significativamente en los resultados obtenidos por la red. Esto se debe a que cada neurona oculta representa una determinada región del espacio de entrada, pudiendo no estar adecuadamente representado, bien sea por la presencia de pocas clases, o por la presencia de muchas neuronas ocultas en una misma zona, que tienen como consecuencia negativa la mala aproximación de la red. Por ello las neuronas pertenecientes a la capa oculta en una RBF poseen un carácter local, lo que supone que deba tenerse un rango delimitado en las entradas de la red, para que esta pueda ser manejable en términos de tiempo de procesamiento, de recursos computacionales y de predicciones acertadas. Dicho en términos prácticos, las RBF pueden ser aplicadas en modelos tales como el de la predicción del consumo de energía eléctrica, donde, bajo ciertas condiciones de entrada como el mes, día, temperatura y hora entre otros factores, la salida tendrá un valor que sólo depende de estos, independientemente del año en que se encuentre; en este

⁴⁶ Véase Martín del Peso, Mónica: "Aplicaciones de las redes neuronales artificiales a problemas de predicción y clasificación financiera", Madrid: Universidad Rey Juan Carlos, 2005. p 63.

sentido puede considerarse este sistema como cuasi-estático en el tiempo. Por el contrario, el comportamiento de la producción en un pozo petrolero no se puede limitar a un rango predefinido porque la línea de tiempo puede extenderse tanto como registros históricos haya y de qué tan amplio sea el plazo escogido de simulación; además, debido a la naturaleza declinante del pozo a lo largo del tiempo, las mismas condiciones de operación no necesariamente suponen un comportamiento similar a uno observado anteriormente.

La desventaja de utilizar este tipo de redes se evidenció de manera clara y concluyente, luego de hacer las pruebas en Matlab y observar el sobreentrenamiento con los datos propuestos y una pésima generalización del patrón. Se concluyó entonces que las redes de base radial no eran las idóneas para el desarrollo de nuestro trabajo. Los resultados gráficos, se pueden apreciar en la fig. 13.

Tabla 9. Errores de entrenamiento y validación para la red RBF

	Error Cuadrático Medio	Error Absoluto Medio
Entrenamiento	4.048×10^{-11}	2.925×10^{-4}
Validación	0.2789	0.5264

3.3.2 Sistemas Neuro-Borrosos

Se estudió la viabilidad de utilizar ANFIS, que es un sistema híbrido que combina técnicas de RNA y sistemas de inferencia borrosa, para la solución de nuestro problema, pero por los pobres resultados obtenidos en el proceso de simulación, se optó por descartar esta opción.

Tabla 10. Errores de entrenamiento y validación para ANFIS

	Error Cuadrático Medio	Error Absoluto Medio
Entrenamiento	0.0046	0.0403
Validación	0.1112	0.0718

Figura 14. Simulación y validación utilizando RBF

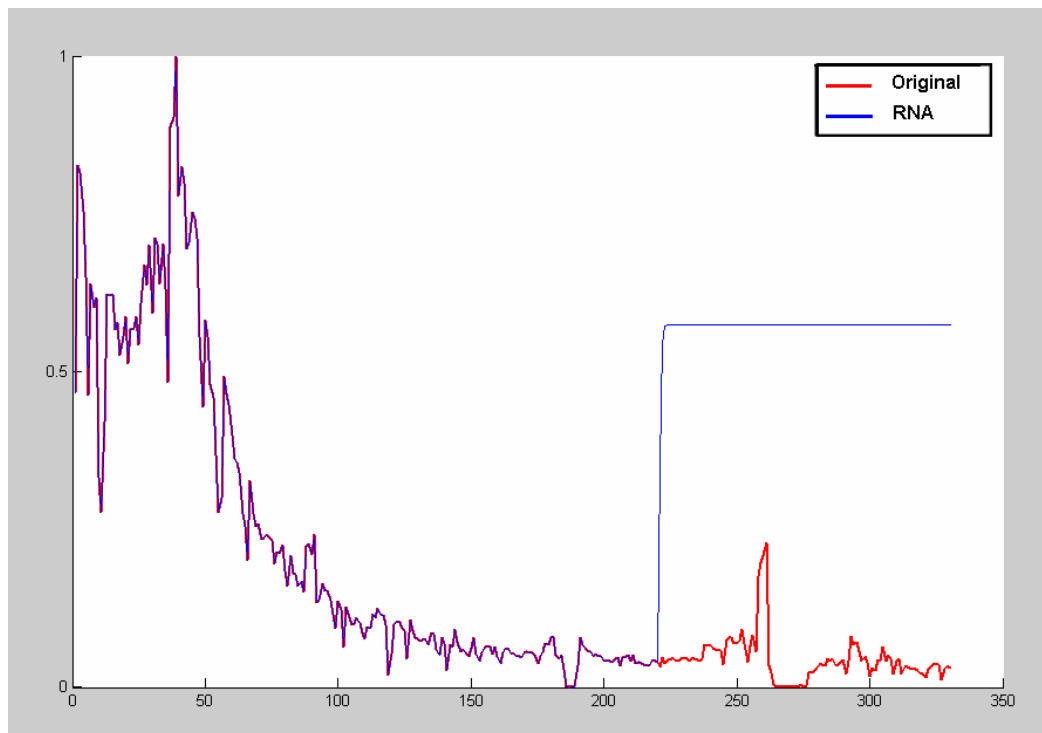
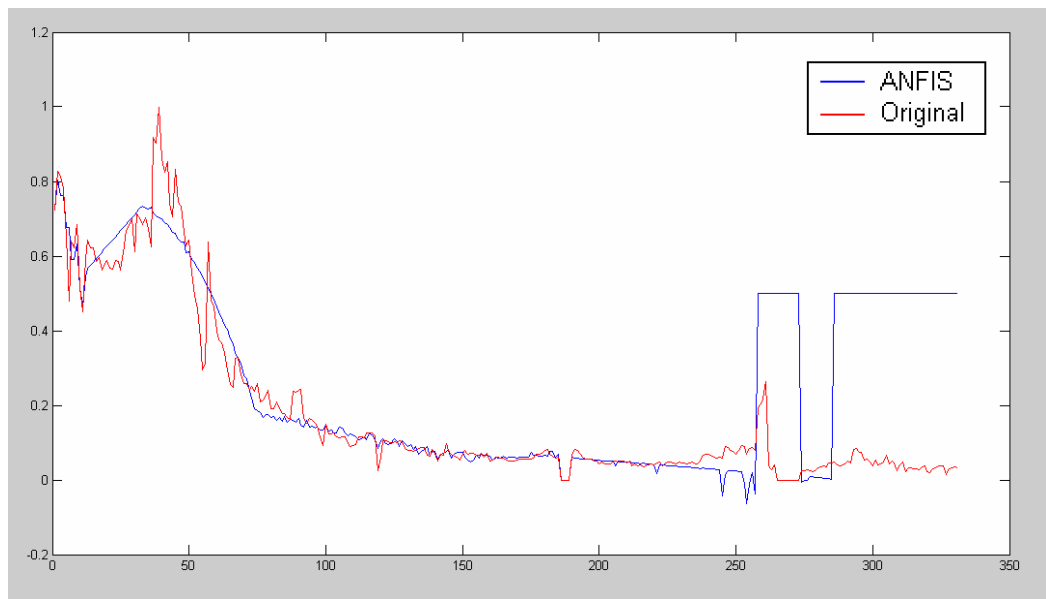


Figura 15. Simulación y validación utilizando ANFIS



3.3.3 Perceptrón Multicapa

Debido al conocimiento previo que se tiene del pozo petrolero a estudiar, representado en los datos de producción, unido a la relación altamente no lineal entre las variables consideradas, y como consecuencia de los pobres resultados obtenidos con las redes de base radial y con los sistemas neuroborrosos, se decidió que una arquitectura perceptrón multicapa, que permite un aprendizaje de tipo supervisado, es la adecuada para el desarrollo del modelo. Algunas de las ventajas que ofrecen las redes tipo perceptrón multicapa son:

- Gran variedad de algoritmos de entrenamiento que permiten que el modelo converja en un tiempo reducido.
- Posibilidad de ser entrenadas y simuladas en línea, suministrando al modelo los datos de producción en el mismo instante en que estos son obtenidos.
- Capacidad de generalización, que permite obtener resultados coherentes a partir de datos no suministrados durante el entrenamiento.

Se empleó el esquema de redes completamente conectadas con alimentación hacia delante (feedforward networks) para evitar la inestabilidad que puede presentarse en redes recurrentes.

3.4 DIMENSIONAMIENTO DE LA RED

El número de capas y neuronas que se necesitan, exactamente, para resolver un problema concreto depende de la naturaleza de este. Las redes con un número elevado de capas pueden resolver problemas muy complejos, pero el elevado tiempo de entrenamiento y la memorización son desventajas inherentes al aumento del número de capas. Luego, se debe establecer una arquitectura intermedia entre rendimiento y precisión.

Para hallar el número de neuronas de las capas de entrada y salida que permitían una mejor aproximación al comportamiento de la producción del pozo petrolero, se utilizó el análisis de variables del sistema, descrito en la sección anterior, y se determinó que el número de neuronas en la capa de entrada es igual al número de variables de entrada, y que el número de neuronas en la capa de salida es igual al número de variables de salida. En este caso 3 y 1, respectivamente.

Determinar el número de neuronas a utilizar en la capa oculta no suele ser tan evidente como lo es para las capas de entrada y de salida. El objetivo principal es minimizar el número de unidades a emplear en esta capa, porque cada unidad adicional supone una carga computacional extra durante las simulaciones. La experiencia indica que, para redes de gran tamaño (cientos o miles de entradas), el tamaño de la capa oculta sólo necesita ser una pequeña porción del que tiene la capa de entrada.

Para establecer el número de neuronas de la capa oculta, se tuvo en cuenta el criterio de convergencia de la siguiente forma: si la red no llega a converger

para llegar a una solución, o dicha tendencia es muy lenta, es posible que se necesiten más neuronas. Si converge, se puede probar con un número menor de neuronas. El tamaño final de la capa se define basándose en el rendimiento global del sistema. Otro criterio de selección se basa en que el error cuadrático medio (MSE) para el conjunto de entrenamiento y validación sea el menor, lo que sugiere que el número de neuronas correspondientes son las adecuadas para la fase de entrenamiento.

La configuración de la Red Backpropagation con el mejor desempeño se calculó utilizando una rutina en matlab 6.5 , que variaba el número de neuronas para una y dos capas ocultas. Dicho número se modificó entre un mínimo de 2 y un máximo de 10 neuronas, empezando por la primera capa oculta y continuando con la segunda, hasta probar todas las configuraciones posibles. Este proceso se repetía 10 veces en cada ejecución del programa, asegurando que se encontraría la mejor configuración posible entre todas las redes que se entrenaban. Se analizó en total 10×9 y 10×36 redes en cada ejecución, para pruebas con una y dos capas ocultas respectivamente.

Para evaluar el buen desempeño de la red es necesario comparar los valores reales y los estimados por la red con el fin de calcular el error en la simulación, es por eso, que en una misma ejecución de la rutina, se calculaba el error cuadrático medio y el error total. Estos valores se comparaban con los calculados para la red anterior, y si mejoraban, se guardaba esta nueva configuración como la mejor. En las tablas 11 y 12, se muestran los diferentes errores con los que se evaluó el desempeño de la red con una capa oculta.

Tabla 11. Errores de entrenamiento y tiempo consumido para redes de 1 capa oculta

Errores de entrenamiento			
Configuración De la red	Error cuadrático Medio	Error absoluto Medio	Tiempo de Entrenamiento
3-2-1	0.0094	0.0536	1.0485 S
3-3-1	0.0074	0.0434	1.1548 S
3-4-1	0.0068	0.0404	1.2108 S
3-5-1	0.0070	0.0421	1.2344 S
3-6-1	0.0017	0.0245	1.2812 S
3-7-1	0.0035	0.0326	1.3141 S
3-8-1	0.0015	0.0227	1.2844 S
3-9-1	0.0011	0.0212	1.2750 S
3-10-1	0.0062	0.0369	1.4265 S

Tabla 12. Errores de validación para redes de 1 capa oculta

Errores de Validación		
Configuración de la red	Error cuadrático Medio	Error absoluto Medio
3-2-1	0.0178	0.0970
3-3-1	0.0164	0.0744
3-4-1	0.1641	0.1276
3-5-1	0.0588	0.1160
3-9-1	0.0025	0.0288
3-7-1	1.1821	0.2263
3-8-1	0.0087	0.0481
3-6-1	0.0072	0.0468
3-10-1	0.0529	0.0740

Para las pruebas de error en las etapas de entrenamiento y validación, se encontró que la configuración 3-9-1 era la que brindaba el mejor rendimiento tanto en los tiempos de entrenamientos, como en los errores obtenidos, entre las redes de una sola capa oculta.

Para la configuración de dos capas ocultas se escogieron los 10 mejores resultados de las 81 combinaciones posibles, los cuales se pueden apreciar en las tablas 13 y 14.

Tabla 13. Errores de entrenamiento y tiempo consumido para redes de 2 capas ocultas

Errores de entrenamiento			
Configuración De la red	Error cuadrático Medio	Error absoluto Medio	Tiempo de Entrenamiento
3-7-9-1	0.0011	0.0204	1.2719 S
3-2-9-1	0.0016	0.0434	1.2562 S
3-2-7-1	0.0020	0.0261	1.2407 S
3-10-7-1	0.0021	0.0269	1.2454 S
3-6-4-1	0.0022	0.0270	1.1610 S
3-8-3-1	0.0029	0.0315	1.1032 S
3-10-6-1	0.0031	0.0328	1.2047 S
3-3-10-1	0.0044	0.0363	1.4487 S
3-5-10-1	0.0061	0.0383	1.3452 S
3-9-7-1	0.0072	0.0402	1.1748 S

Tabla 14. Errores de validación para redes de 2 capas ocultas

Errores de Validación		
Configuración de la red	Error cuadrático Medio	Error absoluto Medio
3-8-5-1	0.0014	0.0112
3-3-5-1	0.0019	0.0126
3-7-9-1	0.0014	0.0117
3-3-6-1	0.0015	0.0131
3-3-10-1	0.0016	0.0131
3-2-3-1	0.0017	0.0131
3-4-4-1	0.0015	0.0133
3-6-4-1	0.0015	0.0134
3-10-10-1	0.0016	0.0134
3-9-10-8	0.0018	0.0137
3-2-8-1	0.0015	0.140

La red ganadora en la fase de entrenamiento fue aquella con la configuración de 3-7-9-1, porque presentaba los menores errores (cuadrático y absoluto) y tiempo de ejecución. En la fase de validación se encontraron 2 redes 3-7-9-1 y 3-8-5-1, con el mismo error cuadrático y error absoluto parecido. En definitiva, se escogió la configuración 3-7-9-1 como la representativa de redes con 2 capas ocultas, por tener el mejor rendimiento en las etapas de entrenamiento y validación.

El tiempo total que gastó la rutina en encontrar la mejor topología variaba en cada ejecución, debido a que los valores de inicialización de los pesos eran aleatorios, y por consiguiente, se inicializaban redes diferentes que alcanzaban la meta de error en distintos tiempos. En general el tiempo total de procesamiento de la rutina se consideró alto, debido a la gran cantidad de

operaciones realizadas y datos almacenados. En la Tabla 15, se puede apreciar estos tiempos para diferentes ejecuciones de la rutina.

Tabla 15. Tiempo consumido en pruebas para configuraciones de 1 y 2 capas ocultas

Nº capas ocultas	Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5
1 capa	133.906 S	109.634 S	116.032 S	104.703 S	114.297 S
2 capas	932.402 S	1104.359 S	976.216 S	1038.740 S	913.051 S

Para determinar el número de capas a utilizar, se realizaron pruebas en matlab 6.5 con las mejores configuraciones de 1 y 2 capas ocultas, buscando obtener una respuesta razonable que tuviera en cuenta el compromiso entre error cuadrático medio y tiempo de entrenamiento.

Tabla 16. Prueba para número de capas según error cuadrático medio

Arquitectura de red	# de épocas	Tiempo de entrenamiento (S)	Error cuadrático medio ⁴⁷
3-9-1	10	0.6249	0.0027
3-7-9-1	10	0.9342	0.0029

Como puede verse, el aumento de las capas ocultas de una a dos no ofrece mejoría significativa, porque tanto los errores como los tiempos de entrenamiento se incrementan. Luego, una capa oculta fue suficiente para aproximar nuestro problema.

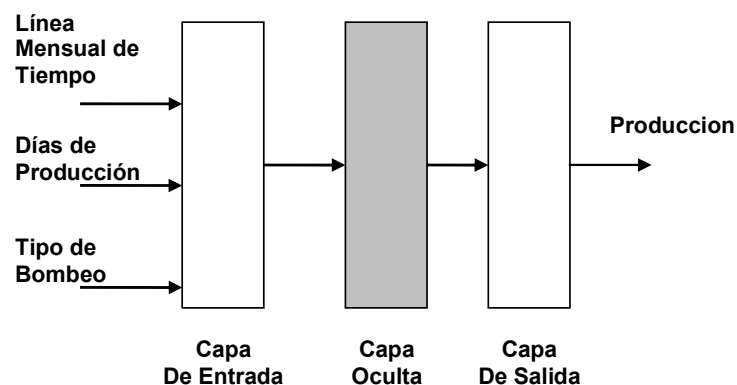
⁴⁷ El error cuadrático medio fue calculado en base a los datos suministrados a la red, los cuales pasaron por un proceso de normalización y ajuste, que se explicó en la sección 3.2.3, correspondiente al preprocesamiento de la información.

Los parámetros utilizados en las pruebas anteriores se resumen en la tabla 17 y se conservan constantes a excepción de aquél que será evaluado.

Tabla 17. Parámetros más importantes de las redes entrenadas

Algoritmo de entrenamiento	Levenberg-Marquardt
Función de activación capa de entrada	Sigmoidal
Función de activación capas ocultas	Sigmoidal
Criterio de desempeño	Error Cuadrático Medio
Función de activación capa de salida	Lineal
Número máximo de épocas	100
Error cuadrático medio límite	0.0001

Figura 16. Modelo de red con las variables de entrada y salida



3.5 FUNCION DE ACTIVACIÓN

Redes multicapa con funciones de activación no lineales permiten a la red aprender relaciones, tanto lineales como no lineales entre los vectores de entrada y salida. Luego, por la naturaleza no lineal de nuestro problema, se escogió la función de activación tipo sigmoidal para las neuronas de la capa de entrada y la capa oculta. Para la capa de salida, conformada por una sola

neurona, se escogió la función de activación lineal, que no restringe los valores de salida de la red a un intervalo específico.

3.6 ALGORITMO DE ENTRENAMIENTO

El algoritmo de aprendizaje de retropropagación del error (Backpropagation error) fue escogido para el proceso de entrenamiento de la red neuronal, debido a su versatilidad y número de variantes que por sus características, se adaptan a la mayoría de problemas.

Se realizaron pruebas en Matlab 6.5 con las diferentes variantes de éste algoritmo, utilizando los errores de entrenamiento como criterio de selección de aquella que brindara el mejor desempeño. El parámetro correspondiente al número máximo de épocas difería para cada variante, se flexibilizó de manera que pudieran medirse los errores y tiempo de entrenamiento para aquellos métodos de convergencia lenta, los cuales necesitan un número elevado de épocas para converger a valores de error admisibles, comparables a los de otros métodos.

La prueba para cada variación del backpropagation, se realizó repitiendo 10 veces el entrenamiento, al final de la prueba, se tomaron los menores errores y tiempos de entrenamiento. En la tabla N° 18 se muestran los resultados de la prueba realizada.

Tabla 18. Tiempo requerido en pruebas para configuraciones de 1 y 2 capas ocultas

Variante del Algoritmo de Retropropagación		Función de Matlab	Nº de épocas	Tiempo de Entrenamiento	Error Cuadrático Medio
Gradiente Descendiente	Sin Momentum	Traingd	2000	20.812 S	0.0592
	Con Momentum	Traingdm	1500	15.672 S	0.0513
Gradiente Conjugado		Traincgf	283	6.594 S	0.0414
		Traincgb	324	7.656 S	0.0326
		Traincgp	189	4.407 S	0.0375
		Trainscg	411	7.390 S	0.0338
Levenberg-Marquardt		Trainlm	35	1.125 S	0.0187
Quasi-Newton		Trainbfg	197	4.406 S	0.0242
		Trainoss	235	4.735 S	0.0381
Rata de Aprendizaje Adaptativa		Traingda	418	4.672 S	0.0602
		Traingdx	346	3.828 S	0.0425

Como se puede ver en la tabla anterior, se escogió el algoritmo de Levenberg-Marquardt para el entrenamiento de nuestro modelo, porque tiene el mejor desempeño, tanto en las pruebas de menor error cuadrático y absoluto, como en la de menor tiempo de entrenamiento. Los parámetros de las RNA utilizadas en las pruebas de desempeño se muestran en la tabla 19.

Tabla 19. Parámetros de las redes para las pruebas de desempeño

Clasificación	Denominación	Valor	Observación
Función de inicialización	Todas las Capas		Aleatoria
Funciones de activación	Capas Entrada y Oculta		Sigmoidal
	Capa de Salida		Lineal
Función de desempeño	Error cuadrático medio (ecm)		
Criterios de parada del entrenamiento	Ecm límite	0.001	Mínimo
	Factor de descenso	1e9	Máximo
	Gradiente	1e-8	Mínimo
	Número de fracasos	5	Máximo
	Tiempo de entrenamiento	Inf	máximo
	Épocas ⁴⁸	100	Máximo
Otros parámetros	Rata de aprendizaje	0.01	
	Valor inicial de μ	1e-3	

3.7 INICIALIZACIÓN DE PESOS

El método Levenberg – Marquardt minimiza una función utilizando de manera combinada el algoritmo de retropropagación y el método de la inversa de la matriz Hessiana. Se utiliza retropropagación cuando se está alejado del mínimo, y se cambia al segundo método a medida que la búsqueda se aproxima al mismo. Los métodos de entrenamiento mencionados utilizan el cálculo del gradiente para minimizar una función de error, luego, en ocasiones tienden a quedar atrapados en mínimos locales de la función objetivo. Esta situación no es deseable, porque la red deja de aprender creyendo que ya llegó a su objetivo, esto genera grandes errores en la fase de validación.

⁴⁸ Cómo se indicó en las pruebas anteriores, el número de épocas varía dependiendo de la variante del algoritmo de entrenamiento.

Básicamente el mínimo alcanzado está sujeto al conjunto de pesos iniciales que se empleen en el entrenamiento. No existen métodos para seleccionar los pesos iniciales que garanticen encontrar el mínimo global, en general, dichos pesos se generan aleatoriamente, por lo que es conveniente realizar varias sesiones de entrenamiento sobre la misma red con diferentes pesos, en forma tal de comenzar el aprendizaje desde diferentes puntos de la función objetivo y así tener más posibilidad de alcanzar el mínimo global.

Se han desarrollado algoritmos de inicialización de pesos, como el método de Nguyen-Widrow⁴⁹, que escoge valores para distribuir la región activa de cada neurona en la capa, aproximada y uniformemente sobre el espacio de entrada de dicha capa, para optimizar resultados y tiempos de entrenamiento. Se probó este algoritmo de inicialización, pero los resultados obtenidos no fueron satisfactorios, lo que llevó a que se utilizara otra técnica.

Se implementó entonces un procedimiento de inicialización de pesos utilizando la siguiente metodología:

1. Seleccionar aleatoriamente el valor del conjunto de pesos, especificando el número de iteraciones.
2. A partir de los datos de entrada y salida y con un número de iteraciones preestablecidas, simular la red con el conjunto de pesos anteriores y determinar el error cuadrático (SSE).
3. Almacenar el conjunto de pesos de la red que obtenga el SSE mínimo.
4. Repetir los pasos 1 al 3 hasta alcanzar el número de evaluaciones máximo, o hasta alcanzar una solución aceptable.

⁴⁹ Mathworkks Neural Networks Toolbox Chapter 14 P 14-89

5. Este proceso de inicialización se debe realizar como mínimo tres veces, lo que garantiza que los conjuntos de pesos iniciales para el entrenamiento sean diferentes y aleatorios permitiendo acercarse al mínimo global del modelo.
6. Luego de varias pruebas, seleccionar los diferentes conjuntos de pesos que garantizan la convergencia hacia un mínimo global, éstos se utilizarán en el entrenamiento definitivo de todos los pozos del yacimiento.

La colección de pesos, definida anteriormente, se ha identificado como buenos conjuntos para iniciar el entrenamiento de la red; además se obtuvo con base en pozos de morfologías diversas y se aplicó con resultados satisfactorios en todos los casos estudiados, donde cada uno de ellos alcanzó el mínimo global.

3.8 RESULTADOS DEL ENTRENAMIENTO

El conjunto de datos con los que se probará y entrenará la red definitiva varía en cantidad de un pozo a otro. En general, de la totalidad de registros, se asigna el 30% a validación de la red y el otro 70% a entrenamiento.

Para todos los pozos, durante el entrenamiento de la red, se tiene en cuenta el conjunto de pesos sugeridos en la fase de inicialización de pesos, enunciada en el ítem anterior, de la siguiente manera:

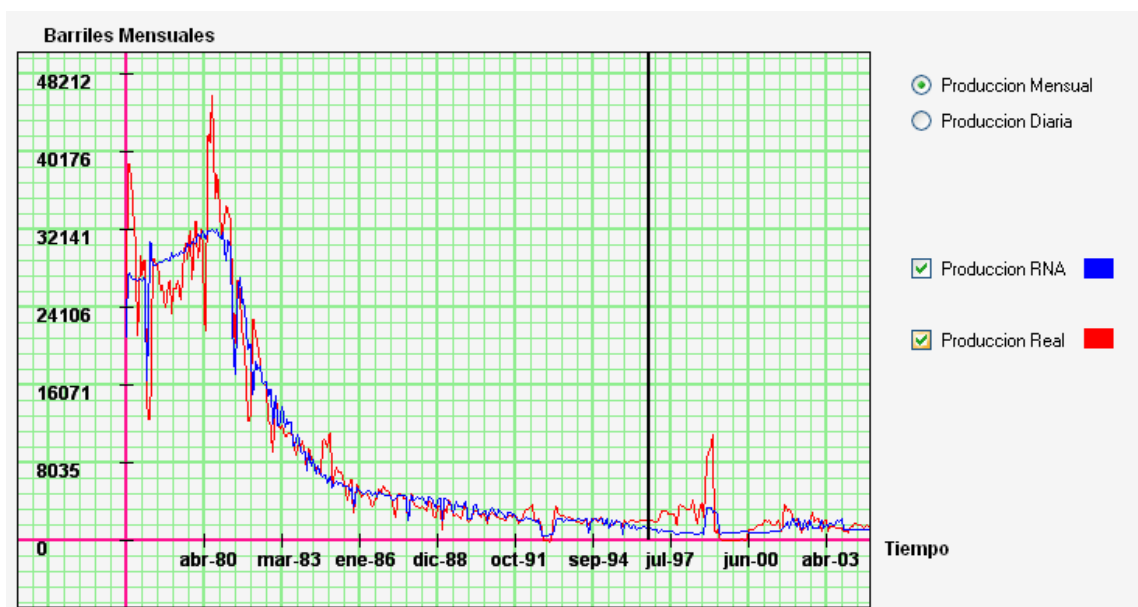
1. Se escoge la primera colección de pesos iniciales y se realiza el entrenamiento.
2. Si el error obtenido está dentro del rango permisible⁵⁰ se detiene la fase de entrenamiento, en caso contrario, se escoge el siguiente conjunto de pesos.

⁵⁰ Este rango se calculó de forma heurística y está definido por el error cuadrático. Se encuentra entre 1×10^{-3} y 1×10^{-4}

3. Si el nuevo conjunto tiene un desempeño superior al anterior, se guarda éste como el mejor, en otro caso se vuelve al paso 2.
4. Este procedimiento se repite hasta probar todos los conjuntos de pesos.
5. Si se han probado todos los conjuntos de pesos y no se encontró alguno que cumpliera con el criterio de desempeño, se procede a escoger el que menor error cuadrático obtuvo, el cual es guardado en el paso 3.

Los resultados de entrenamiento para la configuración red final, 3-9-1, se pueden observar en la fig. 12, donde se compara el vector histórico de producción⁵¹ contra la salida de la red. Los errores de esta prueba, se aprecian en la tabla 20.

Figura 17. Resultados del entrenamiento



⁵¹ Un informe más detallado de los resultados de la simulación para los diferentes pozos de estudio, se pueden apreciar en el capítulo 5, correspondiente al análisis de resultados.

Tabla 20. Errores de entrenamiento para la configuración definitiva de la red

	Error Cuadrático Medio	Error Absoluto Medio
Entrenamiento	0.00175	0.0235

Las pruebas hechas en este capítulo fueron realizadas en un equipo con las siguientes especificaciones de hardware y software:

- Procesador: AMD de 2.1 GHz.
- Memoria RAM :704 Mb
- Sistema Operativo: Windows Xp SP2.
- Matlab 6.5

CAPITULO 4

METODOLOGÍA DE DESARROLLO

Después de analizar y comparar los modelos de ciclo de vida del desarrollo de software, para la elaboración de este proyecto, se decidió utilizar el modelo de prototipado evolutivo, ligeramente modificado, que a partir de una concepción inicial del software a desarrollar, puede incorporar nuevas ideas enmarcadas en prototipos, que contribuyan a construir un sistema más confiable, preciso y robusto, que cumpla con los objetivos trazados en este proyecto. El prototipado, combinado con los fundamentos de la metodología GRAPPLE, brinda flexibilidad y la oportunidad de adaptar e incluir ideas según las necesidades del analista o desarrollador

El plan de trabajo está fundamentado en una serie de iteraciones formadas por actividades de la metodología GRAPPLE, que garantizan que en cada ciclo se obtenga una versión del software más cercana a la final.

Las actividades a realizar tienen la ventaja de no ser secuenciales, luego es posible regresar a un punto anterior para alimentarse de nuevas ideas o requerimientos. Las actividades son:

1. Recopilación de necesidades
2. Análisis
3. Diseño
4. Desarrollo

4.1 RECOPIACIÓN DE NECESIDADES

En esta etapa se realizará un estudio para determinar lo que el software debe estar en capacidad de hacer. Para ello se cuenta con las siguientes tareas:

- Realizar un análisis del dominio y de los términos
- Descubrir las necesidades del sistema y de los usuarios
- Confrontar resultados

4.2 ANÁLISIS

En esta parte se profundizan los resultados obtenidos de la actividad anterior, para aumentar la comprensión del problema; de hecho es posible que porciones del análisis empiecen durante la recopilación de necesidades.

Para llevar a cabo el análisis se cuenta con las siguientes tareas:

- Comprensión del uso del sistema.
- Descripción de los casos de uso.
- Enumeración de las clases necesarias.
- Integración de los conceptos anteriores.

4.3 DISEÑO

En este segmento se trabajará con los resultados del Análisis para diseñar la solución. En el Diseño y Análisis se harán las revisiones necesarias hasta que el diseño se haya completado. Las actividades de esta sección son:

- Análisis de operaciones y diagrama de actividades.
- Planteamiento de la distribución.
- Diseño y prototipos de la interfaz de usuario.
- Pruebas de diseño.
- Iniciar la documentación del sistema.

4.4 DESARROLLO

En este segmento se lleva a cabo la programación; con suficiente Análisis y Diseño, el cual, debería realizarse con relativa rapidez y pocos problemas. Las actividades correspondientes son:

- Generación del código.
- Verificación del código.
- Generación de interfaz de usuario.
- Conexión de la interfaz con el código.
- Pruebas.
- Finalización de la documentación.

Cada prototipo se someterá a evaluaciones y pruebas con el fin de determinar su validez. En el transcurso de esta etapa y gracias a los resultados de las pruebas realizadas se descubrirán fallas que deben ser corregidas, y se harán evidentes nuevas necesidades susceptibles de estudio, análisis y diseño para ser posteriormente implementadas.

En la aplicación de la metodología puede que no se realicen todas las actividades anteriormente descritas, y en algunos casos, dependiendo del estado de desarrollo del proyecto, se le dará más importancia a algunas.

4.5 HERRAMIENTA DE DESARROLLO

A pesar de que existen herramientas especializadas en el manejo de redes neuronales y demás funciones matemáticas, como Matlab, prescindimos de su uso por las siguientes razones⁵²:

⁵² El hecho de no usar Matlab como herramienta de desarrollo no implica que no se haya utilizado en la realización del proyecto. La forma como se usó matlab para definir la arquitectura de red y el algoritmo de entrenamiento se encuentra descrito en el capítulo 3, secciones 4 al 6.

- El paquete de servicios que ofrece Matlab es muy amplio y en el desarrollo de nuestro trabajo sólo necesitamos algunas funcionalidades de unos pocos paquetes específicos; por ello no se justifica la inversión en esta herramienta.
- Utilizar Matlab como herramienta de desarrollo, implica que necesariamente éste deba estar instalado en la máquina donde se desee utilizar nuestra aplicación; esto afecta notoriamente la portabilidad del sistema.
- Si bien Matlab provee algunos controles para interfaces gráficas, estos se quedan cortos comparados con los ofrecidos por entornos especializados de programación.
- Otras herramientas como NEURO SOLUTIONS⁵³ son especializadas en el desarrollo de arquitecturas de redes neuronales, y se caracterizan por ser poco personalizables, es decir, no ofrecen ciertas características específicas de nuestro problema a solucionar, entre ellas, una interfaz amigable y una presentación de resultados dinámica..
- Adicionalmente y de mayor importancia, la razón que nos motivó a utilizar un entorno de desarrollo basado en lenguajes de programación, es la de poder construir, totalmente, la aplicación utilizando la programación orientada a objetos, el manejo de funciones de aproximación numérica, funciones de graficación y otras técnicas aprendidas durante la carrera de Ingeniería de Sistemas; de tal forma que la aplicación desarrollada contribuya a la ingeniería del software nacional y sea planteada como una nueva alternativa frente a paquetes comerciales que aparte de sus elevados costos tienen la desventaja de funcionar como cajas negras.

Por las razones descritas con antelación, unido a los convenios de la Universidad Industrial de Santander con la casa matriz Microsoft para utilizar software de última generación mediante licencia Campus Agreement, decidimos utilizar como entorno de desarrollo Visual Studio 2005,

⁵³ Para mayor información, véase <http://www.nd.com/> , revisado el 7 de junio de 2007.

específicamente el lenguaje de programación Visual Basic .Net. Este lenguaje de alto nivel, a diferencia de su predecesor, Visual Basic 6.0, es totalmente orientado a objetos, pero conservando la simplicidad de su sintaxis; además posee gran cantidad de componentes para un avanzado diseño de la interfaz de usuario, manipulación de gráficos, archivos y estructuras para el manejo de excepciones y todas las funcionalidades que ofrece el Framework 2 de la tecnología .Net.

4.6 APLICACIÓN DE LA METODOLOGÍA

4.6.1 Prototipo 1

Dado el inicial desconocimiento del campo de acción que requería el desarrollo de la investigación, en este prototipo, las acciones predominantes son las que conducen a la comprensión del dominio y de términos. Es por eso que en el primer paso de la construcción del software, la mayor parte del tiempo se invirtió en recopilar información que condujera a un mayor entendimiento, tanto de las necesidades del sistema, como de los factores a tener en cuenta para un buen desarrollo de un modelo de red neuronal. Para ello se consultó en libros, artículos de revistas especializadas, tesis de grado y entrevistas informales con algunos profesionales del área del petróleo.

Para las pruebas y validaciones futuras del sistema, es necesario contar con archivos históricos de producción, que se consideran la materia prima de nuestro proyecto. Es por este motivo que se hace necesario conseguir datos de gran calidad, para este propósito, se hizo una intensa consulta en diferentes medios hasta hallar tal información, la cual no se detalla por motivos de confidencialidad.

De igual forma también fue necesario consultar información de las redes neuronales y sus aplicaciones, para ganar experiencia en el tema y aplicar metodologías y técnicas, que conduzcan a la selección de la aquella que mejor se ajustara a nuestras necesidades, tal como se describió en el capítulo 3.

Dentro de las tareas desarrolladas en este prototipo, se destaca la posibilidad de leer y almacenar la información en archivos planos, adicionalmente y de igual importancia, se incluyó el preprocesamiento de la información leída. Para ello se desarrolló una primera versión de la clase cDatos, que precisamente contaba con las funciones básicas de lectura, de preprocesamiento y de escritura en archivo plano como medio de salida.

Después de leer la información, y antes de obtener las salidas, es necesario que los datos pasen por la red neuronal que en síntesis es el corazón del sistema y es la que nos permite hacer las predicciones y simulaciones. Para tal fin, se implementó la primera versión de la clase cRed y cCapa, con las cuales se puede armar la arquitectura de red deseada, basadas inicialmente en el algoritmo de entrenamiento conocido como GDX, que corresponde al método de backpropagation simple.

A pesar de que el principal objetivo de este prototipo era la profundización y comprensión de los procesos necesarios, es una versión medianamente funcional, cuya interfaz tiene la finalidad de interactuar con las funciones básicas, como entrenamiento de la RNA, para probar su correcto funcionamiento, y no para que un usuario interactúe directamente con el sistema.

La funcionalidad de este primer prototipo se resume así:

- La clase cDatos tenía como objetivos traer los datos desde archivo hasta la aplicación e imprimir la salida mediante archivo.
- Aplicar algún algoritmo que permitiera filtrar los datos.
- La clase cRed tenía como objetivo construir eficazmente una arquitectura de red neuronal.
- Implementar un algoritmo de entrenamiento para la red.

En términos generales el primer prototipo hizo lo que se esperaba, a excepción del algoritmo de entrenamiento, que no ajustaba los pesos porque este método

era poco eficiente para aproximar el problema en cuestión⁵⁴ y además las funciones de filtrado estaban en su fase de mejoramiento.

4.6.2 Prototipo 2

En este segundo prototipo se mejoraron varias funciones, algunas a nivel de funcionamiento y otras a nivel visual. La primera de estas mejoras es respecto a la lectura de los datos: El prototipo 1 suponía que el usuario debía migrar la información desde Excel hasta un archivo plano y organizarla por columnas, para que el sistema pudiera leerla adecuadamente. En este prototipo, apoyados en la tecnología ADO.NET, se implementó una función dentro de la clase cDatos que permitiera leer la información directamente de una hoja de Excel. Esta es una gran ventaja, pues la información original está consignada en archivos de Excel, que es una herramienta de amplio uso y de fácil manejo.

Seguidamente, para guardar las similitudes de presentación entre los datos originales y los arrojados por el sistema, estos se muestran mediante estructuras avanzadas de muy buena visualización y fácil manejo, como lo son los controles DataGrid. Adicionalmente en la clase cDatos se implementaron nuevas funciones para el filtrado y preprocesamiento de los datos. Las funciones de lectura y escritura evolucionaron tal que permitieran manejar la información de una forma sencilla, organizada y estructurada.

Por su parte la clase cRed fue mejorada al introducirle nuevas funciones que dieran soporte al nuevo algoritmo de entrenamiento Levenberg-Marquadt; por ello se profundizó en el manejo de operaciones con vectores y matrices que se utilizan en los datos previamente preprocesados. Además, se incluyeron funciones robustas para calcular numéricamente el Jacobiano y las derivadas parciales, que eran necesarias para la ejecución correcta del método de entrenamiento.

⁵⁴ Las pruebas realizadas con el algoritmo GDX se describen en el capítulo anterior.

En resumen, la funcionalidad del prototipo 2 es la siguiente:

- La clase cDatos tenía como objetivo la lectura de datos desde Excel y la implementación de nuevos algoritmos para el preprocesamiento de la información.
- La clase cRed incluye nuevas funcionalidades, como la implementación del algoritmo de entrenamiento de Levenberg-Marquadt.
- Mejoras visuales en la presentación de la información al usuario en cuanto a los datos históricos de producción y a los resultados de la simulación.

En este prototipo todas las expectativas fueron cubiertas, el programa ya brindaba resultados satisfactorios, pero cabe anotar que la interfaz estaba aún en crecimiento.

4.6.3 Prototipo 3

En esta etapa del desarrollo, el programa ya era bastante funcional y sólo restaba mejorar la visualización de los resultados; para ello se implementó la nueva clase cGrafica, para presentar los datos mediante gráficas 2D dentro de la aplicación, y con esto mostrar de forma más clara y diciente el análisis de los resultados obtenidos. También se crearon funciones para el análisis de resultados como el cálculo del límite económico, visualización de la tasa de declinación, producción acumulada, entre otras.

En la clase cDatos, se creó una nueva función que permitía exportar los datos de la simulación a una hoja de Excel, para su posterior manipulación.

Asimismo la interfaz pasó de ser una consola primitiva y un conjunto de tablas, a una aplicación totalmente amigable con muchas opciones para el usuario, con un tratamiento riguroso de las excepciones, reduciendo al mínimo las

posibilidades de error, informando oportunamente las operaciones indebidas junto a la sugerencia de solución respectiva en cada caso.

Para la presentación gráfica, la simulación y consultas se emplearon los controles apropiados para guiar al usuario hacia lo que desee hacer.

4.7 DIAGRAMAS UML

Del desarrollo de la metodología anterior surgen para el prototipo 3 los siguientes diagramas UML

4.7.1 Diagrama de Casos de Uso

Tabla 21. Casos de uso: Cargar información

Cargar Información	
Actor	Usuario
Propósito	Cargar en la aplicación los datos históricos de producción de un pozo petrolero.
Resumen	Lee un archivo de Excel y lo carga en el sistema.
Flujo Principal	<ol style="list-style-type: none"> 1. El actor selecciona el archivo deseado. 2. El sistema tabula los datos y se los muestra al usuario. 3. El sistema carga la información en sus estructuras. 1a: El actor cancela el proceso El sistema regresa al módulo principal 1b: El actor carga un archivo con información no válida o con el formato incorrecto. El sistema Presenta un mensaje de error y retorna al módulo principal
Precondición	Ninguna

Tabla 22. Casos de uso: Entrenar RNA

Entrenar RNA	
Actor	Usuario
Propósito	Ajustar los pesos de la RNA.
Resumen	Aplica el algoritmo de entrenamiento Levenberg-Marquardt para un conjunto de datos, y establece el valor de los pesos.
Flujo Principal	<ol style="list-style-type: none"> 1. El actor selecciona el volumen de datos para el entrenamiento. 2. El sistema ajusta sus pesos mediante el algoritmo de entrenamiento. 3. El sistema asigna los resultados del entrenamiento a los respectivos módulos del sistema. <p>1a: El actor no selecciona ningún volumen de datos. El sistema toma por defecto el 70 % del total de datos para el entrenamiento.</p>
Precondición	Debe estar cargado el conjunto de datos de algun pozo específico.

Tabla 23. Casos de uso: Simular RNA

Simular RNA	
Actor	Usuario
Propósito	Simular una RNA de acuerdo a los datos de simulación suministrados.
Resumen	Simular el comportamiento de la producción de un pozo.
Flujo Principal	<p>1. El actor escoge el rango de fechas de simulación.</p> <p>2. El sistema procesa los datos a través de la red neuronal.</p> <p>3. El sistema presenta en pantalla los resultados de la simulación.</p> <p>1a: El actor selecciona un rango de fechas erróneo.</p> <p>El sistema Presenta un mensaje de error y retorna al módulo principal</p> <p>1b: El actor no selecciona ningún rango de fechas.</p> <p>El sistema toma por defecto las fechas de inicio y finalización presentes en el archivo de producción cargado anteriormente.</p>
Precondición	Debe estar cargado el conjunto de datos de producción, y la RNA debe estar entrenada.

Tabla 24. Casos de uso: Curva de declinación

Curva de Declinación	
Actor	Usuario
Propósito	Mostrar la curva de declinación.
Resumen	Genera la curva de declinación, bien sea la real y/o la simulada.
Flujo Principal	<p>1. El actor escoge el tipo(s) de visualización.</p> <p>2. El sistema, muestra la visualización escogida.</p>
Precondición	Debe estar cargado el conjunto de datos de producción, la RNA debe haber sido entrenada y simulada.

Tabla 25. Casos de uso: Analizar resultados

Analizar Resultados	
Actor	Usuario
Propósito	Mostrar información relevante del pozo simulado.
Resumen	Mediante gráficas y cifras se presentan los resultados analizados.
Flujo Principal	<ol style="list-style-type: none"> 1. El actor escoge el tipo de análisis. 2. El sistema, en una nueva ventana, genera las gráficas y datos pertinentes. . 3. El actor interactúa con las diferentes opciones de visualización. 4. El actor cierra la ventana de análisis. 5. El sistema regresa al módulo principal.
Precondición	Debe estar cargado el conjunto de datos de producción, la RNA debe haber sido entrenada y simulada.

Tabla 24. Casos de uso: Producción acumulada

Producción acumulada	
Actor	Usuario
Propósito	Mostrar la curva de producción acumulada.
Resumen	Genera la curva de producción acumulada, bien sea la real y/o la simulada.
Flujo Principal	<ol style="list-style-type: none"> 1. El actor escoge el tipo(s) de visualización. 2. El sistema, muestra la visualización escogida.
Precondición	Debe estar cargado el conjunto de datos de producción, la RNA debe haber sido entrenada y simulada.

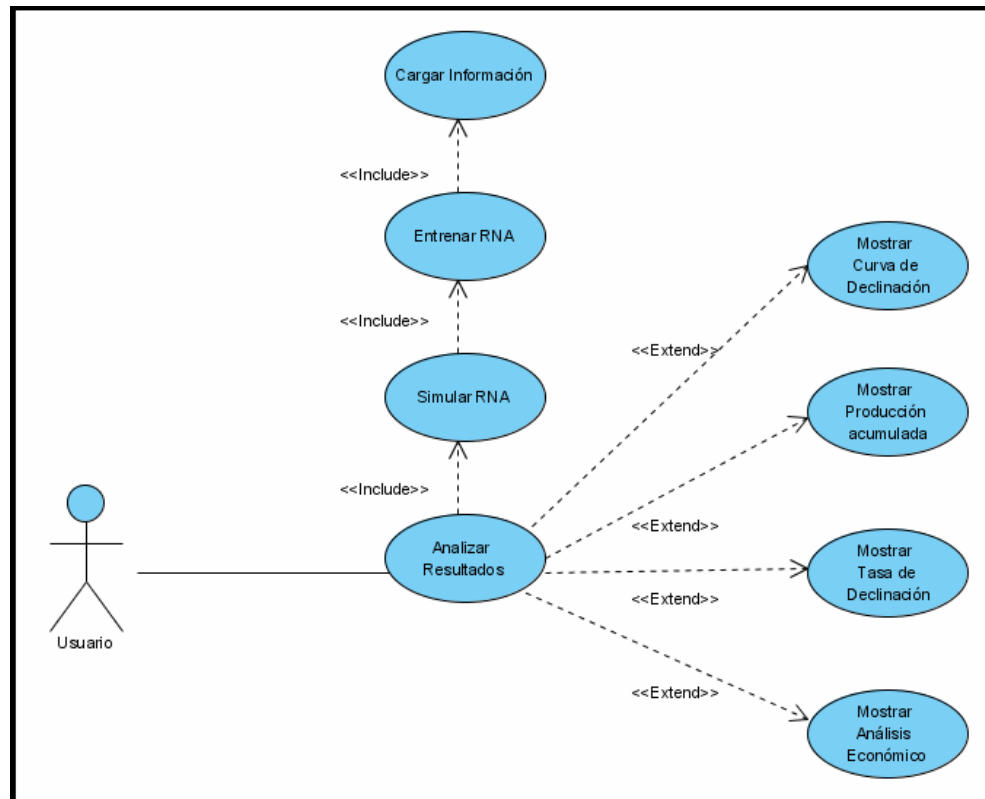
Tabla 25. Casos de uso: Tasa de declinación

Tasa de Declinación	
Actor	Usuario
Propósito	Construir la gráfica de la variación de la tasa de declinación.
Resumen	El sistema evalúa las variaciones en la tasa de declinación, dada una periodicidad, y la muestra en una gráfica.
Flujo Principal	<ol style="list-style-type: none"> 1. El actor escoge la periodicidad. 2. El sistema genera la grafica correspondiente. <p>1a: El actor no selecciona ninguna periodicidad. El sistema toma por defecto una periodicidad de 12 meses.</p>
Precondición	Debe estar cargado el conjunto de datos de producción, la RNA debe haber sido entrenada y simulada.

Tabla 26. Casos de uso: Análisis económico

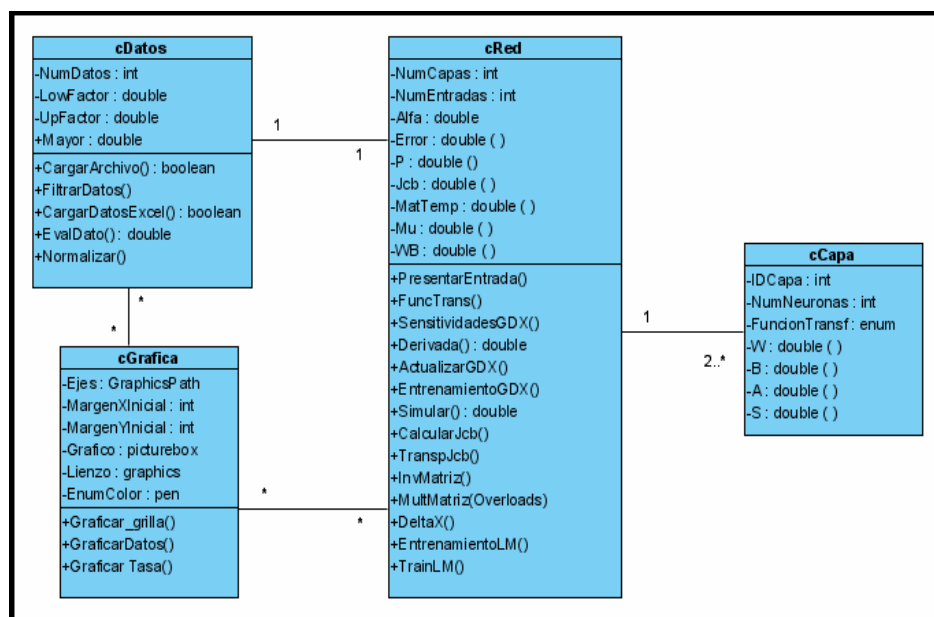
Análisis Económico	
Actor	Usuario
Propósito	Brindar la información económica de la simulación..
Resumen	El sistema determina el límite económico.
Flujo Principal	<ol style="list-style-type: none"> 1. El actor establece el costo de operación del pozo. 2. El usuario establece el precio del barril de petróleo. 3. El sistema genera la gráfica e identifica el punto donde es alcanzado el límite económico. <p>1a: El actor ingresa datos erróneos. El sistema Presenta un mensaje de error y retorna al módulo actual.</p>
Precondición	Debe estar cargado el conjunto de datos de producción, la RNA debe haber sido entrenada y simulada.

Figura 18. Diagrama de casos de uso



4.7.2 Diagramas de Clases

Figura 19. Diagrama de Clases



4.7.2.1 Diccionario de Datos

- **Clase cCapa:** Es una clase sencilla que no tiene métodos, pero los miembros son sumamente importantes para construir las capas dentro de las redes neuronales. Esta clase describe una capa cualquiera de una red neuronal y se compone así:

Tabla 27. Miembros principales de la clase cCapa

Nombre	Tipo de datos	Descripción
IDCapa	Entero	Identificador de la capa.
NumNeuronas	Entero	Número de neuronas presentes en la capa
FuncionTransf	Enumerado	Tipo de función de transferencia
W(,)	Vector Doble	Vector de pesos
B()	Vector Doble	Vector de ganancias
S()	Vector Doble	Vector de sensibilidades
A()	Vector Doble	Vector de salidas

- **Clase cDatos:** Es la clase que contiene los métodos y miembros necesarios para realizar todo el manejo de los datos, correspondiente a lectura, escritura y preprocesamiento.

Tabla 28. Miembros principales de la clase cDatos

Nombre	Tipo de datos	Descripción
NumDatos	Entero	Número de filas de entrenamiento
LowFactor	Doble	Proporción mínima admitida en datos atípicos
UpFactor	Doble	Proporción máxima admitida en datos atípicos
Mayor	Doble	El mayor valor de los datos de producción.

Tabla 29. Métodos principales de la clase cDatos

Nombre	Devuelve	Finalidad
CargarArchivo	Booleano	Cargar los datos desde archivos planos.
FiltrarDatos		Filtrar los datos, corrigiendo los atípicos.
CargarDatosExcel	Booleano	Cargar los datos desde archivos de Excel.
EvalDato	Doble	Devuelve el valor que reemplaza al dato atípico.
Normalizar		Normalizar los datos a valores entre 0 y 1.

- **Clase cRed:** Es la clase que nos permite construir las redes neuronales, hacer el entrenamiento y la simulación.

Tabla 30. Miembros principales de la clase cRed

Nombre	Tipo de datos	Descripción
NumCapas	Entero	Número de capas de la red.
NumEntradas	Doble	Número de entradas.
Alfa	Doble	Rata de aprendizaje.
P(,)	Vector Doble	Vector de Entrenamiento.
Error()	Vector Doble	Almacena los errores de los entrenamientos.
Jcb(,)	Vector Doble	Matriz jacobiana del algoritmo Lev-Mar.
MatTemp(,)	Vector Doble	Matriz que almacena los DeltaX en Lev-Mar.
Mu	Doble	Factor de entrenamiento.
WB()	Vector Doble	Vector de pesos y ganancias

Tabla 31. Métodos principales de la clase cRed

Nombre	Devuelve	Finalidad
PresentarEntrada		Propagar una entrada a través de las capas, se obtienen las salidas.
FuncTrans		Evaluar y asignar el valor de la función de transferencia correspondiente.
SensitividadGDX		Calcular las sensitividades para el entrenamiento GDX.
Derivada	Doble	Evaluar la derivada de la función de transferencia.
ActualizarGDX		Actualizar los parámetros (W, B) para GDX.
EntrenamientoGDX		Entrenamiento tipo GDX
Simular	Doble	Simular a partir de unos datos de entrada.

Nombre	Devuelve	Finalidad
TranspJcb		Calcular la transpuesta de la matriz jacobiana.
InvMatriz		Calcular la inversa de una matriz parámetro
MultMatriz (Overloads)		Multiplicar dos matrices
Deltas		Obtener los nuevos pesos y ganancias que arroja el algoritmo L-M.
EntrenamientoLM		Entrenar a la red neuronal mediante Lev-Mar
TrainLM		Escoger el mejor vector de pesos iniciales.

- **Clase cGrafica:** Esta clase se encarga de todas las rutinas gráficas que permiten la correcta visualización de los resultados obtenidos.

Tabla 32. Miembros principales de la clase cGráfica

Nombre	Tipo de datos	Descripción
Ejes	GraphicsPath()	Objeto para dibujar los ejes coordenados.
MargenXInicial	Entero	Distancia del borde en que se dibuja el eje X.
MargenYInicial	Entero	Distancia del borde en que se dibuja el eje Y.
Grafico	PictureBox	Control en que se dibuja la gráfica.
Lienzo	Graphics	Objeto que soporta las funciones de graficación.
EnumColor	Pen	Objeto que especifica el color y grosor del trazo.

Tabla 33. Métodos principales de la clase cGraficas

Nombre	Devuelve	Finalidad
graficar_grilla		Dibujar la cuadrícula de la gráfica.
Graficar		Dibujar la gráfica en el lienzo especificado.
GraficarTasa		Graficar la tasa de declinación.
Derivada	Doble	Evaluar la derivada de la función de transferencia.

4.7.3 Diagrama de Secuencias

Figura 20. Diagrama de secuencias: Cargar datos, entrenar y simular la RNA

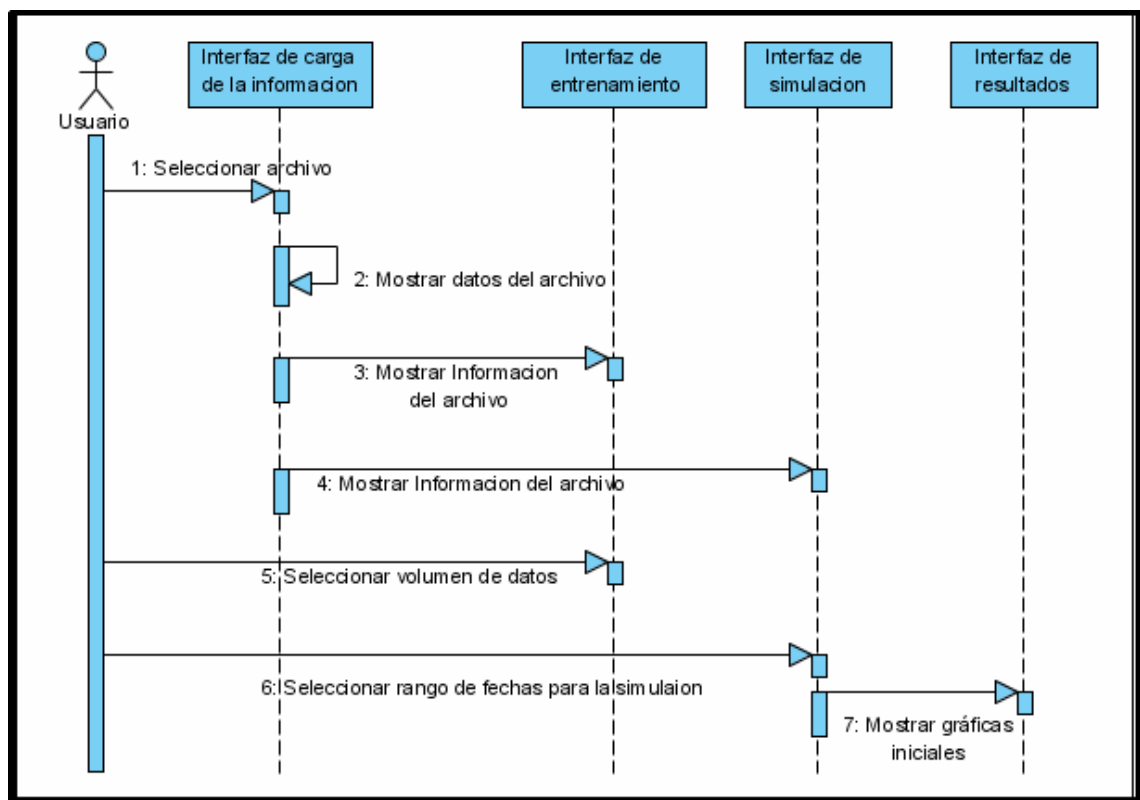


Figura 21. Diagrama de secuencias: Curva de declinación

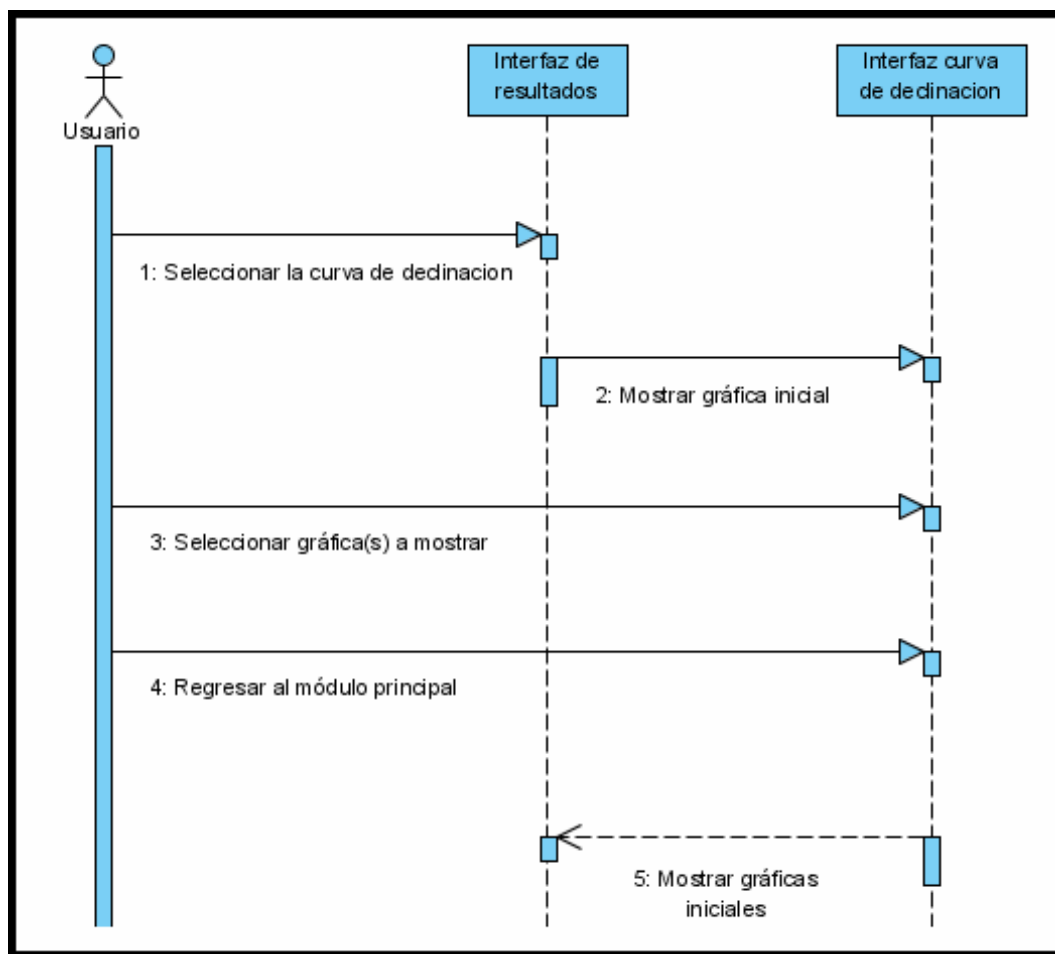


Figura 22. Diagrama de secuencias: Producción acumulada

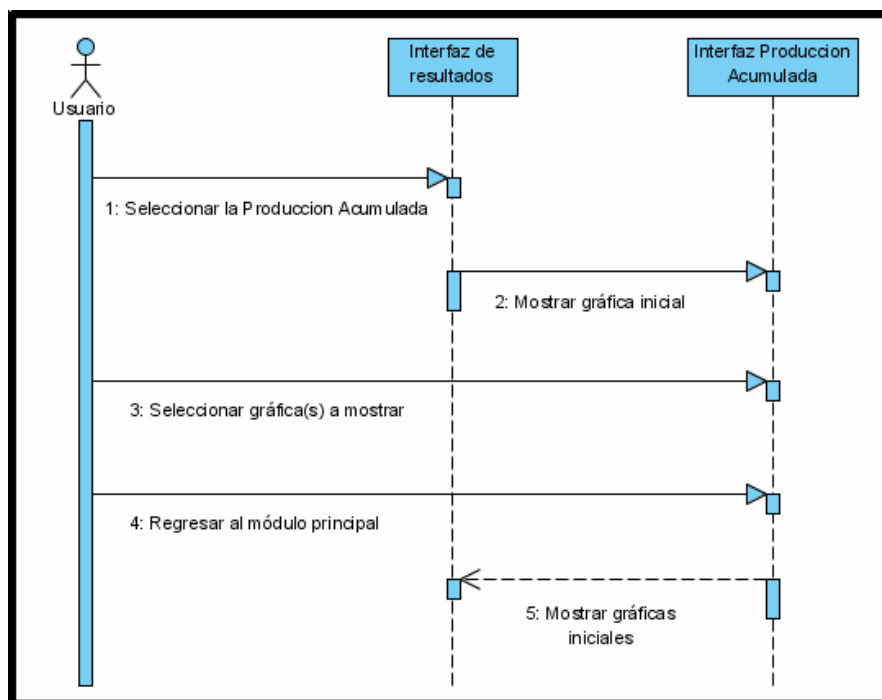


Figura 23. Diagrama de secuencias: Tasa de declinación

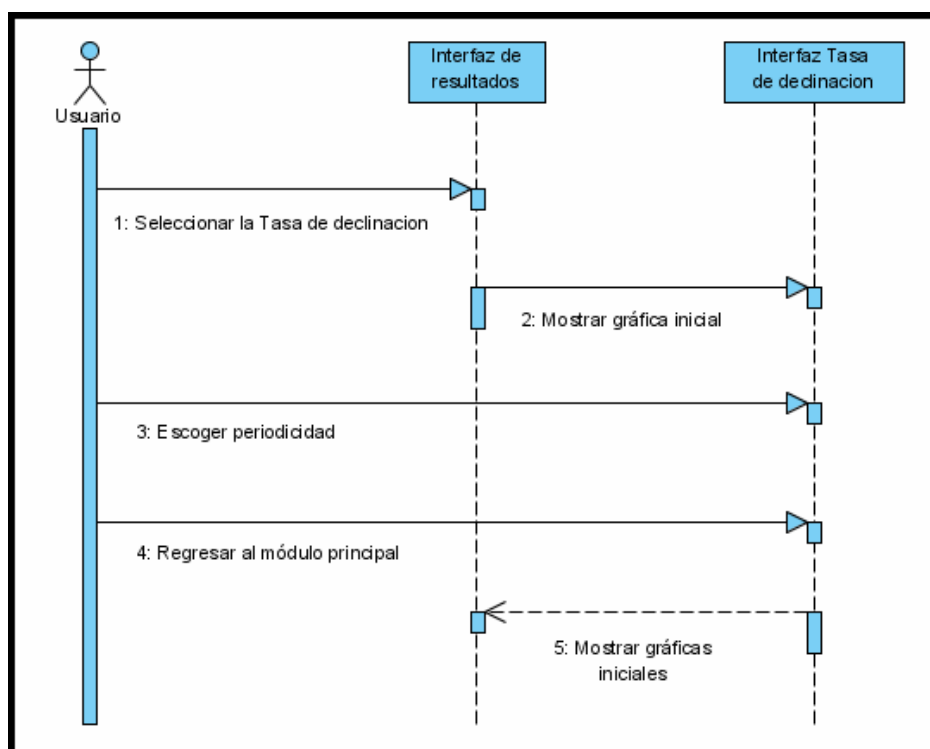
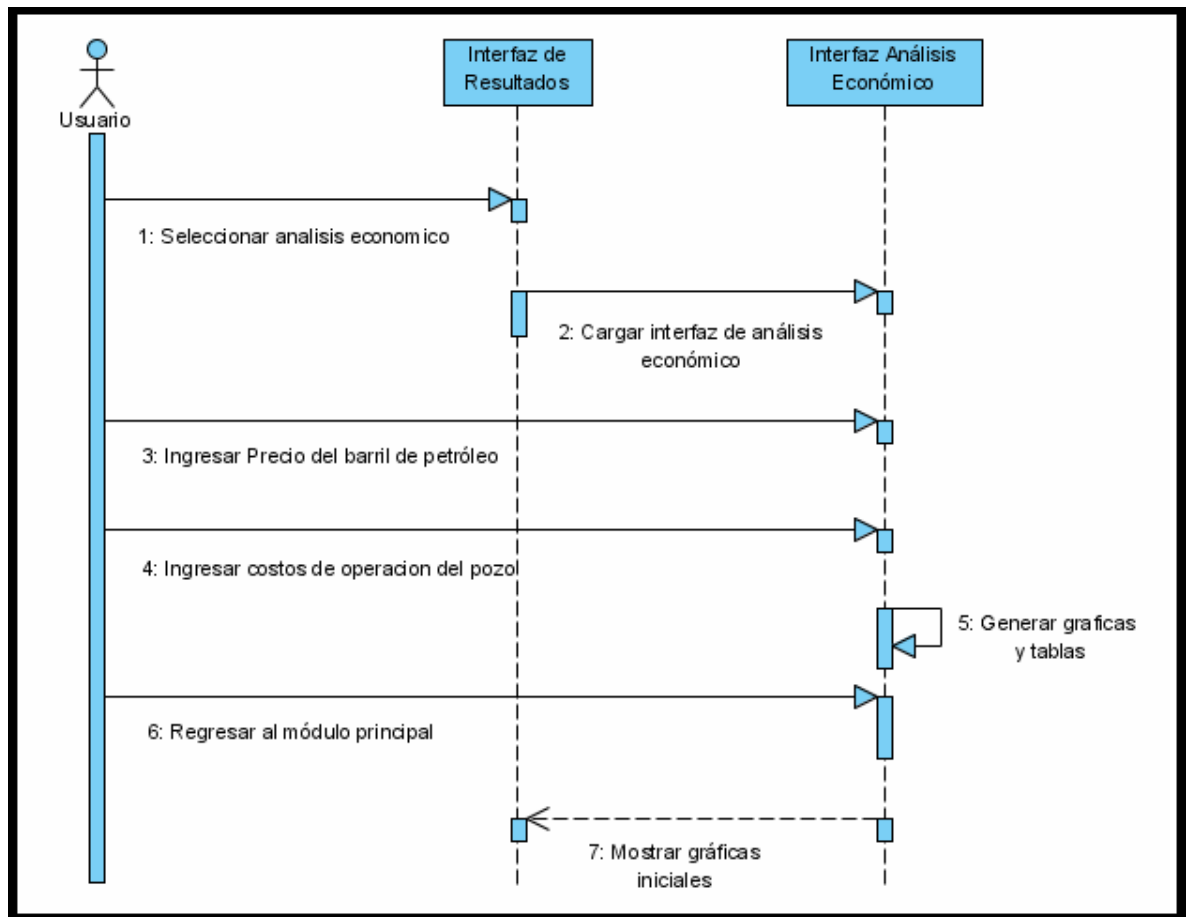


Figura 24. Diagrama de secuencias: Análisis Económico



CAPITULO 5

ANÁLISIS DE RESULTADOS

Para la realización de las distintas comparaciones y gráficos entre la herramienta DECLINET y las técnicas clásicas de análisis de curvas de declinación, se utilizaron 8 pozos, pertenecientes a un mismo yacimiento, cada uno con características individuales. Estas comparaciones incluyen el error absoluto promedio y la tasa de ajuste obtenidos durante la fase de predicción de la producción futura. La tasa de ajuste es un indicador que muestra el porcentaje de semejanza entre los datos reales de producción y los valores simulados ya sea por la RNA o por las técnicas clásicas, y está basado en el error relativo promedio obtenido durante la etapa de predicción y está dado por:

$$Ta = 100 - \frac{\sum_{i=1}^n \frac{|P_i - Ps_i|}{P_i}}{n} 100\% \quad (5.1)$$

Donde P_i : Producción real

Ps_i : Producción simulada.

n: Numero de meses simulados.

La implementación de los métodos para calcular la declinación hiperbólica, armónica y exponencial, y sus respectivas gráficas, fue desarrollada en matlab 6.5. El parangón entre estas técnicas y DECLINET contra los datos reales incluyen cuatro gráficas por cada pozo: la primera y segunda gráfica muestran la producción original frente a la predicción de las técnicas clásicas y de DECLINET respectivamente; la tercera y cuarta, por su parte, muestran la producción real acumulada frente a los acumulados de las técnicas clásicas y

de DECLINET; esta separación se hizo con el objeto de apreciar mejor los resultados con cada técnica.

5.1 POZO 1

Inicio de la producción: Febrero de 1977.

Final de producción: Diciembre de 2004.

Total de datos: 334

Tabla 34. Error absoluto promedio y tasa de ajuste de cada método para el pozo 1

	Armónico	Exponencial	Hiperbólico	RNA
Error Absoluto Promedio ⁵⁵	5286.72	6602.95	6857.03	2556.46
Tasa de ajuste (%)	56.15	44.52	40.84	71.81

⁵⁵ El error absoluto esta dado en barriles de petróleo (BLS) por mes,

Figura 25. Pozo 1: Comparación entre las técnicas clásicas y la producción original

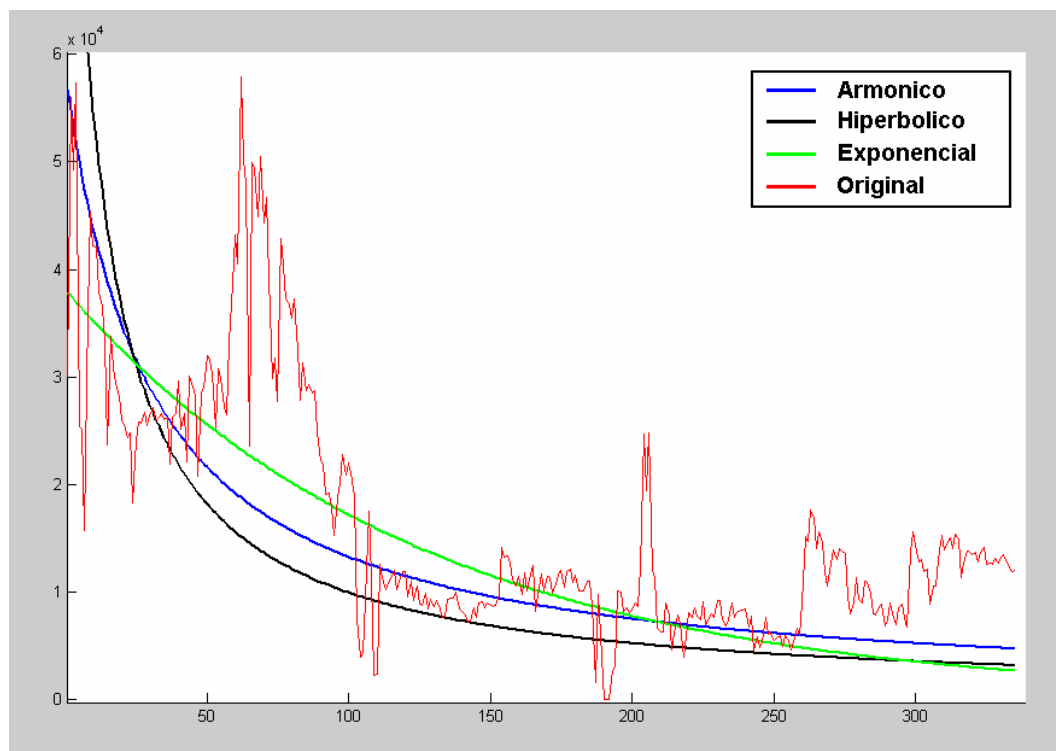


Figura 26. Pozo 1: Comparación entre DECLINET y la producción original

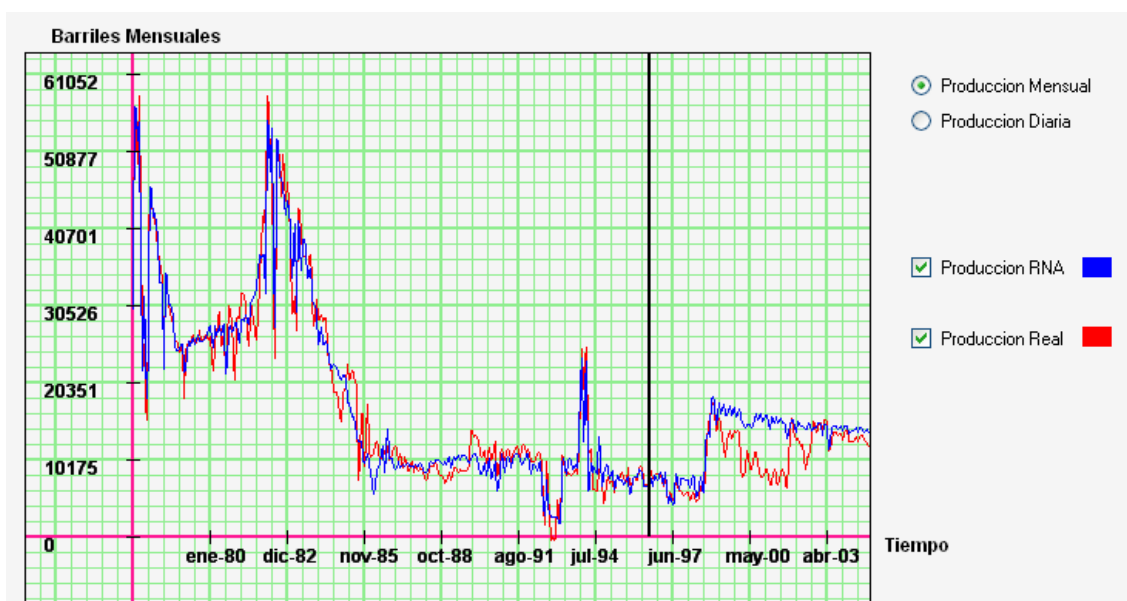


Figura 27. Pozo 1: Comparación entre las técnicas clásicas y la producción acumulada original

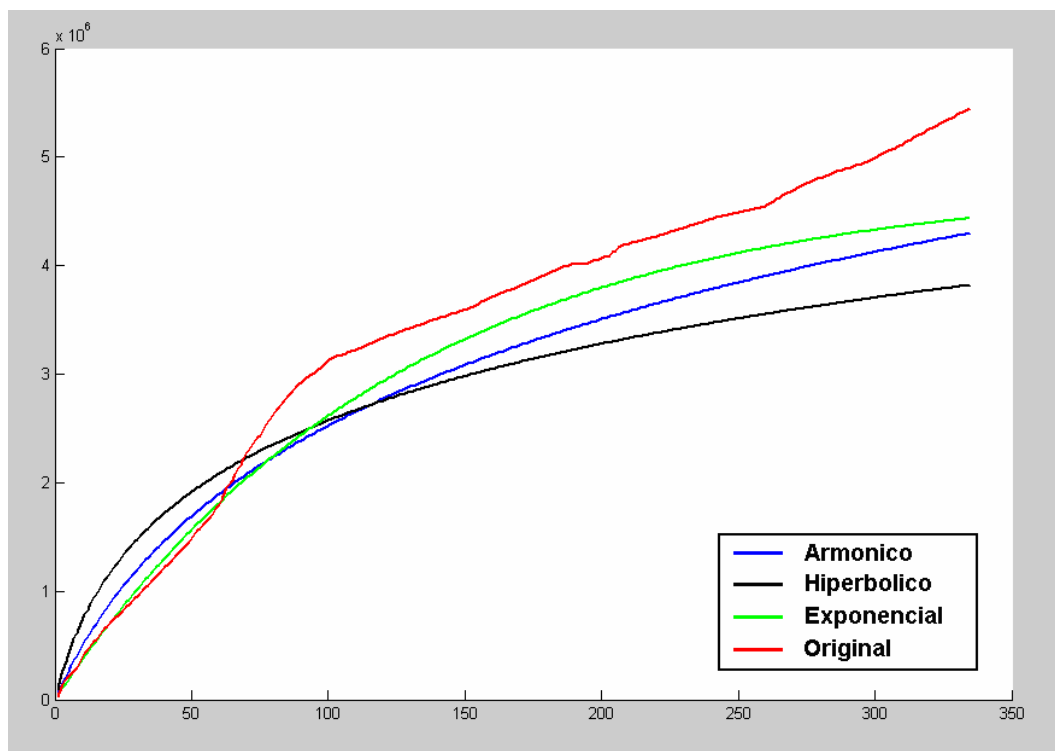
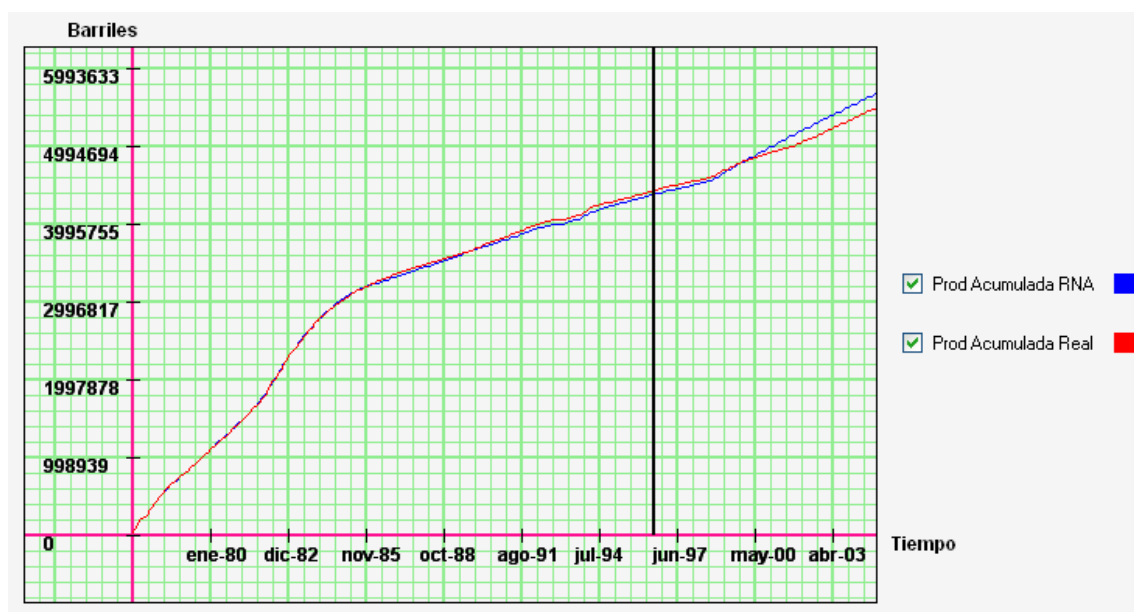


Figura 28. Pozo 1: Comparación entre DECLINET y la producción acumulada original



5.2 POZO 2

Inicio de la producción: Junio de 1977.

Final de producción: Diciembre de 2004.

Total de datos: 331

Tabla 35. Error absoluto promedio y tasa de ajuste de cada método para el pozo 2

	Armónico	Exponencial	Hiperbólico	RNA
Error Absoluto Promedio	1088.03	1629.69	1748.37	1211.50
Tasa de ajuste (%)	31.15	24.12	73.85	26.32

Figura 29. Pozo 2: Comparación entre las técnicas clásicas y la producción original

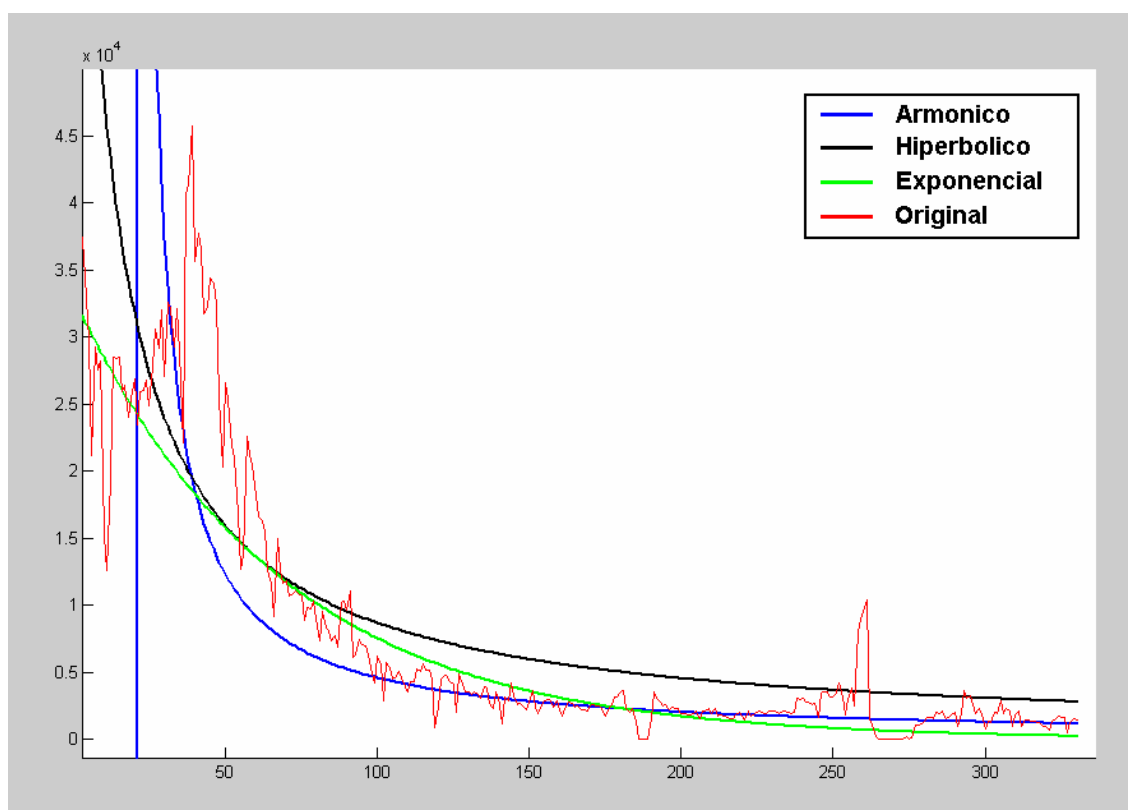


Figura 30. Pozo 2: Comparación entre DECLINET y la producción original

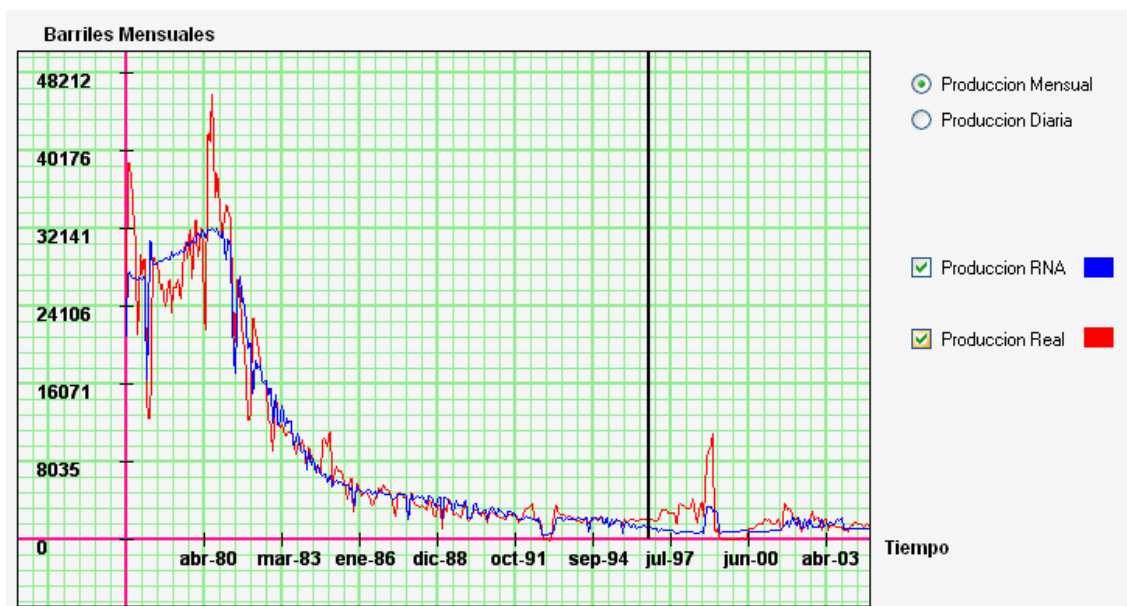


Figura 31. Pozo 2: Comparación entre las técnicas clásicas y la producción acumulada original

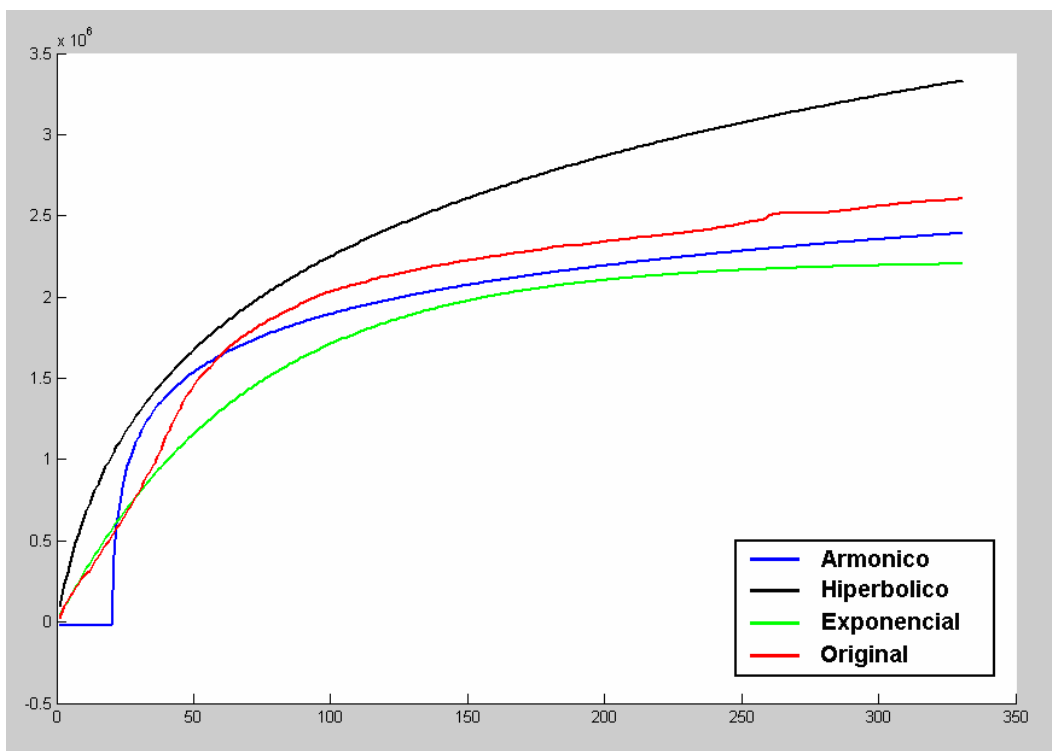
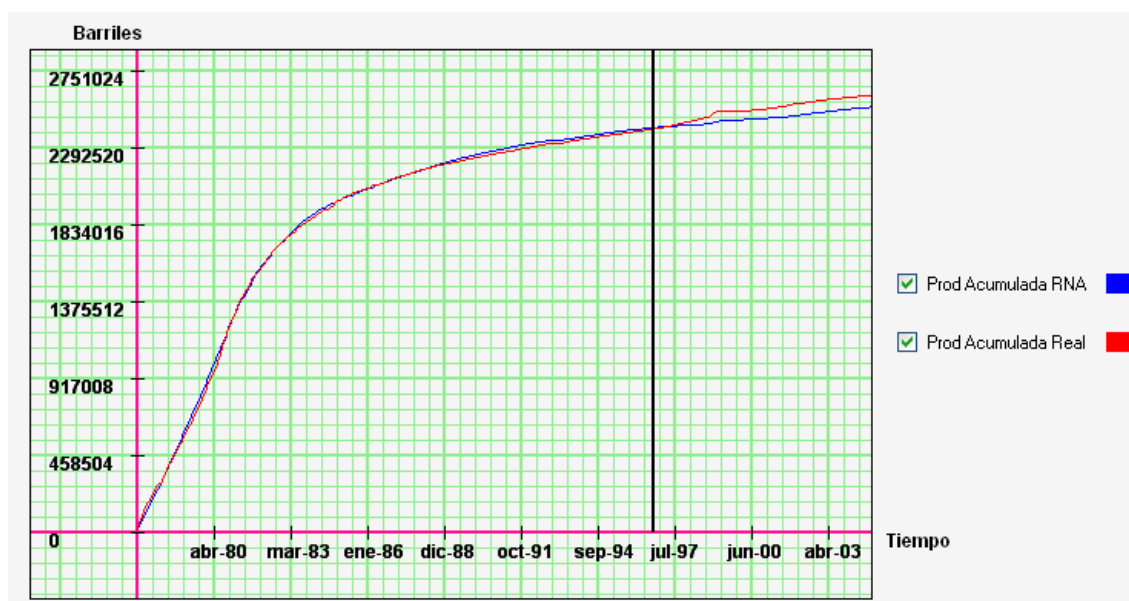


Figura 32. Pozo 2: Comparación entre DECLINET y la producción acumulada original



5.3 POZO 3

Inicio de la producción: Julio de 2000.

Final de producción: Diciembre de 2004.

Total de datos: 53

Tabla 36. Error absoluto promedio y tasa de ajuste de cada método para el pozo 3

	Armónico	Exponencial	Hiperbolico	RNA
Error Absoluto Promedio	10722.53	4225.34	4752.08	5084.10
Tasa de ajuste (%)	54.03	80.92	78.64	76.92

Figura 33. Pozo 3: Comparación entre las técnicas clásicas y la producción original

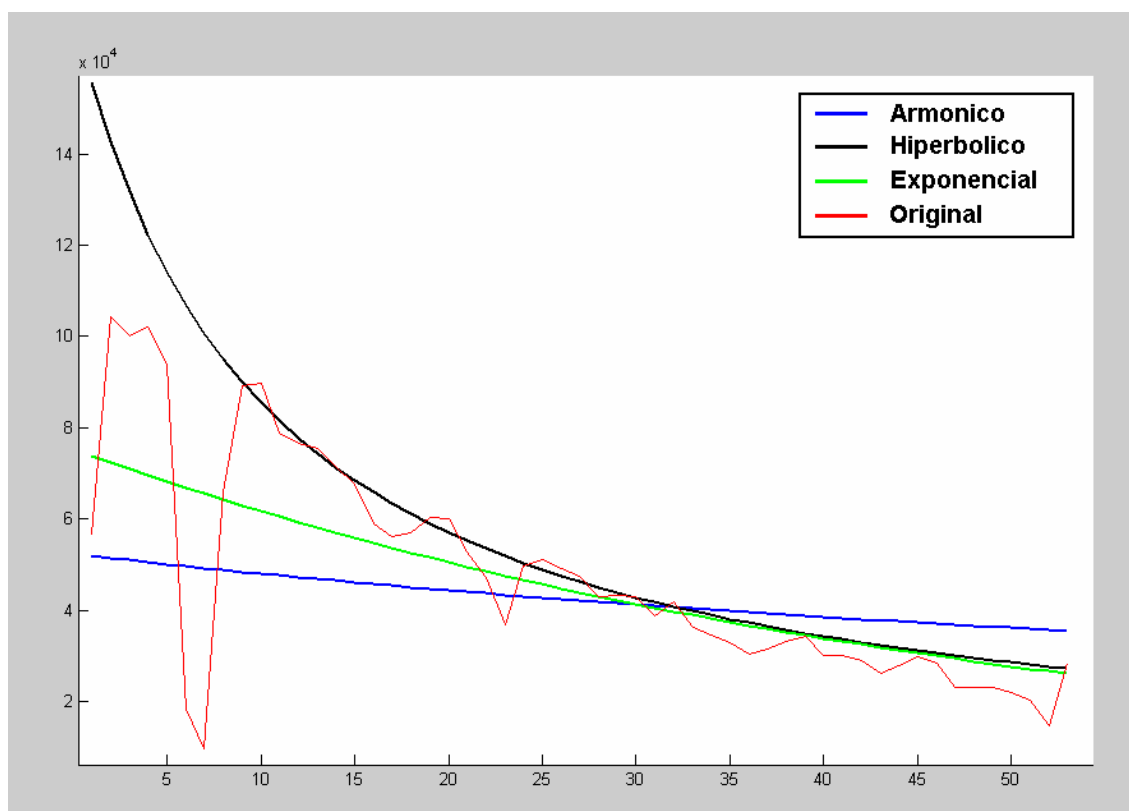


Figura 34. Pozo 3: Comparación entre DECLINET y la producción original

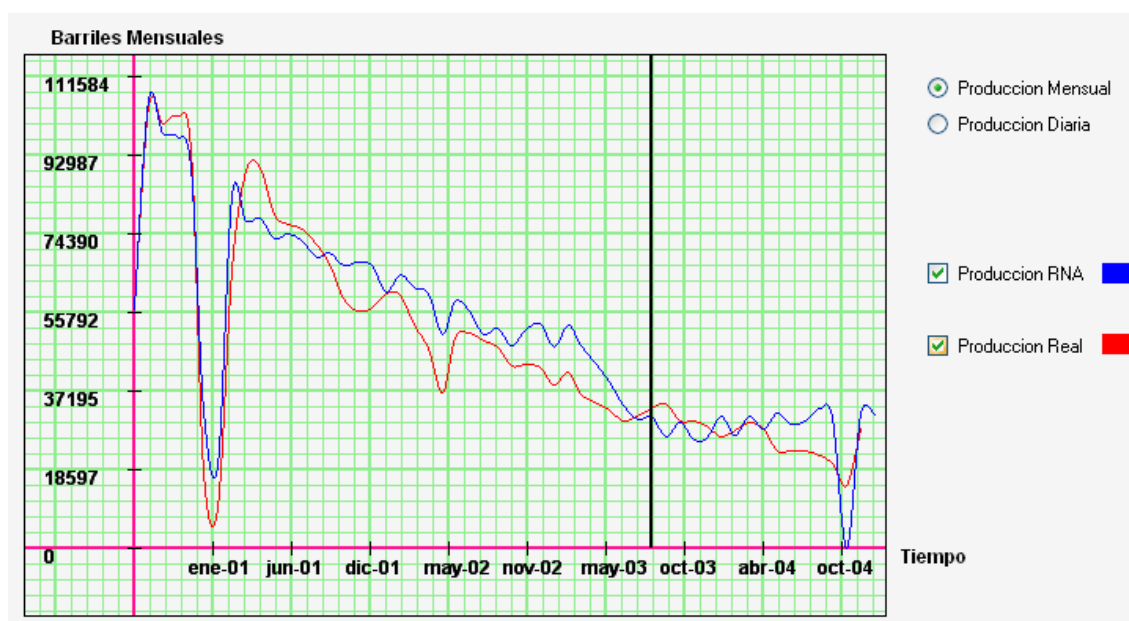


Figura 35. Pozo 3: Comparación entre las técnicas clásicas y la producción acumulada original

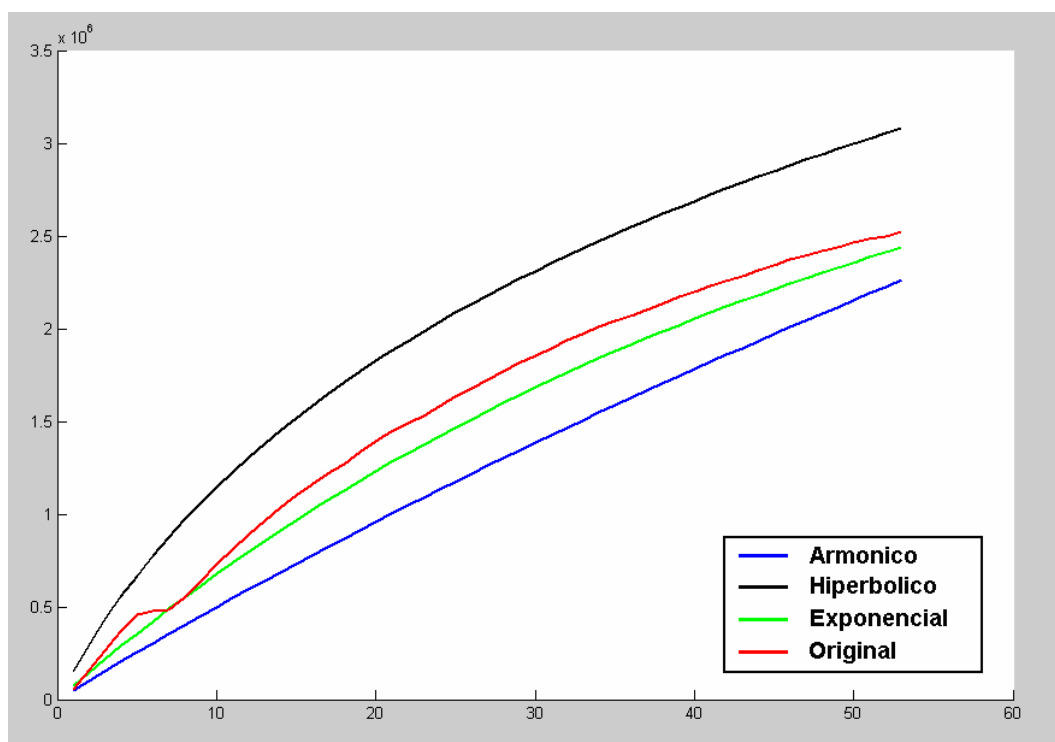
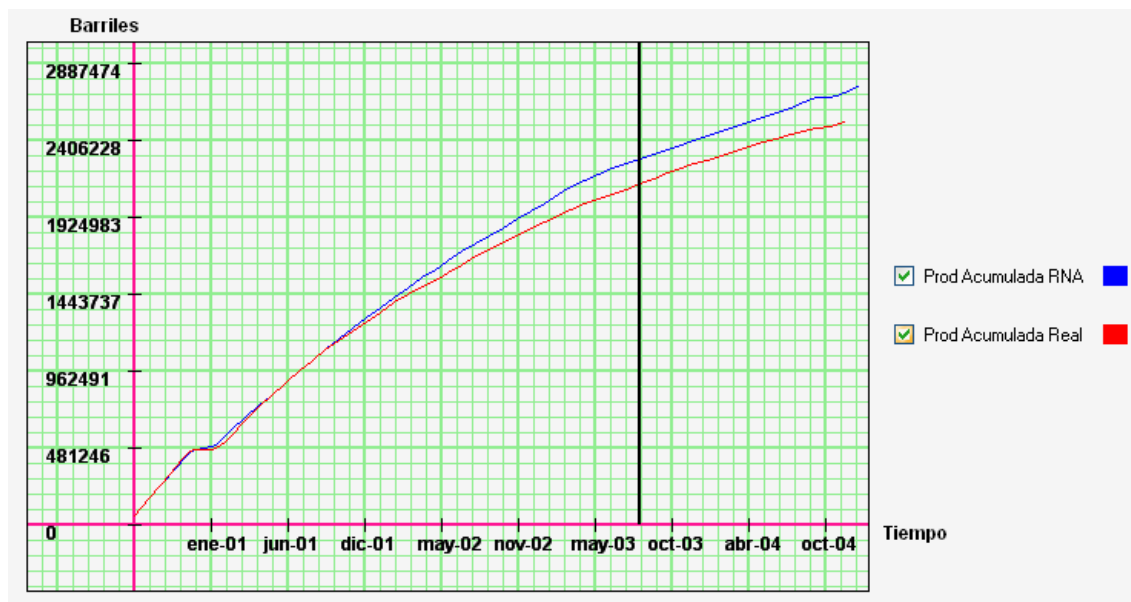


Figura 36. Pozo 3: Comparación entre DECLINET y la producción acumulada original



5.4 POZO 4

Inicio de la producción: Noviembre de 1979.

Final de producción: Diciembre de 2004.

Total de datos: 301

Tabla 37. Error absoluto promedio y tasa de ajuste de cada método para el pozo 4

	Armónico	Exponencial	Hiperbólico	RNA
Error Absoluto Promedio	1245.68	1566.77	1089.52	1343.37
Tasa de ajuste (%)	28.02	22.04	26.33	41.81

Figura 37. Pozo 4: Comparación entre las técnicas clásicas y la producción original

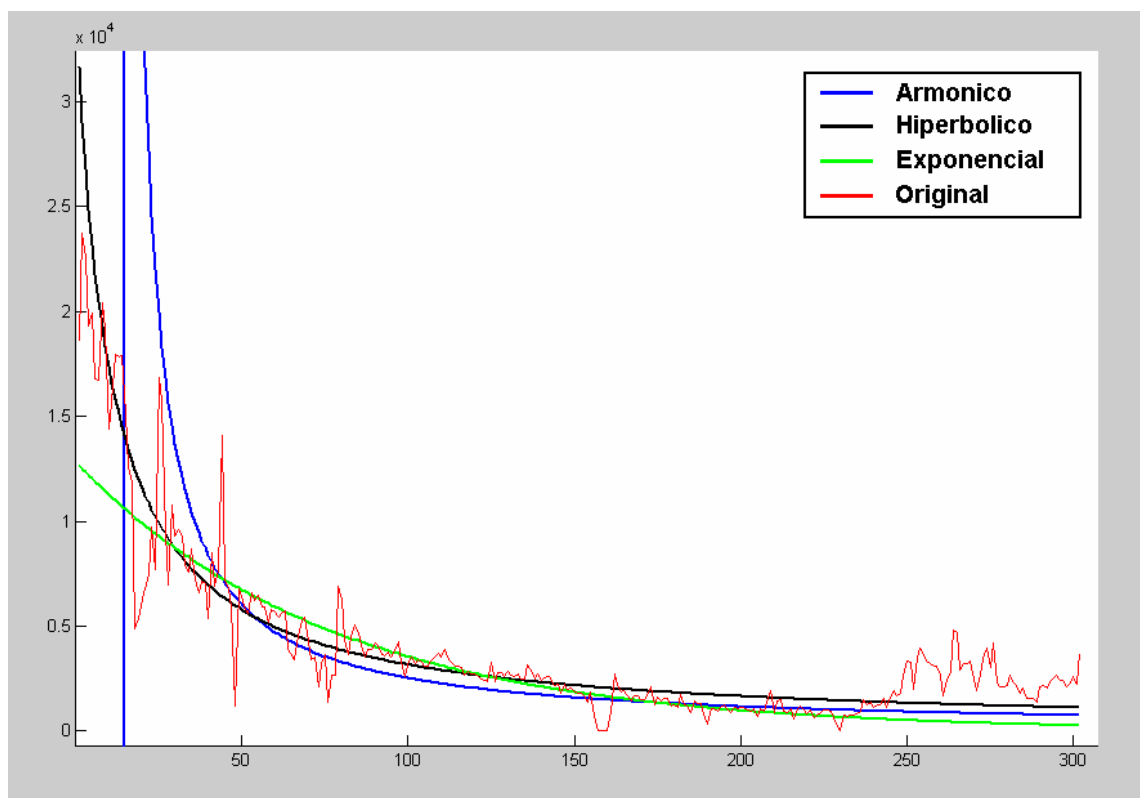


Figura 38. Pozo 4: Comparación entre DECLINET y la producción original

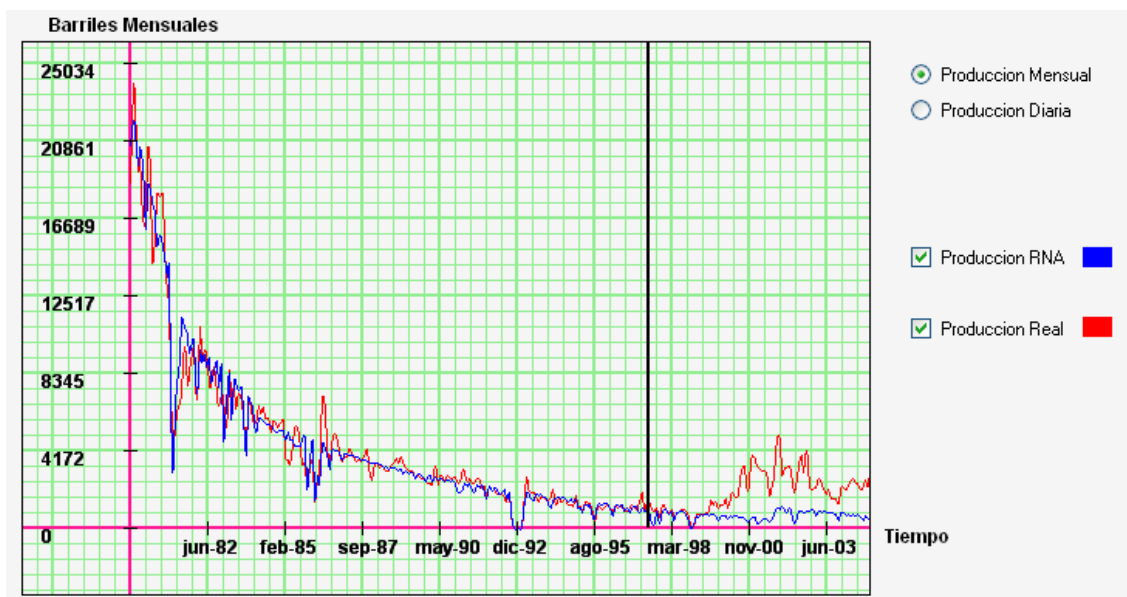


Figura 39. Pozo 4: Comparación entre las técnicas clásicas y la producción acumulada original

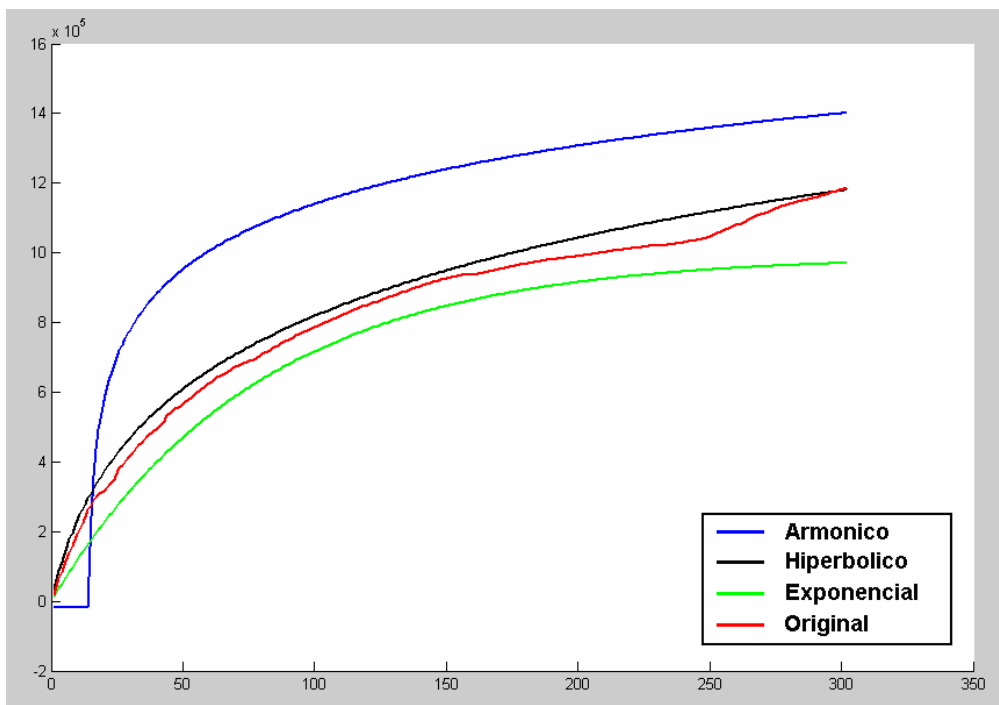
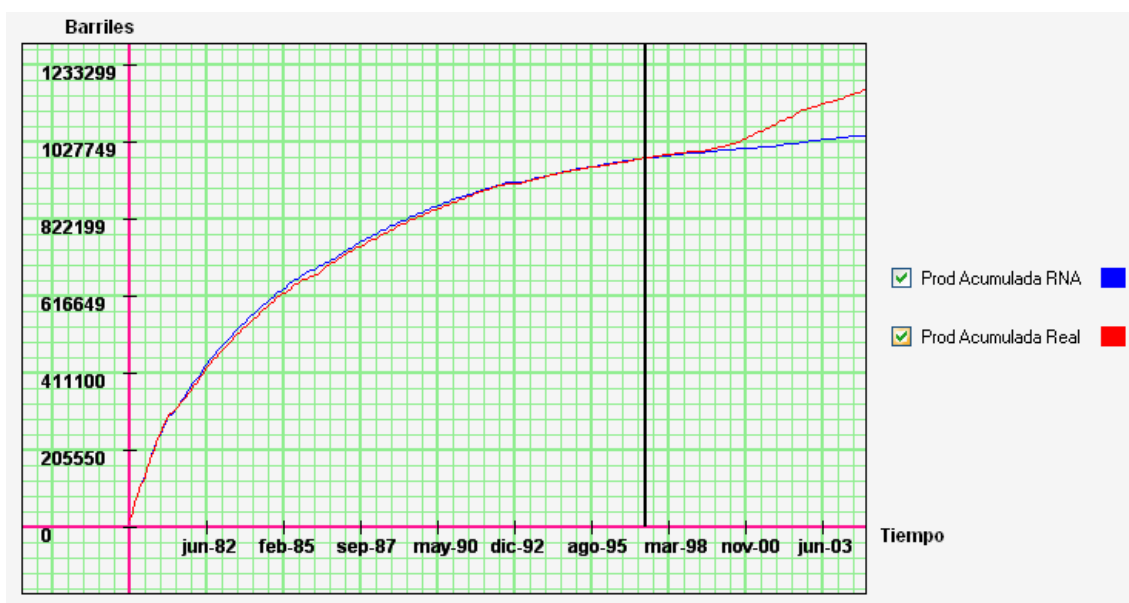


Figura 40. Pozo 4: Comparación entre DECLINET y la producción acumulada original



5.5 POZO 5

Inicio de la producción: Junio de 1980.

Final de producción: Diciembre de 2004.

Total de datos: 291

Tabla 38. Error absoluto promedio y tasa de ajuste de cada método para el pozo 5

	Armónico	Exponencial	Hiperbólico	RNA
Error Absoluto Promedio	9078.14	9922.82	8801.86	4444.74
Tasa de ajuste (%)	20.54	11.34	23.80	60.28

Figura 41. Pozo 5: Comparación entre las técnicas clásicas y la producción original

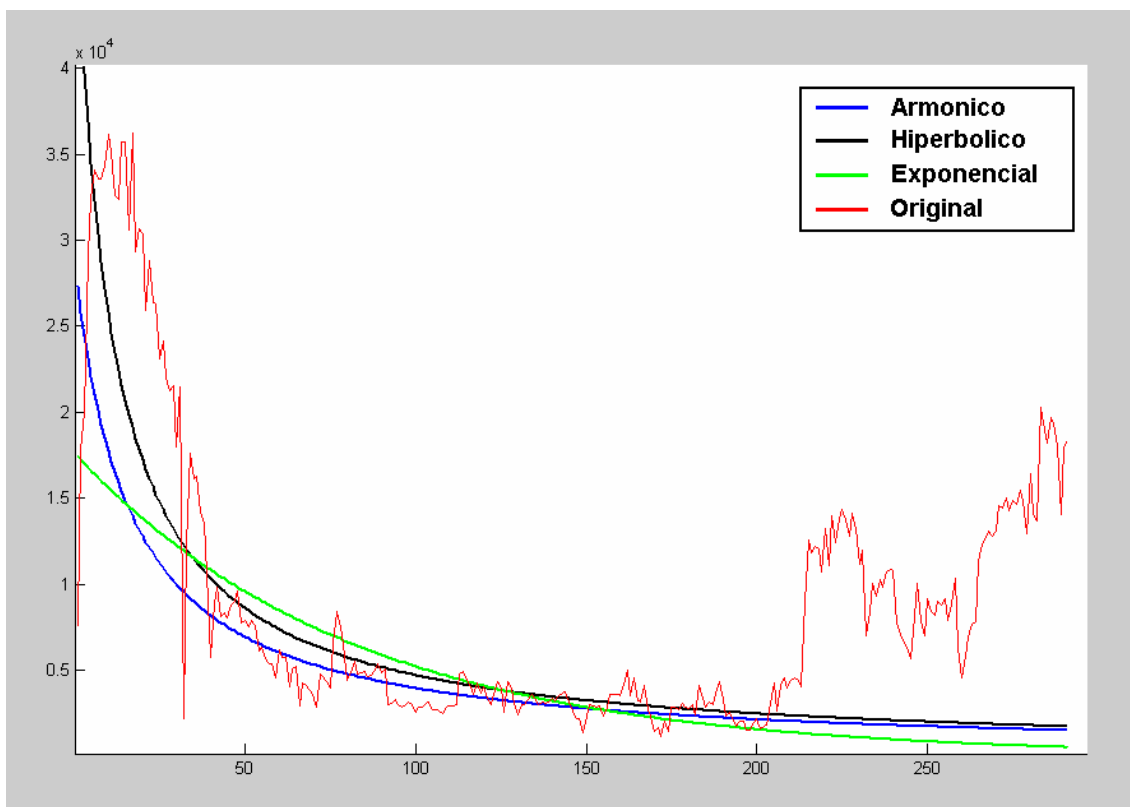


Figura 42 . Pozo 5: Comparación entre DECLINET y la producción original

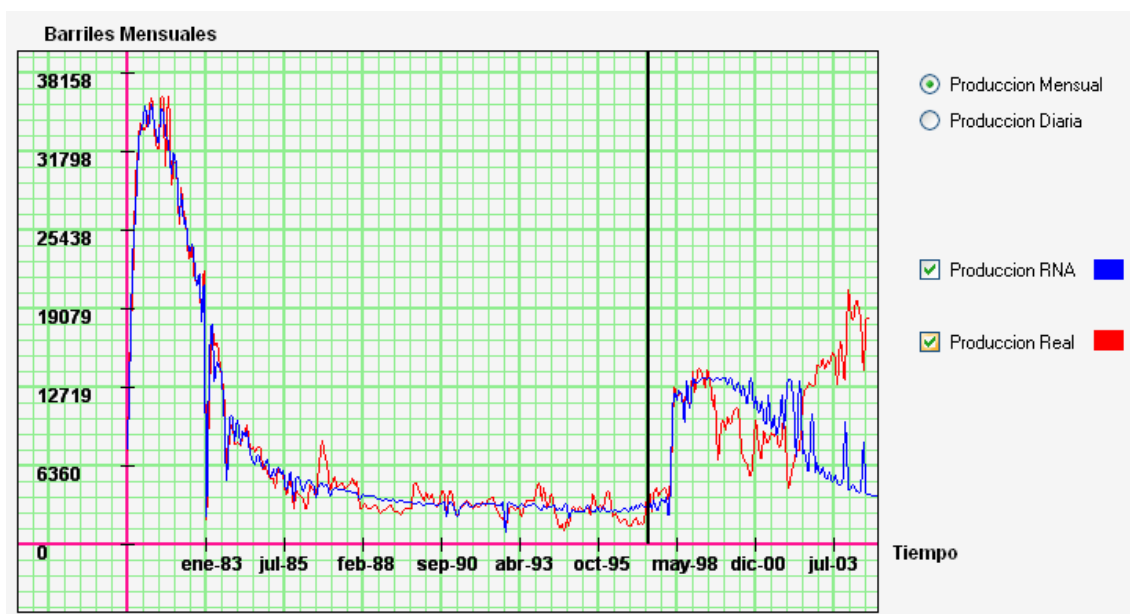


Figura 43. Pozo 5: Comparación entre las técnicas clásicas y la producción acumulada original

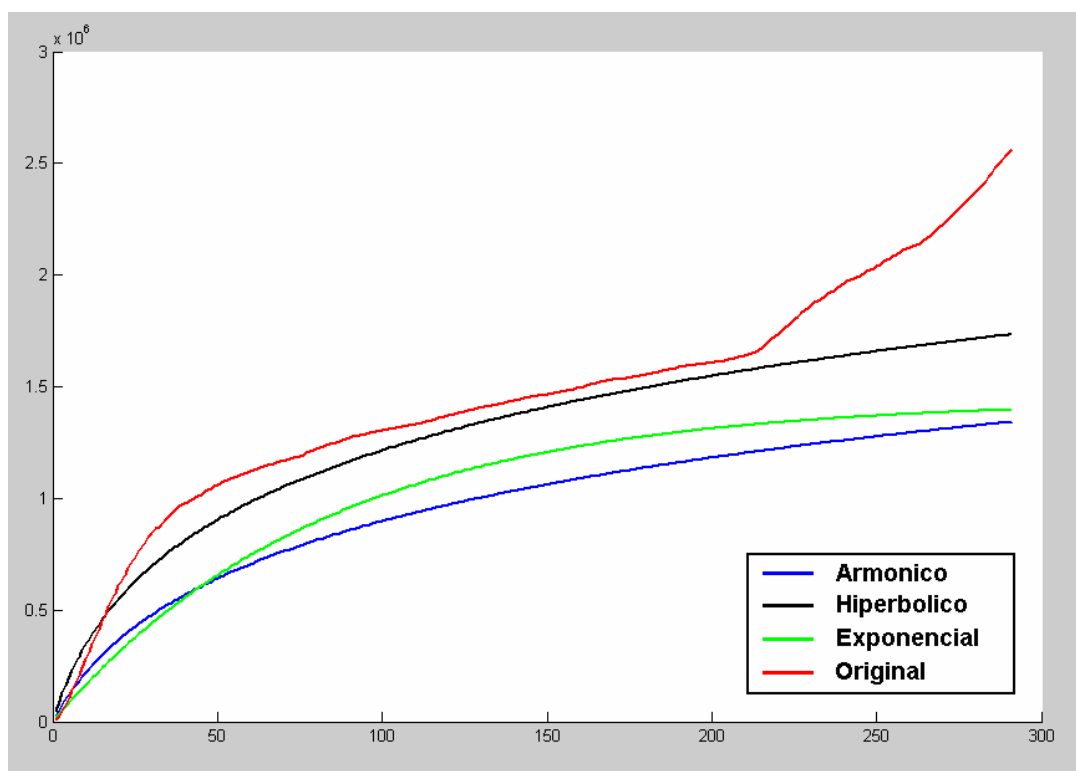
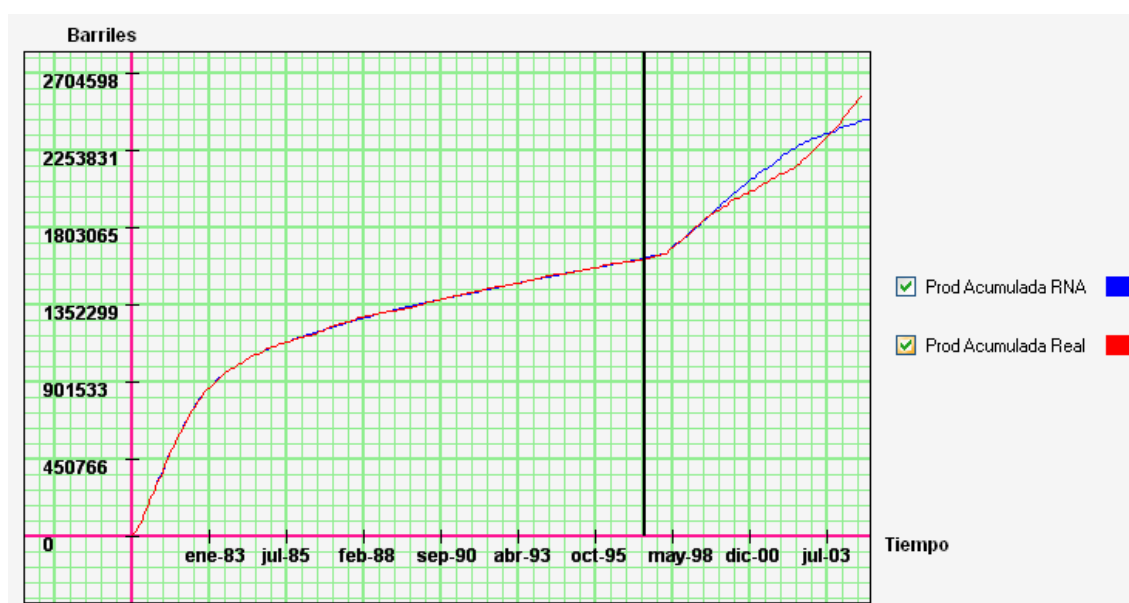


Figura 44. Pozo 5: Comparación entre DECLINET y la producción acumulada original



5.6 POZO 6

Inicio de la producción: Noviembre de 2000.

Final de producción: Diciembre de 2004.

Total de datos: 50

Tabla 39. Error absoluto promedio y tasa de ajuste de cada método para el pozo 6

	Armónico	Exponencial	Hiperbolico	RNA
Error Absoluto Promedio	32930.97	19096.39	32034.28	7988.47
Tasa de ajuste (%)	4.48	40.99	7.53	75.49

Figura 45. Pozo 6: Comparación entre las técnicas clásicas y la producción original

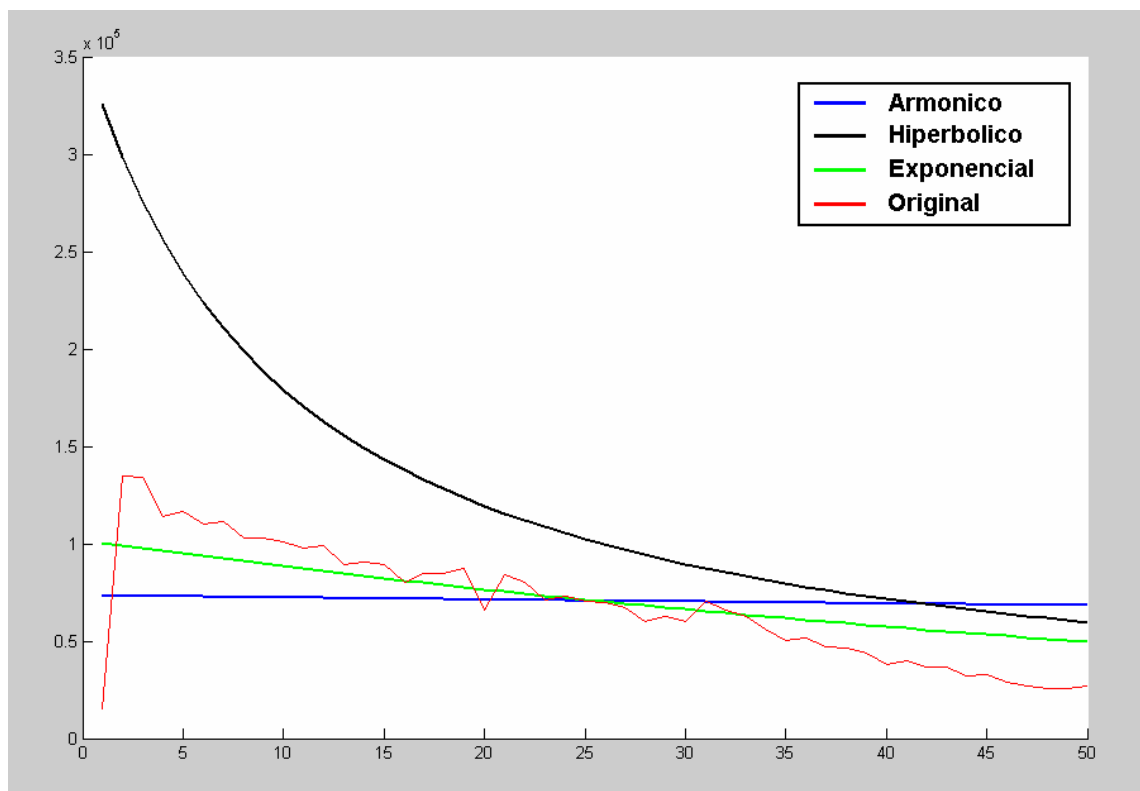


Figura 46. Pozo 6: Comparación entre DECLINET y la producción original



Figura 47. Pozo 6: Comparación entre las técnicas clásicas y la producción acumulada original

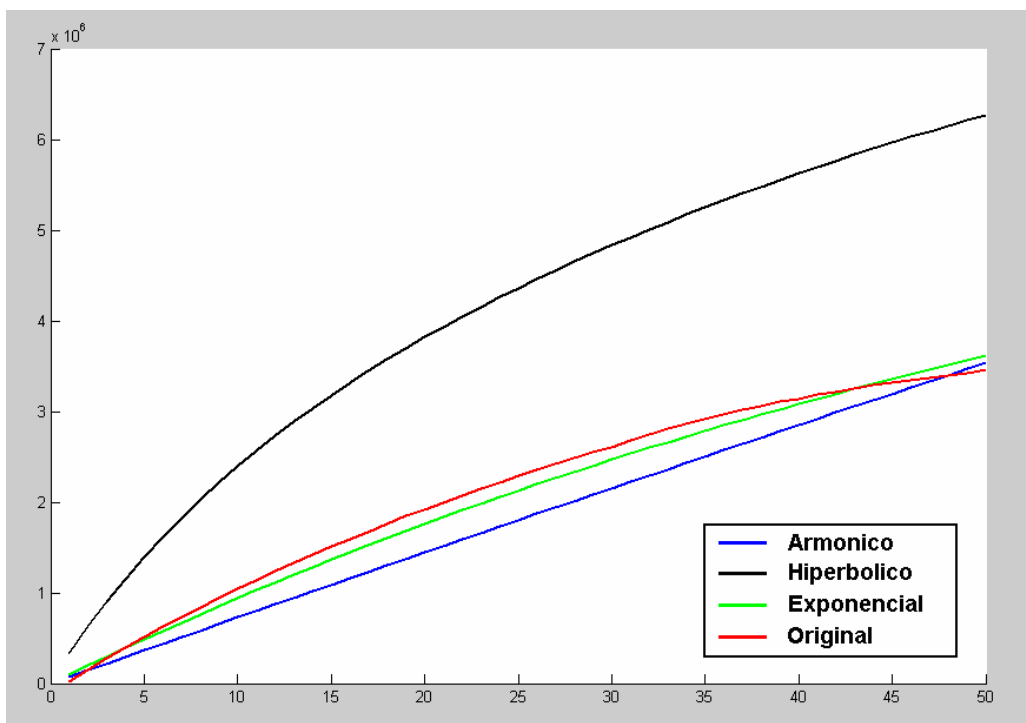
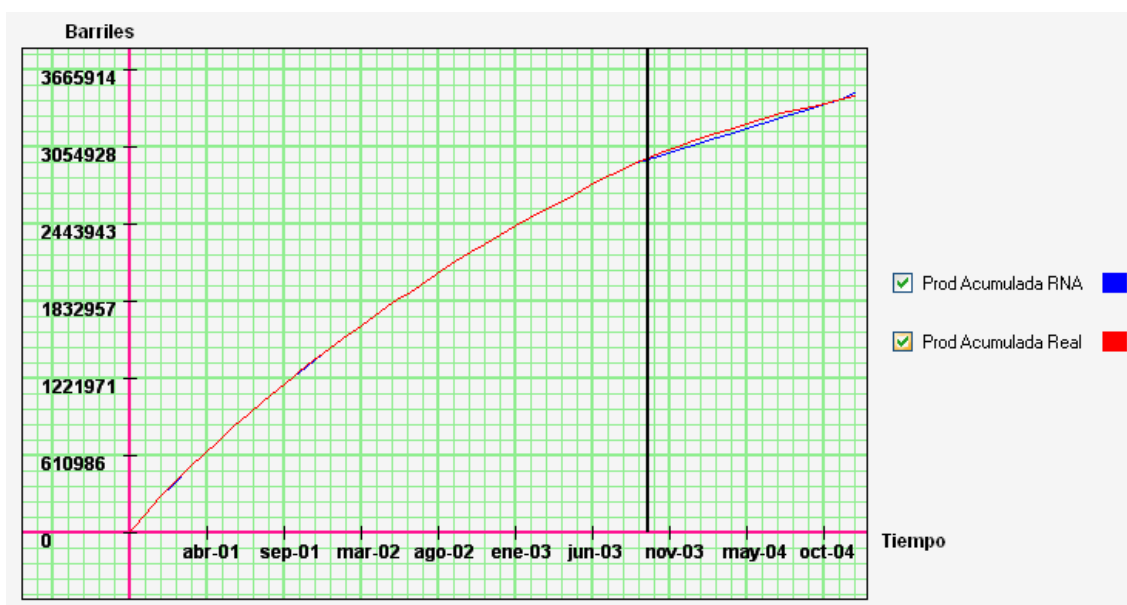


Figura 48. Pozo 6: Comparación entre DECLINET y la producción acumulada original



5.7 POZO 7

Inicio de la producción: Julio de 1999.

Final de producción: Diciembre de 2004.

Total de datos: 66

Tabla 40. Error absoluto promedio y tasa de ajuste de cada método para el pozo 7

	Armónico	Exponencial	Hiperbólico	RNA
Error Absoluto Promedio	1393.62	3991.36	4232.59	3112.58
Tasa de ajuste (%)	83.93	63.73	61.71	66.46

Figura 49. Pozo 7: Comparación entre las técnicas clásicas y la producción original

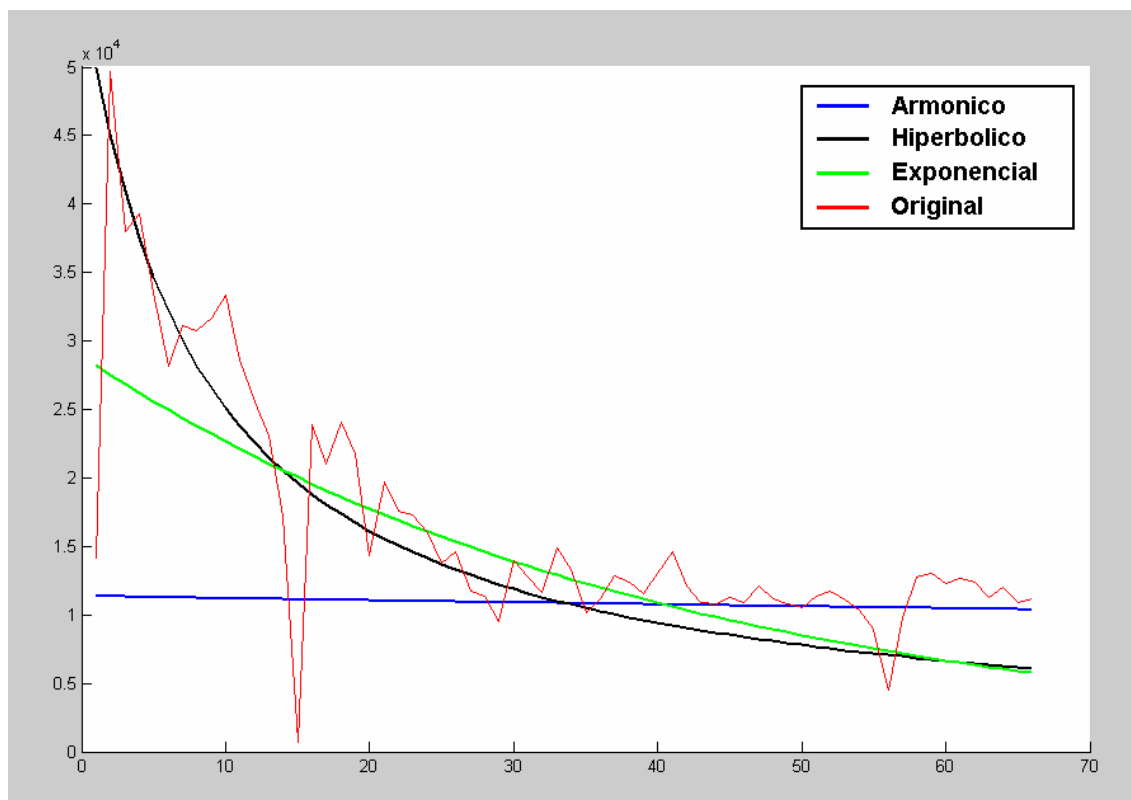


Figura 50. Pozo 7: Comparación entre DECLINET y la producción original

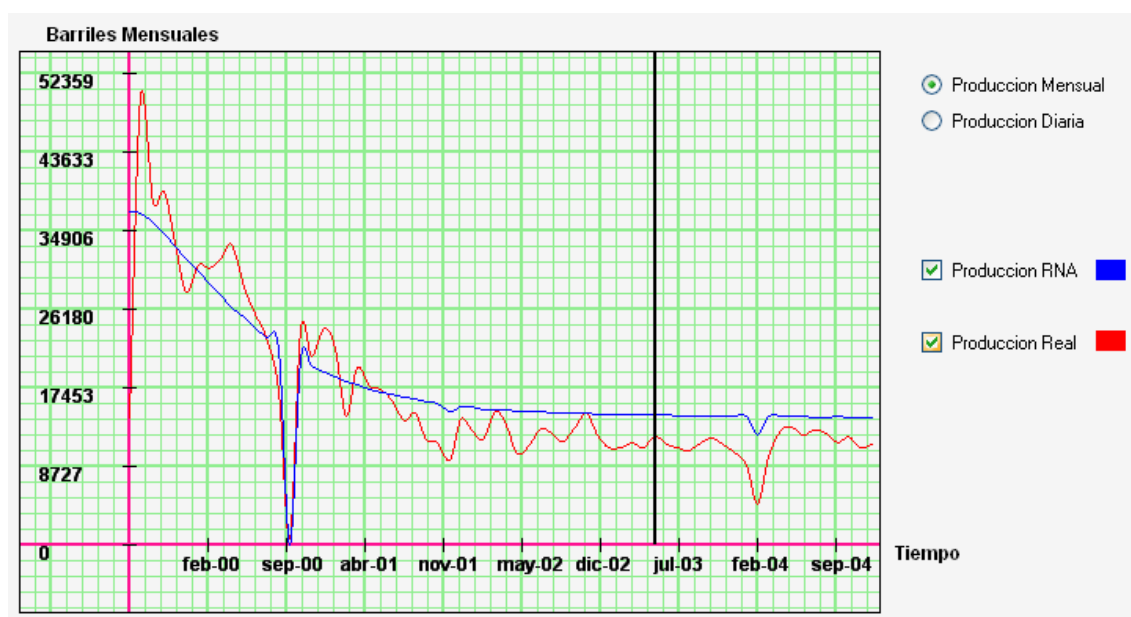


Figura 51. Pozo 7: Comparación entre las técnicas clásicas y la producción acumulada original

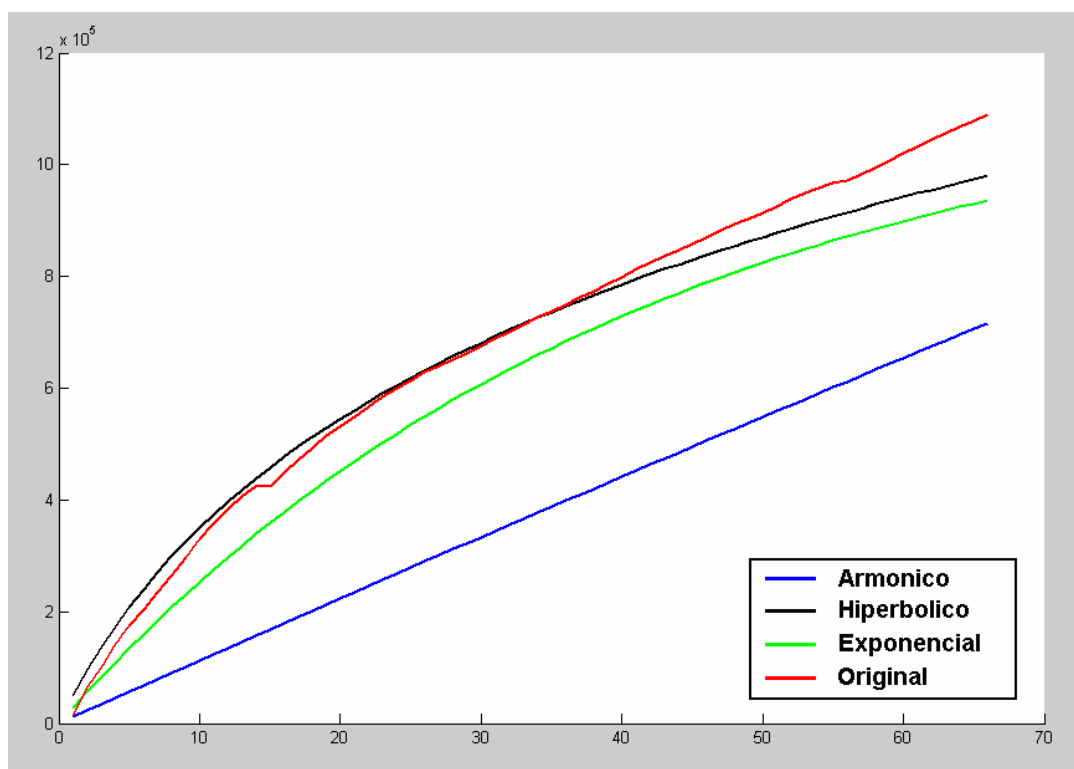
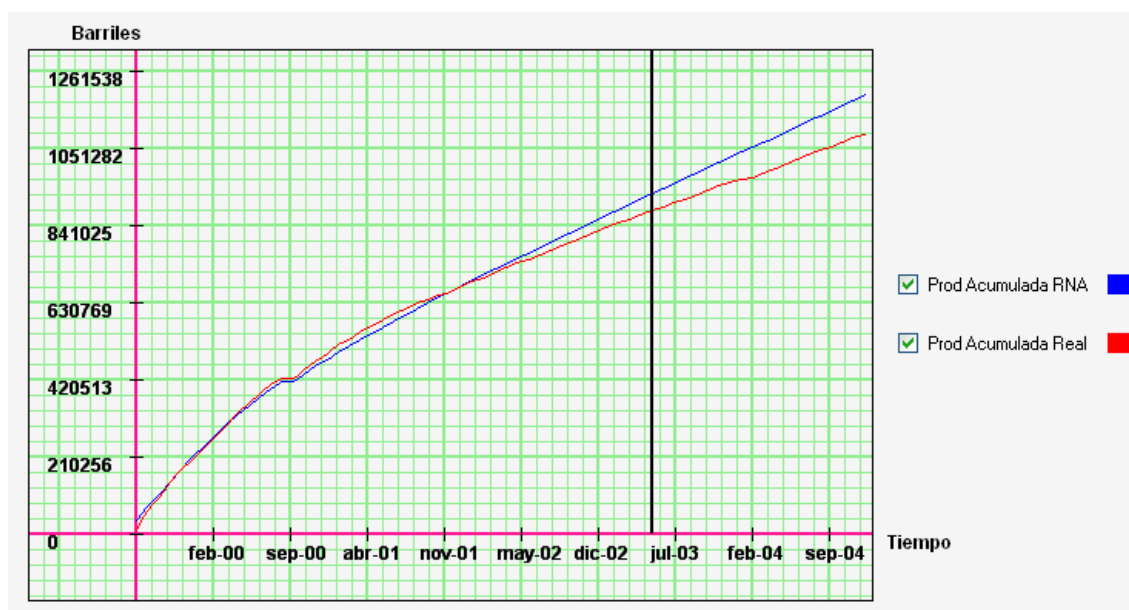


Figura 52. Pozo 7: Comparación entre DECLINET y la producción acumulada original



5.9 OBSERVACIONES

Del proceso de comparación de las técnicas tradicionales, DECLINET y los datos reales de producción, se encontraron las siguientes consideraciones:

A pesar de que en ciertos pozos las técnicas clásicas arrojan menores errores que el modelo neuronal planteado, para los ingenieros de petróleo, en la mayoría de los casos, es más valiosa la tendencia mostrada por la simulación. Es por eso que para la toma de decisiones, en cuanto a la realización de operaciones que conlleven al mejoramiento de la producción del pozo, la aplicación se convierte en una herramienta de gran utilidad, por su carácter predictivo.

La red basa su aprendizaje en los datos históricos de producción, de ahí que si existen pocos datos para el entrenamiento, la red neuronal aprenderá menos patrones, y por consiguiente la calidad en la generalización se verá comprometida.

Adicional a las comparaciones anteriores, se toma un pozo como ejemplo, al que se le realiza un análisis exhaustivo de las bondades ofrecidas por la aplicación desarrollada, entre ellas, la tasa de declinación y el cálculo del límite económico.

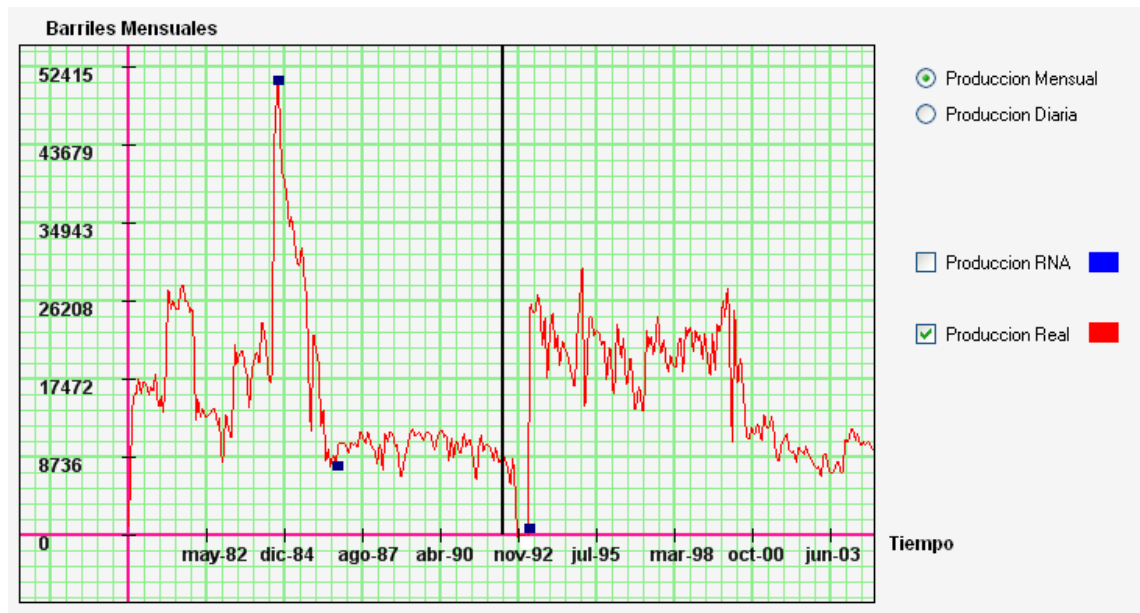
5.10 CASO DE ESTUDIO, POZO 8

5.10.1 Curvas De Declinación

El análisis de las curvas de declinación, realizado mediante los métodos clásicos, se basa en la gráfica de la producción Vs tiempo. La aplicación, en cambio, evita el uso de estas aproximaciones gráficas, porque la información necesaria para dicho análisis la obtiene directamente de los archivos de producción. A pesar de esto, el primer paso incluye la construcción de tal

gráfica, con el objetivo de mostrarle al usuario el comportamiento del pozo y verificar la validez de los resultados que se obtendrán de la simulación.

Figura 53. Pozo ejemplo: Producción original



Tal como lo muestra la fig. 52, el comportamiento de este pozo no cuenta con las características necesarias para aplicar correctamente las técnicas clásicas. Estas características implican que el pozo se encuentre estable, es decir que no presente variaciones bruscas en la producción, y que además cuente con una tasa de declinación aproximadamente constante, tal que la gráfica de la producción contra el tiempo pueda aproximarse a una línea recta en un papel semilog.

La mayoría de los pozos que se encuentran en los yacimientos, no poseen las características de estabilidad y tasa constante arriba descritas, y este es precisamente uno de los motivos que impulsó a la realización de este trabajo, con el fin de poder analizar no sólo aquellos pozos de condiciones ideales sino también otros que presentan altas variaciones en cuanto a su producción y tasa de declinación. En este caso práctico se pretende mostrar la validez de la metodología propuesta, eligiendo un pozo problema al cual se le han aplicado

diferentes procedimientos técnicos, que se ven reflejados en un cambio notable en el curso de la producción; estos procedimientos están marcados como cuadrados azules dentro de la fig. 52.

Para el análisis de la curva de declinación, la herramienta se entrena con el 50% de los datos de producción, que equivalen a 152 meses y van desde septiembre de 1979 hasta abril de 1992. La brecha que delimita la etapa de la simulación y la de validación se muestra en la fig. 52 como una línea negra.

Luego del entrenamiento, la producción es simulada hasta el final de archivo, con el fin de mostrar su validez respecto a la producción real. No obstante, DECLINET es capaz de simular la producción hasta cualquier fecha posterior, pero para este caso no habría forma de medir la validez.

DECLINET es capaz de generar la curva de declinación en dos escalas: la escala de producción mensual, donde se aprecia mejor el comportamiento global de la simulación a nivel gráfico; y la escala de producción diaria, donde pueden presentarse variaciones notables en los valores simulados, respecto a la real, de esta forma puede que la información mostrada difiera de la tendencia generada en la escala mensual, esto se debe a que al hacerse la escala mas pequeña, se gana precisión en la información; a pesar de esto, los datos generados son consistentes con la declinación real.

Figura 54. Pozo ejemplo: Comparación entre DECLINET y la producción original en escala mensual

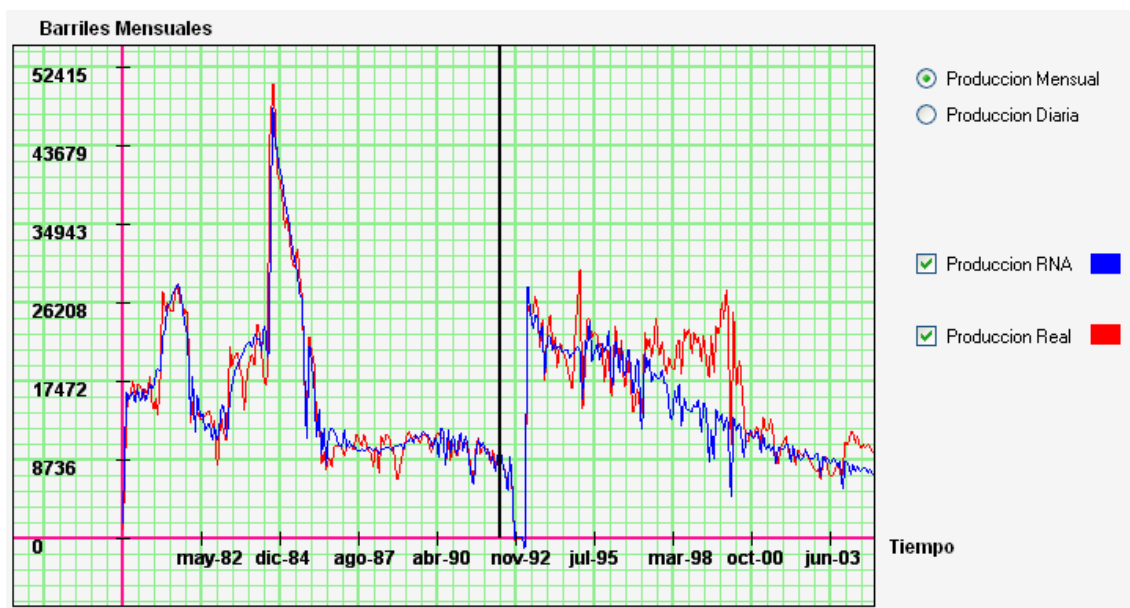
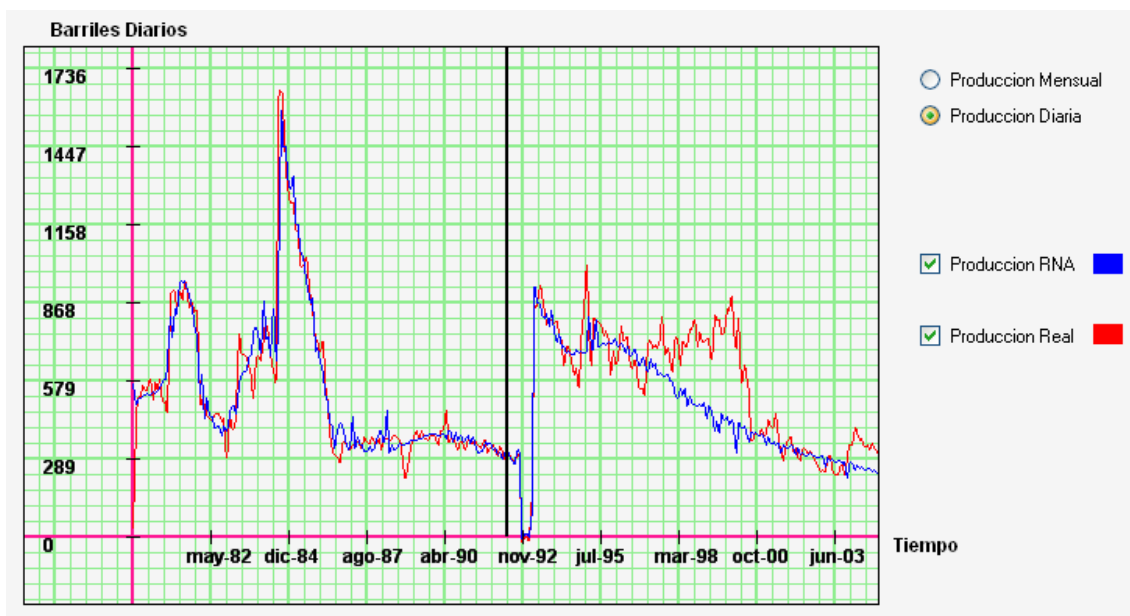
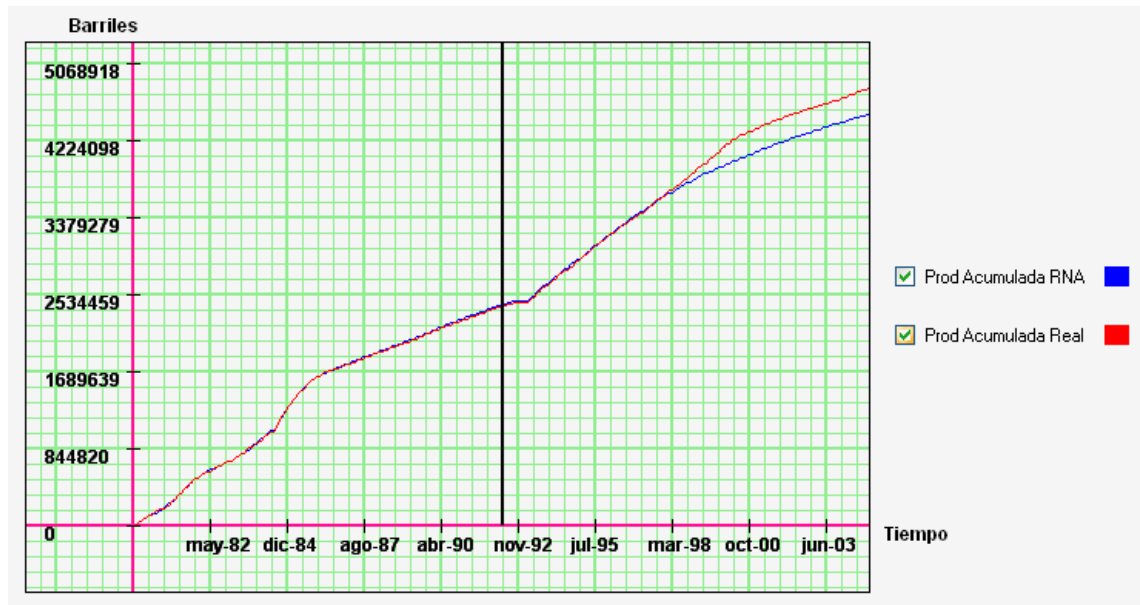


Figura 55. Pozo ejemplo: Comparación entre DECLINET y la producción original en escala diaria



5.10.2 Producción Acumulada

Figura 56. Pozo ejemplo: Comparación entre DECLINET y la producción acumulada original



Como se puede apreciar en la fig. 55, la producción acumulada de la aplicación se asemeja en comportamiento y valores a la real. Este ítem puede ser usado como indicador adicional para medir la validez del modelo, y es ampliamente utilizado por los ingenieros de producción en técnicas complementarias.

5.10.3 Tasa de declinación

En las técnicas clásicas, la tasa de declinación, que se asume constante, se usa para extrapolar la curva de producción y de esta forma predecir su valor en cualquier instante de tiempo. Asumir esta tasa como constante en todos los casos, arroja resultados poco satisfactorios frente a pozos como el escogido para este ejemplo. En nuestra aplicación, la tasa de declinación, a diferencia de ser el factor que determina la producción, aparece como resultado del proceso de simulación, porque primero se obtiene la producción futura, y a

partir de esta información, se procede a calcular la tasa de declinación. La exactitud de esta medición depende de la periodicidad que se escoja.

Como consecuencia de que este modelo tiene en cuenta los cambios en el tipo de bombeo, que casi siempre alteran el normal curso de la producción, es posible observar que en la vida productiva del pozo existan tramos que no tienen tendencia declinante, por el contrario, se presenta un aumento de la producción. Este fenómeno es común en muchos de los pozos reales y no está contemplado dentro de las ecuaciones de Arps, para el análisis de curvas de declinación. DECLINET en cambio, contempla esta situación, luego, podemos encontrar tasas que en lugar de indicar una caída de la producción, representan una mejora, estas se muestran en el sistema con su valor respectivo, precedido del símbolo (^).

Para el caso de estudio, se escogieron 2 periodicidades correspondientes a 24 y 60 meses. Si se quiere hacer un análisis más detallado de la tasa de declinación, pueden escogerse periodos de estudio más cortos. Para fines prácticos en este ejemplo tomamos valores tales que arrojen una cantidad moderada de datos para poder visualizar un conjunto representativo, y tener una idea general del procedimiento.

Figura 57. Tasa de declinación generada por DECLINET, para una periodicidad de 24 meses

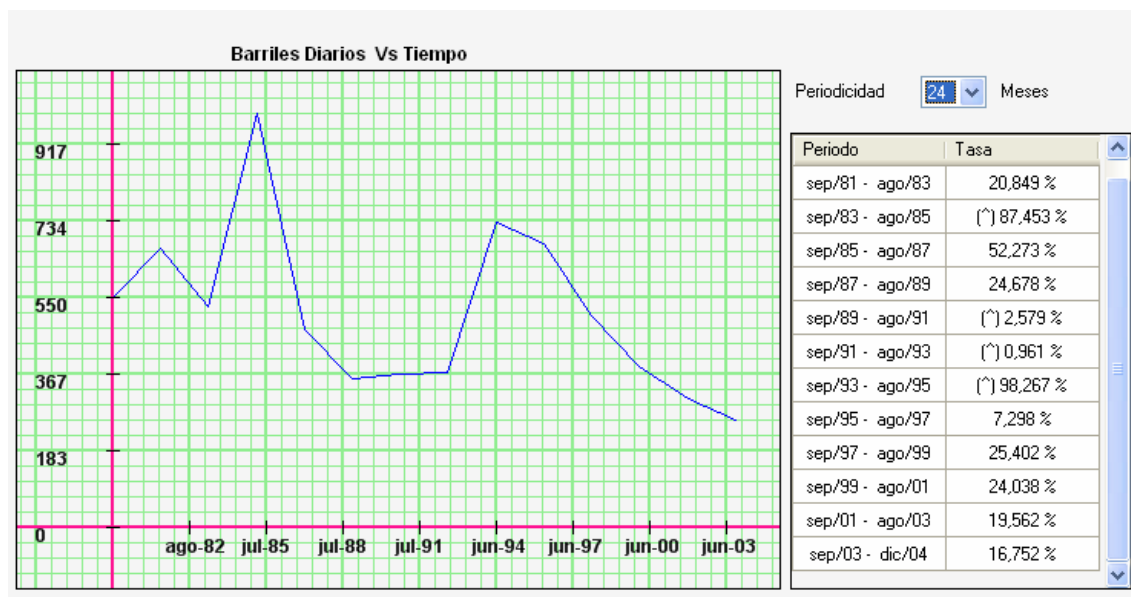
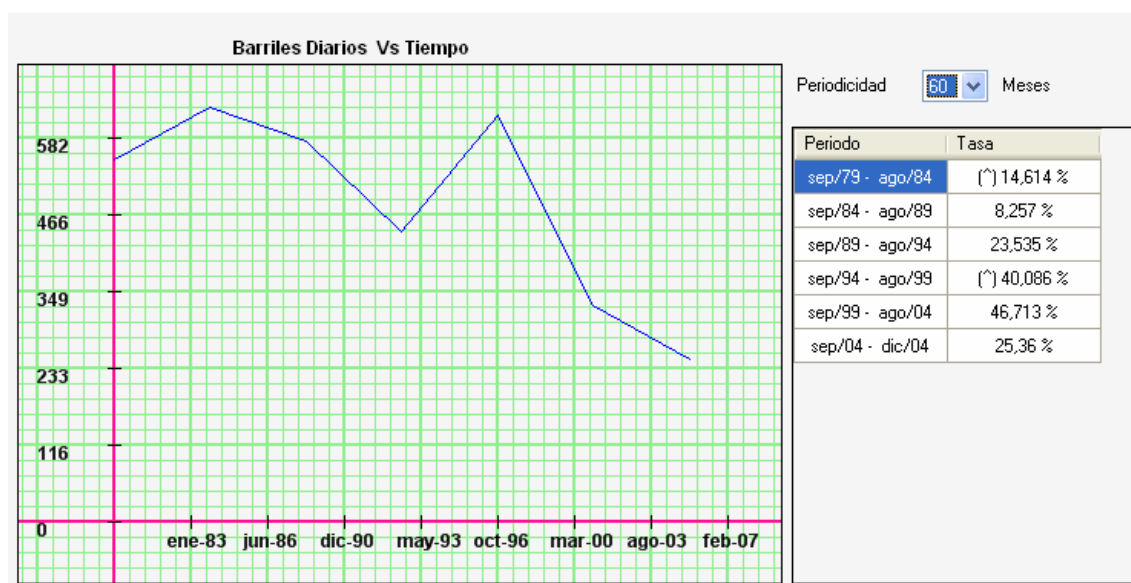


Figura 58. Tasa de declinación generada por DECLINET, para una periodicidad de 60 meses



Para calcular la tasa de declinación se promedian los valores de producción en cada intervalo establecidos por la periodicidad escogida, luego este valor promedio se toma como la marca de cada clase. Para la presentación de los resultados se siguen dos procedimientos: uno gráfico y uno aritmético. En el primero de ellos, el conjunto previamente hallado se une mediante líneas

rectas; de esta manera, se forma una gráfica aproximada del comportamiento de la variación en la producción, donde su mayor utilidad es brindar un panorama amplio al usuario del comportamiento del pozo. Para el segundo de los métodos, el aritmético, al conjunto de valores se le aplica la siguiente formula:

$$tasa_i(\%) = 100 \frac{prom_i - prom_{i+1}}{prom_i} \quad (5.2)$$

5.10.4 Límite Económico

El límite económico se define como el punto, en la vida productiva del pozo, hasta donde su rendimiento es viable en términos de rentabilidad. Esta rentabilidad pocas veces se calcula para un pozo individual, sino para todo el conjunto de pozos pertenecientes al mismo yacimiento. Los ingenieros de producción para determinar la viabilidad económica de cualquier pozo, traducen dicha rentabilidad en los barriles diarios mínimos (umbral mínimo) que debe producir el pozo. A partir de este parámetro y apoyado en los resultados de la simulación, se hacen los cálculos necesarios para determinar en que punto o fecha el pozo alcanza el límite económico.

El procedimiento para dicho cálculo, se basa en el siguiente algoritmo:

1. Se analiza cada mes de la línea de tiempo de forma individual, y en caso de que la producción sea menor que el umbral mínimo establecido, se consultan los 12 meses venideros de producción.
2. Si el promedio de la producción de estos meses es menor que el umbral mínimo, se dice que a partir del mes analizado se alcanza el límite económico.

3. Si el promedio es mayor que el umbral, se vuelve al paso 1 y se prueba con el mes siguiente.
4. Si se llega al final de los meses simulados y no se ha cumplido con la condición del punto 2, se dice que en los meses simulados, no se alcanza el límite económico.

La condición expresada en el punto 2 asegura que no se incurra en un falso límite económico debido a algún mes de baja producción, sino que sea necesario tener la certeza de que el pozo continúe con la tendencia de pérdida por lo menos un año mas.

A continuación se muestran dos escenarios, el primero de ellos, cuando se encuentra el límite y el segundo cuando no hay límite económico.

Figura 59. Pozo ejemplo: calculo del límite económico, caso 1

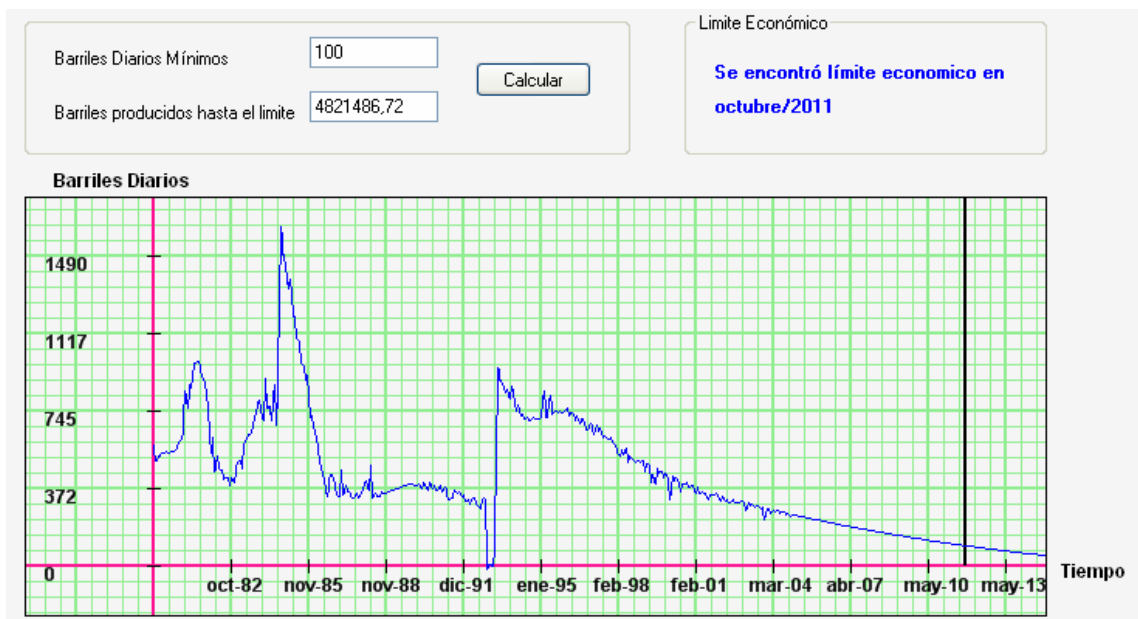
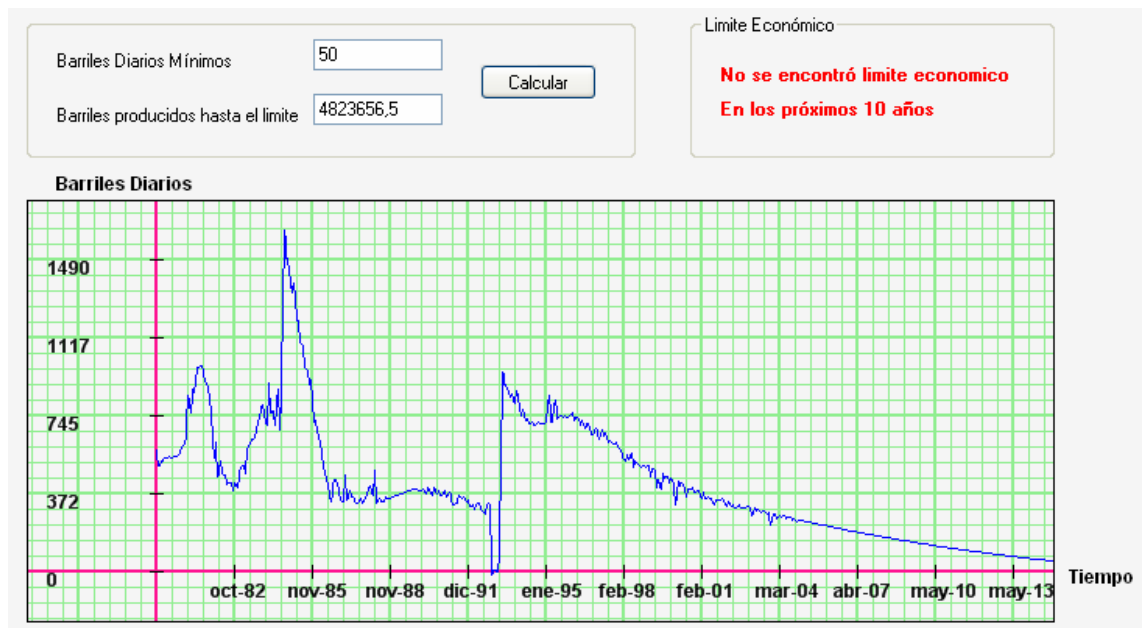


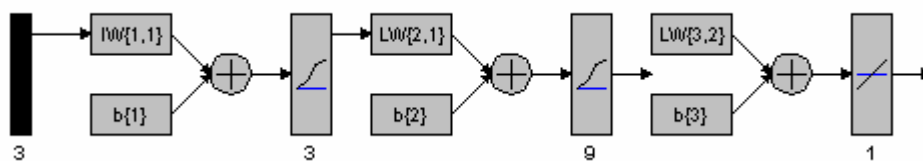
Figura 60 Pozo ejemplo: calculo del límite económico, caso 2



5.11 VERIFICACION DE LA HERRAMIENTA

Para comprobar la consistencia de la información arrojada por la herramienta DECLINET, durante los procesos de entrenamiento y simulación de la red, se realizó una prueba comparativa con matlab 6.5, en la que se usó el mismo conjunto de pesos iniciales para el entrenamiento. La asignación de los pesos y ganancias, y su distribución en la arquitectura de red utilizada, se pueden apreciar mejor en la fig. 60.

Figura 61. Distribución de pesos y ganancias en el modelo de red propuesto



Donde $IW\{1,1\}$ = Pesos del conjunto de entrada.

$LW\{2,1\}$ y $LW\{3,2\}$ = Pesos de la capa de entrada y oculta, respectivamente.

$b\{1\}, b\{2\}$ y $b\{3\}$ = Vectores de ganancias.

El conjunto de pesos iniciales utilizado en la prueba se muestra a continuación.

$IW\{1,1\}$ = [-0.20128 2.1464 -2.4431;
0.039869 -8.5996 -1.1027;
2.0925 6.3383 17.1823]

$LW\{2,1\}$ = [-11.948 5.1451 -3.7449;
-10.017 7.4407 1.23;
-16.6817 -3.5711 4.0365;
3.2938 -11.3902 -5.8809;
3.4632 -3.1909 11.3174;
13.0179 3.9013 4.6732;
4.3438 9.8159 6.2001;
-1.8299 9.3456 -6.5369;
-7.8233 9.017 1.7247]

$LW\{3,2\}$ = [5.1548 1.416 -3.2732 5.0554 -2.2513 1.7981 -1.946 -0.97738
-1.7666]

$b\{1\}$ = [1.7066;
2.8412;
-5.2776]

b{2}= [10.75;
-0.86649;
5.5871;
3.8127;
-6.044;
-3.4314;
-3.0958;
-5.2084;
-6.5231]

b{3}= [-0.29222]

Los datos correspondientes al pozo de prueba, se encuentran consignados en el anexo B.

Figura 62. Simulación de la RNA en matlab, dado un conjunto de pesos iniciales

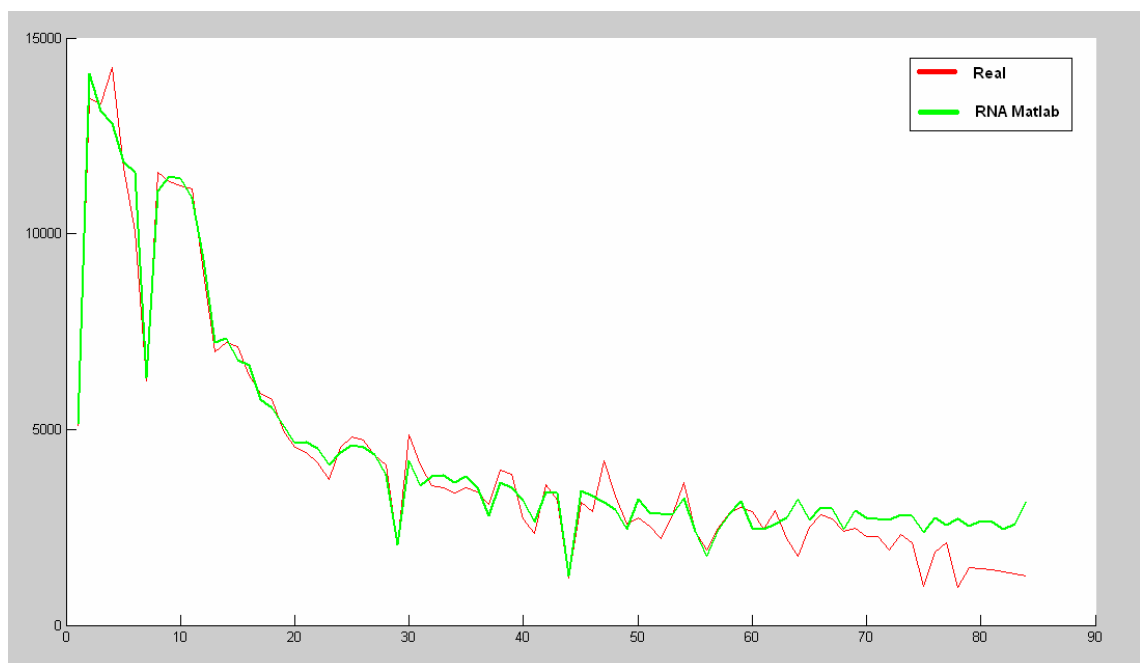
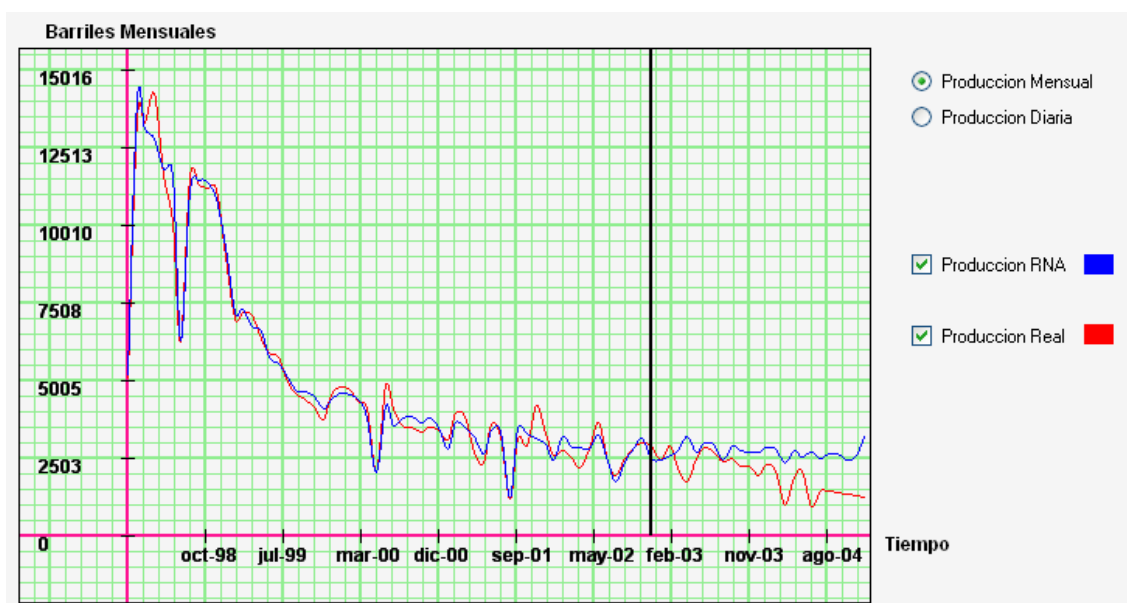


Figura 63. Simulación de la RNA en DECLINET, dado un conjunto de pesos iniciales



6. CONCLUSIONES

El trabajo investigativo introduce mejoras en la predicción de la producción y en los procesos relacionados al análisis de curvas de declinación, frente a técnicas matemáticas clásicas y algunas otras basadas en redes neuronales.

De un buen análisis, del correcto planteamiento y de eficientes técnicas de preprocesamiento de datos, depende el éxito del modelo neuronal.

Contar con un buen modelo neuronal, no siempre es garantía para obtener buenos resultados en la predicción de la producción; las condiciones de operación de pozos depende en gran medida de las decisiones tomadas por los ingenieros de producción; es por esto que la red no podría predecir y ajustarse a malos procedimientos que conlleven a un bajo rendimiento del pozo o a un daño irreparable.

Las técnicas clásicas de análisis de curvas de declinación tienen la limitante de que, luego de una operación de mantenimiento o de mejora, no representan adecuadamente la nueva tendencia de producción.

Si se presentan operaciones de mantenimiento o mejoramiento de la producción, dentro de la etapa de simulación que no hayan sido contempladas durante la fase de entrenamiento, es posible que la red neuronal no arroje los resultados esperados

A diferencia de las técnicas clásicas que aproximan la producción de un pozo petrolero basadas en métodos analíticos, matemáticos y gráficos, nuestra

propuesta obtiene los resultados a partir del procesamiento de las condiciones de operación actuales, lo que permite hacer una predicción más cercana a los datos reales. Es por eso que los análisis derivados de esta predicción, tales como el cálculo del límite económico, la tasa de declinación y la estimación de la producción acumulada, resultan ser más confiables.

Los métodos tradicionales para el análisis de curvas de declinación, suponen que la tasa de declinación es constante a lo largo de la vida del pozo. En la práctica encontramos que este valor presenta, con frecuencia, variaciones en el tiempo. Nuestra herramienta es capaz de construir la gráfica y tabla de valores de estas variaciones dada una periodicidad.

Cuando se cuenta con información histórica reducida, es conveniente utilizar las técnicas clásicas para hacer una aproximación global del rendimiento del pozo.

Luego de hacer las comparaciones de la herramienta desarrollada frente a las técnicas derivadas de las ecuaciones de Arps y basándose en los resultados e indicadores de error obtenidos, es posible afirmar que el trabajo realizado ofrece una forma útil y válida para el análisis de curvas de declinación.

El lenguaje de modelado UML facilitó la documentación de todos los procesos presentes en el desarrollo de la aplicación; para una mejor organización y estructuración del código se empleó la programación orientada a objetos, que brinda simplificación, reutilización y encapsulamiento; además algunas técnicas numéricas fueron aplicadas en procesos que así lo requirieron, dando de esta forma, utilidad práctica, a muchos de los conceptos teóricos aprendidos en nuestra formación como ingenieros de sistemas.

7. RECOMENDACIONES

Complementar el modelo de red actual, para que no solo cubra el análisis de pozos individuales, sino que además pueda analizar y describir el comportamiento de todo un yacimiento.

Buscar Información más completa y detallada de la producción de los pozos a analizar, de tal forma que puedan identificarse nuevas variables que contribuyan al mejoramiento de las predicciones.

Implementar nuevas técnicas de programación que permitan gestionar de una forma más eficiente los recursos del sistema para mejorar el rendimiento de la aplicación, sobre todo en las operaciones numéricas de alto procesamiento.

Aplicar técnicas complementarias al análisis de curvas de declinación, para realizar un análisis más profundo de los resultados obtenidos en nuestra aplicación.

8. BIBLIOGRAFIA

- [1] H.H. Nguyen and C.W. Chan, "Petroleum Production Prediction: A Neural Network Approach", Proceedings of 5th International Joint Conference on Engineering Design and Automation 2001 (EDA 2001), 5-8 August 2001, Las Vegas, USA, pp.85-90, 2000
- [2] H.H. Nguyen and C.W. Chan, "A Comparison of Data Preprocessing Strategies for Neural Network Modeling of Oil Production Prediction", Proceedings of the Third IEEE International Conference on Cognitive Informatics, (ICCI 2004)
- [3] McCulloch, WS and Pitts, W: "A Logical Calculus of Ideas Immanent in Nervous Activity." , Bull. Mathematical Biophysics (1943) 5,15
- [4] Rosenblatt, F: "The Perceptron: Probabilistic Model for Information Storage an Organization in the brain," Psychol. Rev. (1958) 65, 386
- [5] Widrow, B.: "Generalization and Information Storage in Networks of Adeline Neurons," Self-Organizing Systems, M.C. Yovitz, Jacobi, and G.D. Goldstein (eds.), Spartan Books, Washintong, DC.
- [6] Minsky, M.L. and Papert, S.A.: Perceptrons, MIT Press, Cambridge, Massachusetts (1969)
- [7] Hopfield, J.J.: "Neural Networks and Phiysical Systems ulT Emergent Collective Computacional Habilidadéis," Proc., Natl. Academy of science (1982) 79, 2554

- [8] Asociación colombiana de ingenieros de petróleo: "Memorias del seminario: Ingeniería de petróleos para no Petroleros", Bogotá D.C., septiembre 11 al 15 de 2006.
- [9] Hilera, J.R., y V.J. Martínez. 2000. Redes neuronales artificiales. Fundamentos, modelos y aplicaciones. 390 p. Editorial Alfaomega Ra-Ma, Madrid, España.
- [10] Arnold, R. and Anderson R.: "Preliminary Report on Coalinga Oil District," U.S. *Geological Survey Bull.* (1908) 357,79
- [11] Cutler, W. W. Jr.: "Estimation of Underground Oil Reserves by Well Production Curves," U.S. Bureau of Mines Bull. (1924) 228.
- [12] Pirson, S. J.: "Production Decline Curve of Oil Wells May Be Extrapolated by Loss Ratio Method," Oil & Gas J. (1935) 34, N° 46,34.
- [13] Arps, J. J.: "Analysis of Decline Curves." Trans., AIME (1945) 160,228
- [14] Arps, J. J.: "Estimation of Primary Oil Reserves," Trans., AIME (1956) 207,182
- [15] Slider, H. C.: "A Simplified Method of Hyperbolic Decline Curve Analysis," JPT (march/1968) 235.
- [16] Gentry, R. W.: "Decline Curve Analysis," JPT (January/1972) 38.
- [17] Gentry, R. W. and McCray, A. W.: "The Effect of Reservoir and Fluid Properties on Production Decline Curves," JPT (September/1978) 1327.
- [18] Long, D. R. and Davis, M.J.: "A New Approach to the Hyperbolic Curve," JPT (July/1988) 909.

[19] Rodriguez, F. and Cinco-Ley, H.: "A New Model for Production Decline," Paper SPE, (March / 1993)

[20] Fetkovich, M.J., Fetkovich, E.J., and Fetkovich, M.D.: "Useful Concepts for Decline-Curve Forecasting, Reserve Estimation, and Analysis," Paper *SPE* (February/ 1996) 13-22.

[21] Agarwal, R.G., Gardner, D.C., Kleinsteinber, S.W. and Fussell, D.D.: "analyzing well production data using combined type-curve and decline-curve analysis concepts," paper SPE 49222, 1998

[22] Masoner, L.O.: "Decline Analysis' Relationship to Relative Permeability in Secondary and Tertiary Recovery," paper SPE 39928, (April / 1998)

[23] F. Aminzadeh, J. Barhen, and N.B. Toomarian, "Estimation of Reservoir Parameter using a Hybrid Neural Network", *Journal of Petroleum Science and Engineering*, vol. 24, no. 1, 1999, pp. 49-56.

[24] H.H. Nguyen and C.W. Chan, "Petroleum Production Prediction: A Neural Network Approach", Proceedings of 5th International Joint Conference on Engineering Design and Automation 2001 (EDA 2001), 5-8 August 2001, Las Vegas, USA, pp.85-90, 2000

[25] Abdelhafidh, F. and Djebbar, T. : "Application of Decline-Curve Analysis Technique in Oil Reservoir Using A Universal Fitting Equation" , paper SPE 70036, (May /2001)

[26] Li, K. and Horne, R.N.: "A Decline Curve Analysis Model Based on Fluid Flow Mechanisms," paper SPE 83470, (May /2003).

- [27]** H.H. Nguyen and C.W. Chan, "A Decision Support System For Oil Production Prediction", Proceedings of Electrical and Computer Engineering. Canadian Conference, (CCECE 2004)
- [28]** Camacho, R. , Fuentes, G, Vásquez, M.: "Decline Curve Analysis of Fractured Reservoirs with Fractal Geometry", paper SPE 104009, (september/ 2006)
- [29]** Slider, H. C., Worldwide Practical Petroleum Reservoir Engineering Methods. TULSA, OK. : PENNWELL BOOKS, 1983.
- [30]** T.e.w., Nind. Fundamentos De Produccion Y Mantenimiento De Pozos Petroleros. Mexico : Limusa, 1987. 477 p.

9. ANEXO A: MANUAL DE USUARIO

9.1 INTRODUCCION

DECLINET 1.0 es una herramienta computacional basada en redes neuronales para predecir la producción de un pozo petrolero a partir de los datos históricos, que corresponden con su producción pasada. A partir de aquí se le orientará para que maneje el aplicativo arrancando desde cero.

Son varios los procesos que realiza DECLINET 1.0 durante el estudio de los datos de un pozo particular: carga de datos, entrenamiento, simulación y presentación de resultados.

9.2 CARGA DE DATOS.

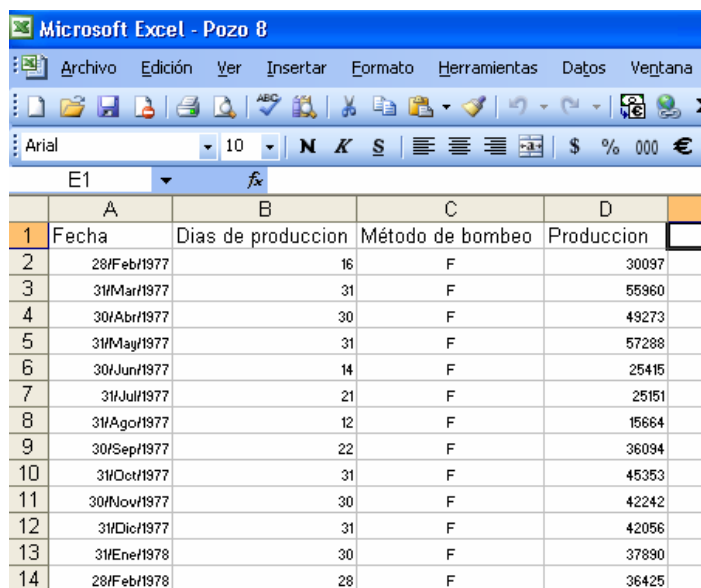
Los datos de los que se alimenta el software se encuentran consignados en archivos de Excel con un formato de 4 columnas, donde la primera de ellas corresponde con la fecha, la segunda indica los días de producción, la tercera pertenece al tipo de bombeo y la última columna es la producción del pozo. La Fig. 63 muestra el formato válido para la creación del archivo.

Los diferentes tipos de bombeo deben ser codificados de la siguiente forma:

Tabla 41. Codificación de los tipos de bombeo

Tipo de Bombeo	Codificación
Flujo Natural	F
Mecánico	B
Electro-sumergible	BE
Hidráulico	BH
Otros	Otros

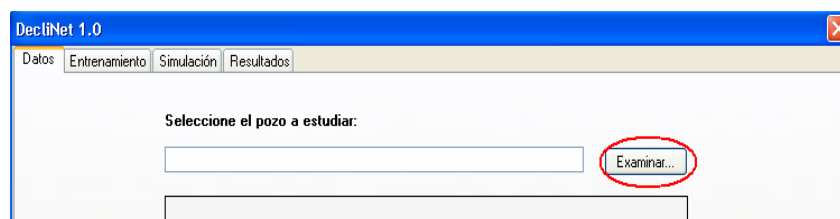
Figura 64. Formato de archivo en Excel



	A	B	C	D
1	Fecha	Dias de produccion	Método de bombeo	Produccion
2	28/Feb/1977	16	F	30097
3	31/Mar/1977	31	F	55960
4	30/Abr/1977	30	F	49273
5	31/May/1977	31	F	57288
6	30/Jun/1977	14	F	25415
7	31/Jul/1977	21	F	25151
8	31/Ago/1977	12	F	15664
9	30/Sep/1977	22	F	36094
10	31/Oct/1977	31	F	45353
11	30/Nov/1977	30	F	42242
12	31/Dic/1977	31	F	42056
13	31/Ene/1978	30	F	37890
14	28/Feb/1978	28	F	36425

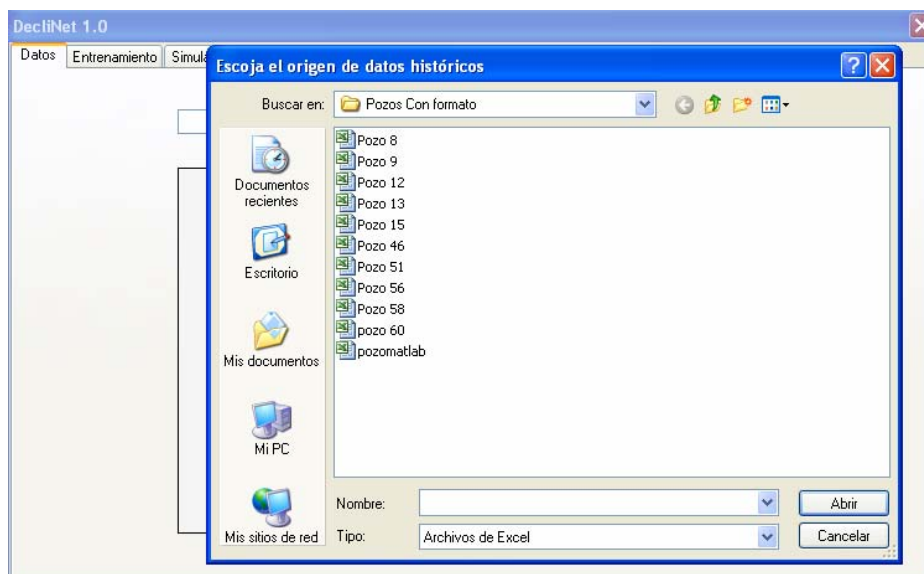
DECLINET provee el método necesario para cargar la información, mediante la forma estándar para abrir archivos, verificando en el proceso que este sea valido.

Figura 65. DECLINET 1.0: Formas de seleccionar el pozo a estudiar



La herramienta le permite escribir la ubicación del archivo directamente o en su defecto explorar en el disco duro pulsando sobre examinar.

Figura 66. DECLINET 1.0: Abrir archivos de alimentación



Esta ventana nos permite seleccionar el archivo deseado. Después de cargar la información deben aparecer los datos como se muestra en la siguiente figura.

Figura 67. DECLINET 1.0: Datos del pozo escogido

DECLINET 1.0

Datos Entrenamiento Simulación Resultados

C:\Documents and Settings\Guido Polo Nule\Mis documentos\PRD\Pozos C Examinar...

	Fecha	Días de producción	Método de bombeo	Producción
►	28/02/1977	16	F	30097
	31/03/1977	31	F	55960
	30/04/1977	30	F	49273
	31/05/1977	31	F	57288
	30/06/1977	14	F	25415
	31/07/1977	21	F	25151
	31/08/1977	12	F	15664
	30/09/1977	22	F	36094
	31/10/1977	31	F	45353
	30/11/1977	30	F	42242
	31/12/1977	31	F	42056
	31/01/1978	30	F	37890

9.3 ENTRENAMIENTO.

El proceso de entrenamiento es quizás el más importante dentro del análisis de curvas de declinación, aplicando redes neuronales. DECLINET 1.0 no permite realizar entrenamiento si previamente no se han cargado datos o si estos son erróneos.

Figura 68. DECLINET 1.0: Selección de parámetros de entrenamiento

The screenshot shows the DECLINET 1.0 software window with the 'Entrenamiento' (Training) tab selected. The interface is divided into several sections:

- El sistema ha detectado (System has detected):** This section contains three input fields: 'Inicio de producción' (Start of production) with the value '28/02/1977', 'final de producción' (End of production) with the value '31/12/2004', and 'Total de datos' (Total data) with the value '335'.
- Entrenamiento (Training):** This section contains four input fields: 'Utilizar el' (Use) with a dropdown menu set to '70' (circled in red), followed by '% de los datos' (of the data); 'Num. datos' (Number of data) with the value '234'; 'Fin entrenamiento' (End training) with the value '31/07/1996'; and 'Algoritmo de entrenamiento:' (Training algorithm:) with the value 'Levenberg-Marquadt'.
- Resultados de entrenamiento (Training results):** This section contains four input fields: 'Tiempo' (Time) with a unit 'S', 'Error cuadrático medio' (Mean square error), 'Error absoluto' (Absolute error), and 'Tasa de ajuste' (Adjustment rate) with a unit '%'. Below these fields is a progress bar ranging from 0% to 100%.
- Buttons:** At the bottom right, there is a button labeled 'Entrenar' (Train), which is circled in red.

En este paso el programa detecta algunas características del pozo tales como las fechas de inicio y fin de la producción, al igual que el total de datos.

Para el entrenamiento el usuario puede decidir que cantidad de datos desea utilizar. Emplear muchos datos no es sinónimo ni garantía de una buena generalización por parte de la red, se recomienda usar el 70% o el 80% en la mayoría de los casos.

Cuando se haya establecido el volumen de entrenamiento se pulsa sobre el botón de entrenar para iniciar el proceso. En la siguiente gráfica se puede apreciar los resultados de esta operación.

Figura 69. DECLINET 1.0: Resultados del entrenamiento

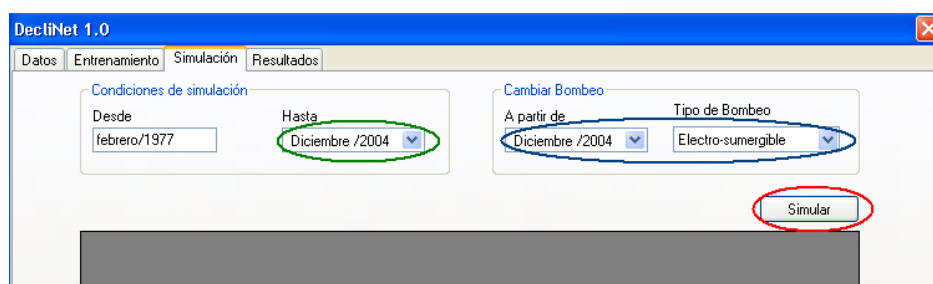
The screenshot shows the DECLINET 1.0 software interface with the 'Resultados' tab selected. The interface is divided into several sections. On the left, under 'El sistema ha detectado', there are three input fields: 'Inicio de producción' (28/02/1977), 'final de producción' (31/12/2004), and 'Total de datos' (335). Below this, under 'Entrenamiento', there are four input fields: 'Utilizar el' (70 % de los datos), 'Num. datos' (234), 'Fin entrenamiento' (31/07/1996), and 'Algoritmo de entrenamiento' (Levenberg-Marquadt). On the right, under 'Resultados de entrenamiento', there are four input fields: 'Tiempo' (1.828 s), 'Error cuadrático medio' (0.0019567), 'Error absoluto' (0.0332769), and 'Tasa de ajuste' (85.74 %). At the bottom right, there is a progress bar labeled 'Entrenamiento Finalizado' showing 100% completion. A green arrow points to the 'Entrenar' button at the bottom center.

Los valores señalados corresponden a indicadores que informan sobre la efectividad del entrenamiento. En el cuadro de resultados tenemos en primer lugar el tiempo de procesamiento consumido, que brinda información acerca del rendimiento del algoritmo de entrenamiento; seguidamente se encuentran los errores cuadrático y absoluto y adicional a los valores anteriores está la tasa de ajuste que es un estimativo expresado en porcentaje de que tanto se aproxima la red a los datos reales durante el entrenamiento, siendo el 100 % una aproximación perfecta.

9.4 SIMULACIÓN

En la etapa de simulación es donde DECLINET calcula los valores de producción. Para realizar este proceso, se hace necesario que la red haya sido entrenada.

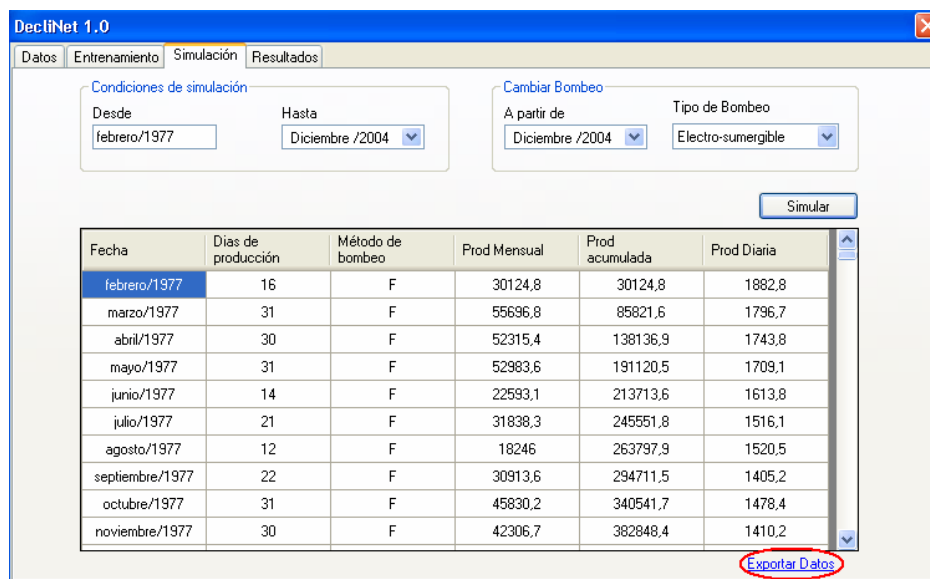
Figura 70. DECLINET 1.0: Selección de los parámetros de simulación



En la ventana de simulación se encuentran varias opciones que hacen de este proceso uno muy flexible. El usuario puede escoger la fecha hasta donde desee simular la producción, escogiendo el valor correspondiente (óvalo verde); adicionalmente puede simular un cambio en el método de bombeo escogiendo la fecha y el nuevo tipo de bombeo (óvalo azul); cuando se hayan establecidos todos los valores de simulación, se pulsa sobre el botón simular.

NOTA: Para que el cambio de bombeo tenga efecto, la fecha final de simulación debe ser mayor que la fecha de fin de producción, establecida en el archivo de entrada.

Figura 71. DECLINET 1.0: Resultados de la simulación

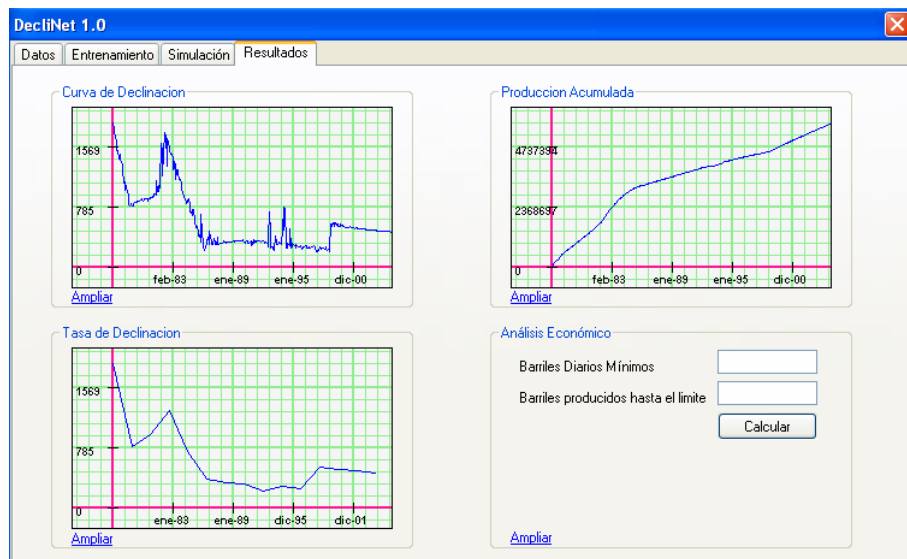


Luego de simular, aparecen los datos tal como lo muestra la figura anterior. DECLINET 1.0 brinda la opción de exportar esta información a un archivo de Excel, de tal forma que el usuario pueda darle un tratamiento diferente a los datos o para la realización de reportes personalizados. Al pulsar el link [Exportar datos](#), aparece un cuadro de diálogo, en el cual se debe indicar la ubicación donde se desee guardar los resultados de la simulación.

9.5 PRESENTACIÓN DE RESULTADOS

Esta es una opción que brinda al usuario los resultados de la simulación de forma gráfica, lo que da un panorama más amplio del comportamiento de la producción del pozo.

Figura 72. DECLINET 1.0: Presentación grafica de resultados

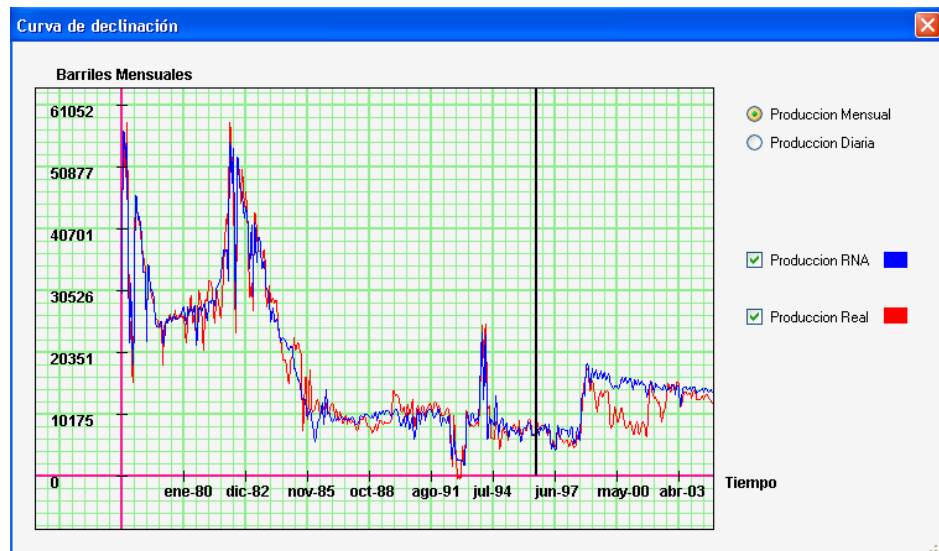


Como se observa en la figura esta ventana ofrece diferentes tipos de resultados, entre ellos la curva de declinación, la producción acumulada, la tasa de declinación y el cálculo del límite económico. Cada una de estos resultados tiene la opción de ser ampliados para la realización de un análisis más detallado.

En las gráficas del análisis detallado de las curvas de declinación y producción acumulada, aparece una línea negra que indica el final de los datos de entrenamiento.

9.5.1 Curvas De Declinación.

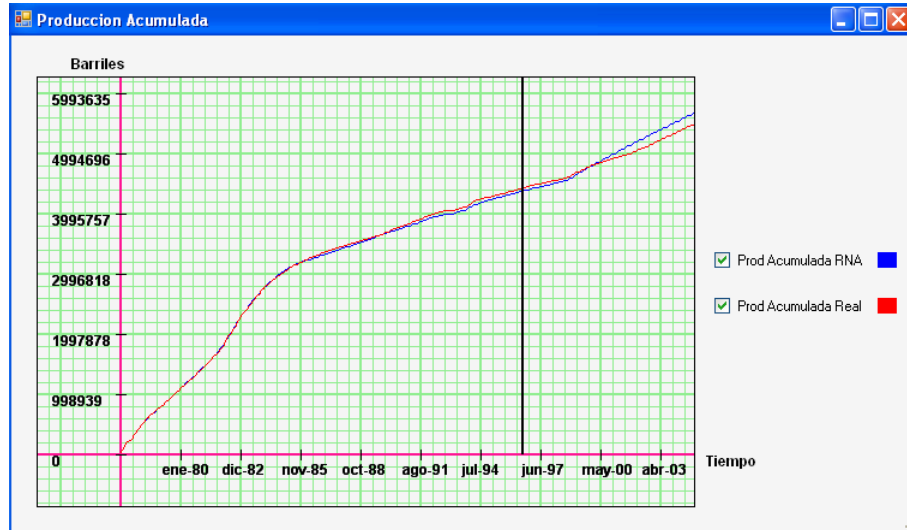
Figura 73. DECLINET 1.0: Comparación gráfica entre la curva de producción real y la simulada



La figura anterior es la ventana ampliada de la curva de declinación donde además de poder graficar la curva simulada, esta puede ser comparada contra la curva real de producción para verificar la validez del procedimiento; adicionalmente DECLINET ofrece la posibilidad de graficar en 2 escalas diferentes, que son producción mensual o diaria según la preferencia del usuario.

9.5.2 Producción Acumulada

Figura 74. DECLINET 1.0: Comparación grafica entre la producción acumulada real y la simulada



La gráfica anterior es la ventana ampliada de la producción acumulada, donde el usuario tiene también la posibilidad de poder graficar una o ambas curvas, para conocer la distribución global de la producción.

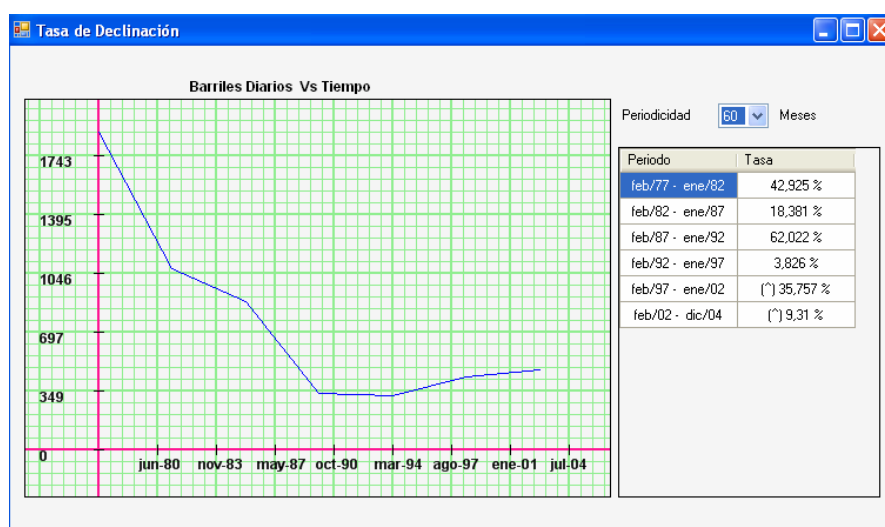
9.5.3 Tasa De Declinación

Figura 75. DECLINET 1.0: Gráfica de la tasa de declinación con periodicidad de 24 meses



La ventana ampliada de la tasa de declinación muestra información completa acerca de la tendencia declinante o ascendente de la producción, tanto en forma gráfica como tabular; además es posible escoger diferentes periodicidades entre 6 y 60 meses para tener una visión completa de las variaciones en la producción. En la siguiente figura se muestra la información del pozo anterior pero con una periodicidad diferente.

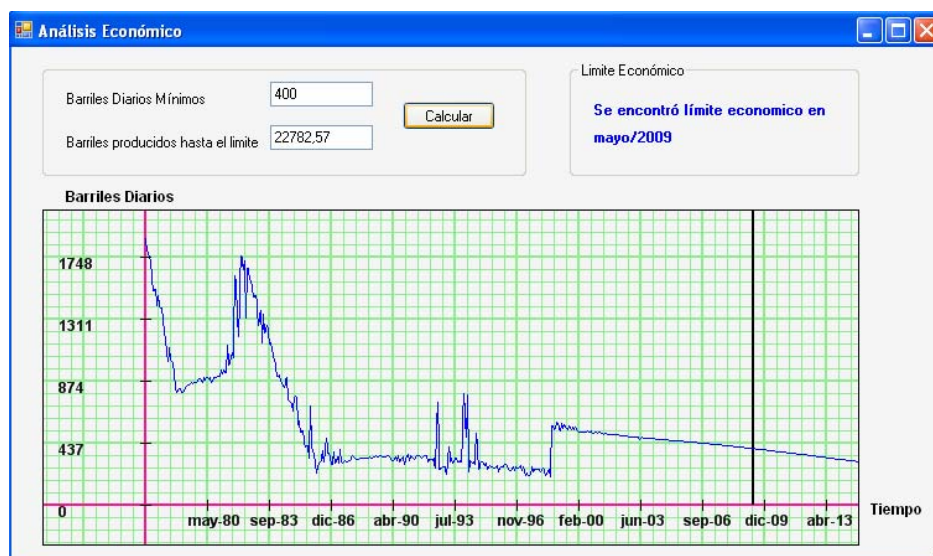
Figura 76. DECLINET 1.0: Gráfica de la tasa de declinación con periodicidad de 60 meses



9.5.4 Limite Económico

Esta es una de las herramientas que ofrece DECLINET para hacer más completo el análisis de curvas de declinación. Es así que, dada cierta información de rentabilidad, la herramienta es capaz de encontrar la fecha en que se estima que el pozo agoto su vida útil.

Figura 77. DECLINET 1.0: Presentación grafica y detallada de el calculo del limite económico



Como muestra la figura, dado unos barriles mínimos diarios de producción, DECLINET se encarga de estimar los barriles que se producen hasta alcanzar el límite económico, además lo muestra de manera intuitiva mediante una gráfica, con un línea vertical que indica la fecha (mes y año) en la que el pozo alcanza el limite.

ANEXO B: DATOS DE PRODUCCIÓN DEL POZO DE PRUEBA

Tabla 42. Datos de producción del pozo de prueba

Fecha	Días de producción	Método de bombeo	Producción
31-Ene-98	10	BE	5099
28-Feb-98	31	BE	13453
31-Mar-98	30	BE	13326
30-Abr-98	31	BE	14242
31-May-98	30	BE	11686
30-Jun-98	31	BE	10046
31-Jul-98	18	BE	6241
31-Ago-98	28	BE	11573
30-Sep-98	31	BE	11327
31-Oct-98	30	BE	11231
30-Nov-98	31	BE	11155
31-Dic-98	31	BE	9006
31-Ene-99	29	BE	6984
28-Feb-99	31	BE	7216
31-Mar-99	30	BE	7114
30-Abr-99	31	BE	6380
31-May-99	30	BE	5918
30-Jun-99	31	BE	5773
31-Jul-99	31	BE	4964
31-Ago-99	30	BE	4540
30-Sep-99	31	BE	4420
31-Oct-99	30	BE	4140
30-Nov-99	27	BE	3730
31-Dic-99	29	BE	4556
31-Ene-00	31	BE	4823
29-Feb-00	31	BE	4727
31-Mar-00	30	BE	4326
30-Abr-00	27	BE	4087
31-May-00	14	BE	2055
30-Jun-00	31	BE	4860
31-Jul-00	27	BE	4102
31-Ago-00	29	BE	3576
30-Sep-00	30	BE	3518
31-Oct-00	29	BE	3366
30-Nov-00	31	BE	3518
31-Dic-00	29	BE	3406
31-Ene-01	25	BE	3083
28-Feb-01	31	BE	3961
31-Mar-01	30	BE	3863
30-Abr-01	28	BE	2716
31-May-01	25	BE	2354
30-Jun-01	30	BE	3591

31-Jul-01	30	BE	3197
31-Ago-01	16	BE	1193
30-Sep-01	31	BE	3136
31-Oct-01	30	BE	2902
30-Nov-01	29	BE	4210
31-Dic-01	28	BE	3303
31-Ene-02	25	BE	2578
28-Feb-02	30	BE	2746
31-Mar-02	28	BE	2522
30-Abr-02	28	BE	2216
31-May-02	28	BE	2822
30-Jun-02	31	BE	3634
31-Jul-02	24	BE	2393
31-Ago-02	18	BE	1929
30-Sep-02	24	BE	2488
31-Oct-02	29	BE	2864
30-Nov-02	31	BE	3020
31-Dic-02	27	BE	2899
31-Ene-03	27	BE	2454
28-Feb-03	28	BE	2928
31-Mar-03	29	BE	2212
30-Abr-03	22	BE	1763
31-May-03	29	BE	2510
30-Jun-03	31	BE	2831
31-Jul-03	31	BE	2722
31-Ago-03	28	BE	2386
30-Sep-03	31	BE	2482
31-Oct-03	30	BE	2269
30-Nov-03	30	BE	2272
31-Dic-03	30	BE	1934
31-Ene-04	31	BE	2323
29-Feb-04	31	BE	2120
31-Mar-04	28	BE	984
30-Abr-04	31	BE	1880
31-May-04	30	BE	2103
30-Jun-04	31	BE	967
31-Jul-04	30	BE	1460
31-Ago-04	31	BE	1447
30-Sep-04	31	BE	1410
31-Oct-04	30	BE	1354
30-Nov-04	31	BE	1317
31-Dic-04	25	BE	1264