



**INSTITUTO POLITÉCNICO NACIONAL**

ESCUELA SUPERIOR DE INGENIERÍA  
QUÍMICA E INDUSTRIAS EXTRACTIVAS  
SECCIÓN DE ESTUDIOS DE POSGRADO E  
INVESTIGACIÓN



**CONSTRUCCIÓN DE UNA RED NEURONAL ARTIFICIAL  
PARA LA MODELACIÓN DE UN REACTOR DE  
HIDRODESULFURACIÓN DE NAFTAS**

**T E S I S**

QUE PARA OBTENER EL GRADO DE  
MAESTRO EN CIENCIAS CON ESPECIALIDAD  
EN INGENIERÍA QUÍMICA

PRESENTA:

**JOSÉ IGNACIO PAZ PAREDES**

DIRECTOR DE TESIS

**DR. ENRIQUE ARCE MEDINA**

CO-DIRECTOR

**DR. ROBERTO LIMAS BALLESTEROS**

JUNIO, 2004.

MÉXICO, D.F.





CGRI-14

**INSTITUTO POLITECNICO NACIONAL**  
**COORDINACION GENERAL DE POSGRADO E INVESTIGACION**

*ACTA DE REVISION DE TESIS*

En la Ciudad de México, D. F. siendo las 11:30 horas del día 16 del mes de Abril 2004 se reunieron los miembros de la Comisión Revisora de Tesis designada Por el Colegio de Profesores de Estudios de Posgrado e Investigación de la ESIQIE para examinar la tesis de grado titulada:  
"CONSTRUCCIÓN DE UNA RED NEURONAL ARTIFICIAL PARA LA MODELACIÓN DE UN REACTOR DE HIDRODESULFURACIÓN DE NAFTA"

Presentada por el alumno:

PAZ PAREDES JOSE IGNACIO  
Apellido paterno matemp nombre(s)  
Con registro: 

9	2	0	3	2	2
---	---	---	---	---	---

Aspirante al grado de:

MAESTRIA EN CIENCIAS CON ESPECIALIDAD EN INGENIERIA QUÍMICA

Después de intercambiar opiniones los miembros de la Comisión manifestaron **SU APROBACION DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISION REVISORA

Director de Tesis		Co-director
<u>DR. ENRIQUE ARCE MEDINA</u>		<u>DR. ROBERTO LINAS BALLESTEROS</u>
<u>DR. RICARDO MACIAS SALINAS</u>		<u>DR. MIGUEL A. VALENZUELA ZAPATA</u>
<u>DR. JOSE JAVIER CASTRO ARELLANO</u>		<u>DR. DAVID ROMERO ROMERO</u>

EL PRESIDENTE DEL COLEGIO

DR. ROBERTO LINAS BALLESTEROS





**INSTITUTO POLITECNICO NACIONAL**  
**COORDINACIÓN GENERAL DE POSGRADO E INVESTIGACIÓN**

**CARTA CESION DE DERECHOS**

En la Ciudad de México, D.F. el día 28 del mes Septiembre del año 2004, el (la) que suscribe José Ignacio Paz Paredes alumno (a) del Programa de Maestría en Ciencia con Especialidad en Ingeniería Química con número de registro 920322, adscrito a la Sección de Estudios de Posgrado e Inv.-ESIQIE, manifiesta que es autor (a) intelectual del presente trabajo de tesis bajo la dirección de Dr. Enrique Arce Medina/Dr. Roberto Linas Ballesteroy ceden los derechos del trabajo intitulado "Construcción de una Red Neuronal Artificial para la Modelación de un Reactor de Hidrodesulfuración de Naftas", al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y del director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección Av. Sn. Juan No. 183 San Agustín, Edo. de México. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

José Ignacio Paz Paredes

*Nombre y firma*



# **DEDICATORIAS**

El desarrollo, elaboración y conclusión del presente trabajo esta dedicado a mi bien amada esposa **Ino Cordero Martínez**. Nunca voy a dejar de dar gracias a la vida por tener la suerte de encontrarte y permitirme compartir la vida contigo. Gracias por tu amor, y apoyo incondicional. Gracias por darme unos hijos maravillosos; siendo tú misma una mujer excepcional, maravillosa, llena de ternura y amor.

También dedico este trabajo a mis hijos, **Julio Alberto, Luis Omar, y Jorge Emilio**, ojalá les sirva como motivación para que forjen sus propios caminos y alcancen sus propias metas.

Con todo mi amor,

José Ignacio Paz Paredes



# **AGRADECIMIENTOS**

Al **Instituto Mexicano del Petróleo** por su apoyo económico durante el periodo de estudio de la maestría; y en la preparación de esta tesis. En particular a las áreas de: **Subdirección de Comercialización, Gerencia de Ingeniería Básica, Gerencia de Aplicación Industrial de Catalizadores, y a la Coordinación del Programa de Tratamiento de Crudo Maya**; y dentro de ellas:

Al **Ing. Antonio Torres Arellano** por su apoyo en la realización de los estudios de la maestría.

Al **Ing. Juan Antonio Rodríguez García** por su motivación en la culminación de este trabajo de tesis, por su apoyo incondicional y por su invaluable amistad.

Al **M. en I. César Edmundo Figueroa Gutiérrez**, por su asesoría y su apoyo.

Al **M. en C. Pedro Martín Vega Merino**, por permitirme el acceso a la información experimental, pero sobre todo por su amistad.

Al **M. en C. Agustín González Ortiz**, por su continua motivación para la culminación de este trabajo.

Al **M. en C. Arturo González Sánchez**, y al **Dr. Martín Aguilera López**, por sus atinados y enriquecedores comentarios al presente trabajo.

A la **M. en C. Claudia de la Paz Zavala**, por proporcionarme la información del simulador cinético *HidroNaf* y por su disposición de apoyo para la comparación entre los modelos.

Al comité de revisores asignado por la **SEPI** de la **ESIQIE**, formado por los **Drs.: Roberto Limas Ballesteros, Miguel Ángel Valenzuela Zapata, Javier Castro Arellano, Ricardo Macías Salinas**; y al **Dr. David Romero Romero**, profesor de la **SEPI-ESIME**; por sus atinados comentarios que permitieron enriquecer ampliamente este trabajo.

Al **Instituto Politécnico Nacional** por permitirme realizar en esa institución los estudios que me han llevado a ser lo que soy.

A la **Sección de Posgrado e Investigación** de la **Escuela Superior de Ingeniería Química e Industrias Extractivas**, y dentro de ella vaya mi más sincero reconocimiento al **Dr. Enrique Arce Medina**, por la acertada dirección de este trabajo, su inagotable paciencia, y por su amistad.



# ÍNDICE

<b>LISTA DE TABLAS</b>	<b><i>i</i></b>
<b>LISTA DE FIGURAS</b>	<b><i>iii</i></b>
<b>NOMENCLATURA</b>	<b><i>vii</i></b>
<b>NOMENCLATURA</b>	<b><i>vii</i></b>
<b>RESUMEN</b>	<b><i>ix</i></b>
<b>ABSTRACT</b>	<b><i>xi</i></b>
<b>INTRODUCCIÓN</b>	<b>1</b>
<b>1. PROCESO DE HIDRODESULFURACIÓN</b>	<b>5</b>
1.1. ANTECEDENTES	5
1.2. FUNDAMENTOS	7
1.3. DESCRIPCIÓN DEL PROCESO	9
1.4. VARIABLES DE OPERACIÓN	12
<b>2. REDES NEURONALES ARTIFICIALES</b>	<b>17</b>
2.1. ANTECEDENTES	17
2.2. DESCRIPCIÓN	21
2.3. MÉTODO DE APRENDIZAJE	24
2.4. APLICACIONES EN LA INGENIERÍA QUÍMICA.	29
SIMULACIÓN DE UNA PLANTA DE TRATAMIENTO DE AGUAS RESIDUALES.	30
MODELACIÓN DE UNA COLUMNA DE EXTRACCIÓN LÍQUIDO-LÍQUIDO.	32
MODELACIÓN DE UNA PLANTA INDUSTRIAL FCC.	35
<b>3. DESARROLLO DEL MODELO</b>	<b>39</b>
3.1. SELECCIÓN DEL ALGORITMO.	39
3.2. DATOS EXPERIMENTALES	41
3.3. DISEÑO DE LA RED NEURONAL.	51
<b>4. DETERMINACIÓN DE LA TOPOLOGÍA ÓPTIMA</b>	<b>57</b>
4.1. TOPOLOGÍA GLOBAL.	57
4.2. TOPOLOGÍA PROCESO.	75
4.3. TOPOLOGÍA PROPIEDADES	85
4.4. SELECCIÓN DE LA TOPOLOGÍA.	93
<b>5. ANÁLISIS DE RESULTADOS</b>	<b>97</b>
5.1. PREDICCIÓN DE DATOS DE PRUEBA.	97
5.2. COMPARACIÓN CON UN MODELO CINÉTICO.	106
<b>CONCLUSIONES</b>	<b>111</b>

<b>GLOSARIO</b>	<b>113</b>
<b>BIBLIOGRAFÍA</b>	<b>115</b>
<b>APÉNDICES</b>	<b>123</b>
<b>A. ALGORITMO DE RETROPROPAGACIÓN</b>	<b>125</b>
<b>B. PROGRAMA NEURONAF</b>	<b>133</b>
<b>C. LISTADO DEL PROGRAMA</b>	<b>137</b>

## **LISTA DE TABLAS**

Tabla 1.1 Unidades Hidrodesulfuradoras de Petróleos Mexicanos .....	6
Tabla 1.2 Características físico-químicas de varias fracciones del petróleo .....	8
Tabla 2.1 Predicción de valores de TOC en una planta de tratamiento de aguas residuales.....	31
Tabla 2.2 Variables del modelo del proceso FCC y sus intervalos. ....	37
Tabla 3.1 Clasificación de los datos experimentales .....	42
Tabla 3.2 Valores de Intervalos de Confianza.....	49
Tabla 3.3 Análisis estadístico de los datos experimentales.....	50
Tabla 3.4 Número de nodos de entrada y salida de los esquemas propuestos.....	54
Tabla 4.1 Clasificación de las variables, topología 10-X-7. ....	57
Tabla 4.2 Errores promedio de la topología 10-6-7, entrenamiento (Caso “Global”). .....	59
Tabla 4.3 Errores promedio de la topología 10-7-7, entrenamiento (Caso “Global”). .....	59
Tabla 4.4 Errores promedio de la topología 10-8-7, entrenamiento (Caso “Global”). .....	59
Tabla 4.5 Errores promedio de la topología 10-9-7, entrenamiento (Caso “Global”). .....	60
Tabla 4.6 Errores promedio de la topología 10-10-7, entrenamiento (Caso “Global”). .....	60
Tabla 4.7 Errores promedio de la topología 10-12-7, entrenamiento (Caso “Global”). .....	60
Tabla 4.8 Errores promedio de la topología 10-14-7, entrenamiento (Caso “Global”). .....	61
Tabla 4.9 Errores promedio de la topología 10-16-7, entrenamiento (Caso “Global”). .....	61
Tabla 4.10 Errores promedio de la topología 10-18-7, entrenamiento (Caso “Global”). ...	61
Tabla 4.11 Errores promedio de la topología 10-20-7, entrenamiento (Caso “Global”). ...	62
Tabla 4.12 Errores de validación de la Red Neuronal “Global”.....	67
Tabla 4.13 Clasificación de las variables, topología 7-X-4. ....	75
Tabla 4.14 Errores promedio de la topología 7-4-4, entrenamiento (Caso “Proceso”). ....	76
Tabla 4.15 Errores promedio de la topología 7-5-4, entrenamiento (Caso “Proceso”). ....	77
Tabla 4.16 Errores promedio de la topología 7-6-4, entrenamiento (Caso “Proceso”). ....	77
Tabla 4.17 Errores promedio de la topología 7-7-4, entrenamiento (Caso “Proceso”). ....	78

Tabla 4.18 Errores promedio de la topología 7-8-4, entrenamiento (Caso “Proceso”). ....	78
Tabla 4.19 Errores promedio de la topología 7-10-4, entrenamiento (Caso “Proceso”)....	78
Tabla 4.20 Errores promedio de la topología 7-12-4, entrenamiento (Caso “Proceso”)....	79
Tabla 4.21 Errores promedio de la topología 7-14-4, entrenamiento (caso “Proceso”). ...	79
Tabla 4.22 Errores de validación de la Red Neuronal “Proceso” .....	79
Tabla 4.23 Clasificación de las variables. Topología “Propiedades” .....	85
Tabla 4.24 Errores promedio de la topología 3-2-3, entrenamiento (Caso “Propiedades”). .....	86
Tabla 4.25 Errores promedio de la topología 3-3-3, entrenamiento (caso “Propiedades”). .....	86
Tabla 4.26 Errores promedio de la topología 3-4-3, entrenamiento (Caso “Propiedades”). .....	87
Tabla 4.27 Errores promedio de la topología 3-5-3, entrenamiento (Caso “Propiedades”). .....	87
Tabla 4.28 Errores promedio de la topología 3-6-3, entrenamiento (Caso “Propiedades”). .....	87
Tabla 4.29 Errores promedio de la topología 3-8-3, entrenamiento (Caso “Propiedades”). .....	87
Tabla 4.30 Errores promedio de la topología 3-10-3, entrenamiento (Caso “Propiedades”). .....	88
Tabla 4.31 Errores de validación para la topología “Propiedades” .....	88
Tabla 4.32 Comparación de parámetros de regresión. ....	94
Tabla 4.33 Valores de los pesos y sesgos para la topología “Proceso”. Capa oculta.....	95
Tabla 4.34 Valores de los pesos y sesgos para la topología “Proceso”. Capa de salida. .	95
Tabla 4.35 Valores de los pesos y sesgos para la topología “Propiedades”. Capa oculta	96
Tabla 4.36 Valores de los pesos y sesgos para la topología “Propiedades”. Capa de salida. ....	96
Tabla 5.1 Por ciento de error relativo promedio de predicción. Caso Prueba.....	104
Tabla 5.2 Resultados obtenidos por el simulador cinético <i>Hidro-Naf</i> .....	107
Tabla 5.3 Resultados obtenidos por el simulador <i>NeuroNaf</i> .....	108
Tabla 5.4 Comparación entre simuladores.....	108

## **LISTA DE FIGURAS**

Figura 1.1 Diagrama de Flujo de Proceso de Hidrodesulfuración de Naftas .....	10
Figura 1.2 Efecto de las variables del proceso de hidrodesulfuración. ....	15
Figura 2.1 Configuración típica de una Red Neuronal Artificial .....	23
Figura 2.2 Etapas de procesamiento dentro de una neurona.....	23
Figura 2.3 Función sigmoideal. ....	25
Figura 3.1 Variación de la carga. ....	42
Figura 3.2 Variación de la temperatura de operación.....	43
Figura 3.3 Variación de la presión de operación. ....	43
Figura 3.4 Variación de la relación H <sub>2</sub> /HC .....	43
Figura 3.5 Variación de la pureza de H <sub>2</sub> .....	44
Figura 3.6 Variación del consumo de H <sub>2</sub> . ....	44
Figura 3.7 Variación de la TIE de carga.....	44
Figura 3.8 Variación de la TFE de carga.....	45
Figura 3.9 Variación del peso específico de carga .....	45
Figura 3.10 Variación del azufre de carga.....	45
Figura 3.11 Variación del flujo de producto .....	46
Figura 3.12 Variación de la TIE de producto. ....	46
Figura 3.13 Variación de la TFE de producto. ....	46
Figura 3.14 Variación del peso específico del producto. ....	47
Figura 3.15 Variación de azufre en el producto.....	47
Figura 3.16 Variación de la $\Delta T$ .....	47
Figura 3.17 Variación de la $\Delta P$ .....	48
Figura 3.18 Esquema de variables de entrada-salida para el caso 10-X-7. ....	52
Figura 3.19 Esquema de la topología “Proceso”. ....	53
Figura 3.20 Esquema de la topología “Propiedades”. ....	53
Figura 3.21 Diagrama de flujo de la determinación de una red neuronal óptima.....	55

Figura 4.1 Producción, fase de entrenamiento.....	62
Figura 4.2 TIE del producto, fase de entrenamiento. ....	63
Figura 4.3 TFE del producto, fase de entrenamiento .....	63
Figura 4.4 Peso específico del producto, fase de entrenamiento .....	64
Figura 4.5 Azufre en el producto, fase de entrenamiento.....	64
Figura 4.6 $\Delta T$ del lecho catalítico, fase de entrenamiento.....	65
Figura 4.7 Caída de presión del lecho catalítico, fase de entrenamiento. ....	65
Figura 4.8 Error RMS de validación. Caso "Global".....	67
Figura 4.9 Comparación de valores de producción. Caso "Global". ....	68
Figura 4.10 Dispersión de los valores de producción. Caso "Global". ....	68
Figura 4.11 Comparación de valores de TIE producto. Caso "Global". ....	69
Figura 4.12 Dispersión de los valores de TIE producto. Caso "Global". ....	69
Figura 4.13 Comparación de valores de TFE producto. Caso "Global". ....	70
Figura 4.14 Dispersión de los valores de TFE producto. Caso "Global". ....	70
Figura 4.15 Comparación de valores de peso específico producto. Caso "Global". ....	71
Figura 4.16 Dispersión de los valores de peso específico producto. Caso "Global". ....	71
Figura 4.17 Comparación de valores de azufre producto. Caso "Global".....	72
Figura 4.18 Dispersión de datos de azufre producto. Caso "Global". ....	72
Figura 4.19 Comparación de valores de $\Delta T$ . Caso "Global". ....	73
Figura 4.20 Dispersión de datos de $\Delta T$ . Caso "Global". ....	73
Figura 4.21 Comparación de datos de $\Delta P$ . Caso "Global".....	74
Figura 4.22 Dispersión de datos de $\Delta P$ . Caso "Global". ....	74
Figura 4.23 Error RMS de validación. Caso "Proceso".....	80
Figura 4.24 Comparación de datos de producción. Caso "Proceso". ....	81
Figura 4.25 Dispersión de datos de producción. Caso "Proceso".....	81
Figura 4.26 Comparación de datos de azufre. Caso "Proceso".....	82
Figura 4.27 Dispersión de datos de azufre. Caso "Proceso". ....	82
Figura 4.28 Comparación de datos de $\Delta T$ . Caso "Proceso". ....	83
Figura 4.29 Dispersión de datos de $\Delta T$ . Caso "Proceso".....	83

Figura 4.30 Comparación de datos de $\Delta P$ . Caso "Proceso".....	84
Figura 4.31 Dispersión de datos de $\Delta P$ . Caso "Proceso".....	84
Figura 4.32 Error RMS de validación. Caso "Propiedades".....	89
Figura 4.33 Comparación de datos TIE producto. Caso "Propiedades".....	90
Figura 4.34 Dispersión de datos de TIE producto. Caso "Propiedades".....	90
Figura 4.35 Comparación de datos TFE producto. Caso "Propiedades".....	91
Figura 4.36 Dispersión de datos de TFE producto. Caso "Propiedades".....	91
Figura 4.37 Comparación de datos peso específico producto. Caso "Propiedades".....	92
Figura 4.38 Dispersión de datos de peso específico producto. Caso "Propiedades".....	92
Figura 4.39 Topología "Proceso" seleccionada.....	94
Figura 4.40 Topología "Propiedades" seleccionada.....	95
Figura 5.1 Comparación de datos de producción. Caso Prueba.....	97
Figura 5.2 Dispersión de datos de producción. Caso Prueba.....	98
Figura 5.3 Comparación de datos de TIE producto. Caso Prueba.....	98
Figura 5.4 Dispersión de datos de TIE producto. Caso Prueba.....	99
Figura 5.5 Comparación de datos de TFE producto. Caso Prueba.....	99
Figura 5.6 Dispersión de datos de TFE producto. Caso Prueba.....	100
Figura 5.7 Comparación de datos de Peso específico producto. Caso Prueba.....	100
Figura 5.8 Dispersión de datos de Peso específico producto. Caso Prueba.....	101
Figura 5.9 Comparación de datos de Azufre en el producto. Caso Prueba.....	101
Figura 5.10 Dispersión de datos de Azufre en el producto. Caso Prueba.....	102
Figura 5.11 Comparación de datos de $\Delta T$ . Caso Prueba.....	102
Figura 5.12 Dispersión de datos de $\Delta T$ . Caso Prueba.....	103
Figura 5.13 Comparación de datos de $\Delta P$ . Caso Prueba.....	103
Figura 5.14 Dispersión de datos de $\Delta P$ . Caso Prueba.....	104



## **NOMENCLATURA**

A	Factor de frecuencia de la expresión de Arrhenius.
$B_j$	Valores de sesgo (bias) de la RNA.
b	Ordenada al origen de la recta de ajuste.
BPD	Barriles por día.
d	Salida deseada de la RNA.
E	Energía de activación de la expresión de Arrhenius, kcal/mol.
F	Factor de carga.
$G^b$	Factor de flujo de gas.
$g(\bullet)$	Función de activación no lineal.
$\Delta H^\circ$	Entalpía de reacción, kcal/mol.
K	Constante de equilibrio de reacción.
$K_p$	Constantes cinéticas del proceso de Hidrodesulfuración.
LHSV	Espacio-Velocidad, $\text{hr}^{-1}$
m	Pendiente de la recta de ajuste.
P	Presión de operación, $\text{kg/cm}^2$ man.
$P^a$	Factor de presión parcial de hidrógeno.
PCA	Análisis de componentes principales (Principal Components Analysis)
$\Delta P$	Caída de presión del lecho catalítico, $\text{kg/cm}^2$ .
R	Constante Universal de los Gases, kcal/mol °K
RMS	Error Mínimo Promedio (Root Mean Squared Error).
RNA	Red Neuronal Artificial.
T	Temperatura de operación, °C
$T_j$	Valor real de las variables de salida de la RNA.
TIE	Temperatura Inicial de Ebullición, °C.
TFE	Temperatura Final de Ebullición, °C.
$W_j$	Factores de ponderación de la RNA.
$X_j$	Variables de entrada de la RNA.
x	Variable independiente de la recta de ajuste.
$Y_j$	Señales de salida de la RNA.
y	Salida calculada de la RNA, variable dependiente.
$z_j$	Activación de una neurona.

### *Superíndices.*

n	Orden de reacción.
p	Número de datos de entrenamiento.

*Subíndices*

- i      Número de nodo de entrada.
- j      Número de capa de la RNA.
- k      Número de capa de la RNA,  $\neq j$ .
- m      Número de nodo de la RNA.

*Letras griegas*

- $\alpha$     Momentum del algoritmo de retro-propagación.
- $\beta$     Velocidad de aprendizaje del algoritmo de retropropagación.
- $\delta$     Gradiente de los errores de la RNA.
- $\eta$     Factor de eficiencia del catalizador.
- $\sigma$    Desviación estándar.

## **RESUMEN**

En este trabajo se construyó el modelo de un Reactor de Hidrodesulfuración de Naftas usando Redes Neuronales Artificiales (RNA).

Los datos utilizados en el desarrollo de este trabajo pertenecen a la Planta Hidrodesulfuradora de Naftas UPH-600 de la Refinería “Fracisco I. Madero”, localizada en Cd. Madero, Tamaulipas. Los juegos de datos de planta se analizaron estadísticamente a fin de eliminar aquellos que presentan una dispersión mayor a dos desviaciones estándar respecto a la media; los datos restantes se dividieron en tres subconjuntos: entrenamiento, validación y prueba.

Se analizaron tres estructuras de la RNA: “Global”, “Proceso” y “Propiedades”. La estructura “Global” incluyó 10 variables de entrada y 7 de salida del proceso. A fin de evaluar el efecto de dividir la estructura “Global” en estructuras más pequeñas se analizaron las topologías “Proceso”, con 7 variables de entrada y 4 de salida; y “Propiedades”, con 3 variables de entrada y 3 de salida. El entrenamiento y la validación se realizó usando el programa comercial QwikNet<sup>®</sup>, utilizando el algoritmo de retropropagación, presentándole 100 juegos de datos de entrenamiento y 63 de validación, y usando el criterio de “validación cruzada” para detener el “aprendizaje”. La mayoría de los procesos de entrenamiento de detuvieron entre las 10,000 y las 25,000 épocas.

El modelo obtenido se evaluó presentándole el conjunto de datos de prueba que la RNA nunca había “visto”; los errores promedio de predicción de las variables de salida están entre 0.6 y 3.0 %, excepto la  $\Delta T$  y la  $\Delta P$  que muestran una desviación del 7.9 y 9.9% respectivamente. Adicionalmente el modelo RNA se comparó contra un modelo cinético obteniéndose resultados similares en la predicción del contenido de azufre en la nafta producto, y siendo superior en la predicción de la  $\Delta T$  y la  $\Delta P$ .

Como parte de este trabajo también se incluye el programa “NeuroNaf” el cual hace uso del modelo de RNA desarrollado para simular el comportamiento del Reactor de Hidrodesulfuración de Naftas analizado.



# **ABSTRACT**

This thesis demonstrates how Artificial Neural Networks (ANN) can be used to model a Naphtha Hydrodesulfuration Reactor.

The data used to get the ANN model belong to Naphtha Hydrodesulfuration Plant UPH-600, at “Francisco I. Madero” Refinery, located at Cd. Madero, Tamaulipas, Mexico. The plant data were statistically analyzed in order to eliminate those data showing a dispersion greater than two standard deviations with respect to the mean; the remaining data were divided into three sets: training, validation and test.

Three ANN topologies were analyzed: “Global”, “Process” and “Properties”. The global topology included 10 input variables and 7 output variables. In order to evaluate the effect of dividing the “Global” topology into smaller topologies, two additional structures were analyzed: “Process” topology considering 7 input variables and 4 output variables; and “Properties” topology considering 3 input variables and 3 output variables. The training and validation procedures were made using QwikNet<sup>®</sup>, it uses the Retropropagation algorithm, using 100 data points on the training set, and 63 data points on the validation set; the “cross-validation” criteria was used to stop training. Almost all the training experiments spend between 10,000 to 25,000 epochs to converge.

The ANN model was evaluated using the test set that was not included during the training experiments. The average error prediction was 0.6 to 3.0% except for  $\Delta T$  and  $\Delta P$ , which have shown a deviation of 7.9 and 9.0%, respectively. In addition, the ANN model was compared with a kinetic model; giving almost the same results predicting the naphtha product sulfur content, but it is superior predicting the  $\Delta T$  and  $\Delta P$  values.

The “Neuronaf” program is included as part of this thesis, which use the ANN model obtained to simulate the Naphtha Hydrodesulfuration Reactor.



# **INTRODUCCIÓN**

Uno de los procesos que ha venido cobrando mayor importancia dentro de los complejos industriales de Petróleos Mexicanos (PEMEX) es el de Hidrodesulfuración. La presencia de azufre en las diversas corrientes provoca problemas tales como: corrosión de equipo, deterioro en la calidad del producto terminado, envenenamiento de catalizadores, y contaminación ambiental entre otros. (Carrillo, et. al., 1987)

Actualmente una de las principales preocupaciones a nivel mundial es el creciente índice de contaminación en todos los ámbitos de la naturaleza: aire, agua, y tierra; la cual está directamente vinculada con las actividades industriales tendientes a satisfacer las necesidades de la sociedad moderna.

Durante los últimos 40 años, la Ciudad de México ha crecido notablemente, tanto en población como en asentamientos industriales, esto ha traído como consecuencia que se hayan incrementado también los requerimientos de diversos recursos naturales, así como energéticos. Dentro de estos últimos destacan los grandes volúmenes de combustibles usados cada día, lo cual ha provocado un considerable aumento de los contaminantes emitidos a la atmósfera como resultado de la combustión de gasolina y diesel. Es pertinente mencionar que actualmente, los principales contaminantes provienen de la combustión de gasolina.

El proceso de hidrodesulfuración ha venido ganando importancia desde que entró en operación la primera Planta Hidrodesulfuradora en la refinería “Gral. Lázaro Cárdenas” ubicada en Minatitlán, Ver. en el año de 1964; hasta la fecha en que PEMEX cuenta con 29 plantas de hidrotratamiento. De estas plantas, 10 están dedicadas al procesamiento de DIESEL-SIN, 11 al procesamiento de gasolina y 8 para destilados intermedios y lubricantes. (Rodríguez, 1994)

Es esta creciente importancia de la Hidrodesulfuración lo que motiva la necesidad de contar con una herramienta de simulación que permita predecir el comportamiento de sus variables. Debido a que el proceso en sí es muy complejo, los modelos tradicionales (Carrillo, et. al., 1987; De la Paz, et. al., 1999) –estadísticos, y matemáticos-, han dado una exactitud relativa, ya que normalmente sólo son aplicables en cierto intervalo de condiciones de operación, y si se requiere conocer algún parámetro fuera de este intervalo es necesario obtener un nuevo modelo.

También se puede mencionar que a la fecha los trabajos desarrollados alrededor de las RNA han hecho uso extensivo de datos provenientes de literatura, de experimentos de

laboratorio o de plantas piloto; y aquellos que han utilizado datos de planta industrial, se han limitado a predecir sólo una variable de salida (Michalopoulos, *et. al.*, 2001)

Actualmente, queda claro lo importante que es el hidrotratamiento y que lo seguirá siendo en el futuro, no sólo por el incremento de la conciencia ecológica, sino por el hecho de que el contenido de azufre en el crudo se ha incrementado lenta pero inevitablemente debido a que cada vez son más pesados y están más contaminados. Sin duda ha pasado el tiempo en que el hidrotratamiento era un proceso secundario, dado que actualmente para asegurar el suministro de hidrógeno en estas unidades se instalan unidades productoras de H<sub>2</sub> del tipo reformación con vapor (Steam reforming), o Pressure Swing Adsorption (PSA).

Adicionalmente a lo anterior, una de las principales preocupaciones del IMP ha sido evaluar continuamente el desempeño de sus catalizadores. Por tal motivo, y de acuerdo a las herramientas matemáticas y computacionales disponibles en su momento, se han realizado algunos trabajos encaminados a este fin, dentro de los cuales podemos mencionar los siguientes:

- \* “Modelo de predicción del rendimiento y propiedades de productos en unidades hidrodesulfuradoras catalíticas”. (Carrillo; *et. al.*, 1987)
- \* “Sistema de cálculo para predicción de rendimientos y propiedades de productos en una unidad reductora de viscosidad” (Bravo, Campos; 1984)
- \* “Sistema de cálculo para la estimación de rendimientos y propiedades de productos de las unidades desintegradoras catalíticas tipo fluidas” (Bravo; 1985)
- \* “Modelo de predicción del comportamiento en plantas hidrodesulfuradoras” actualmente en uso en la División de Aplicación Industrial de Catalizadores (DAIC) del IMP.
- \* HIDRO-NAF. Simulador del proceso de Hidrodesulfuración de Naftas (De la Paz, *et. al.*, 1999).

En los últimos años los recursos computacionales han tenido un avance espectacular, pasando de los grandes ordenadores a disposición sólo de los grandes centros educativos y de investigación; al uso de las computadoras personales, accesibles prácticamente a cualquier profesionalista. Esto sólo hablando del desarrollo de equipo, sin embargo los algoritmos, métodos de solución y lenguajes de programación también han evolucionado, siendo en la actualidad más eficientes, y rápidos.

En particular, el área computacional que está resurgiendo a gran escala es la denominada “Inteligencia Artificial”; y dentro de esta, los sistemas expertos, y las RNA, ambos son implementaciones de programas y/o equipos que tratan de emular los procesos cognoscitivos que tienen lugar en el cerebro humano. (Arce, 1994).

Por lo antes expuesto, el presente trabajo propone como objetivo principal la construcción de un modelo que utilice RNA para simular el comportamiento del Reactor de una Planta Hidrodesulfuradora de Naftas, considerando adicionalmente que:

- El proceso de hidrodesulfuración actualmente no cuenta con un modelo matemático que relacione adecuadamente sus parámetros de operación, lo cual provoca, al igual que para otros procesos complejos, que los simuladores convencionales no representen adecuadamente su comportamiento.
- Las redes neuronales no requieren de un modelo matemático para su implementación, sólo datos de las variables de entrada y sus correspondientes datos de salida, creando un “modelo interno no lineal” para representar el proceso.
- Una de las principales características de las RNA, son capaces de representar el comportamiento de procesos que por su complejidad, no se tiene una idea precisa de los fenómenos a través de los cuales se lleva a cabo.
- Las redes neuronales tienen la capacidad de manejar datos incompletos o viciados.

El modelo de RNA propuesto además permitirá conciliar y predecir el comportamiento de las principales variables de operación de un Reactor de Hidrodesulfuración de Naftas, para lo cual en el desarrollo de esta tesis se proponen los siguientes objetivos:

- ✓ Recopilar datos de operación de una Planta Hidrodesulfuradora de Naftas.
- ✓ Establecer el procedimiento para el diseño de la Red Neuronal Artificial óptima para la representación del proceso de Hidrodesulfuración de Naftas.
- ✓ Aplicar el algoritmo de retropropagación para entrenar a la Red Neuronal Artificial (Aprendizaje).
- ✓ Validar la Red Neuronal Artificial “entrenada”.
- ✓ Desarrollar un programa de cómputo que utilice la Red Neuronal construida, como modelo de predicción de las principales variables del proceso.
- ✓ Realizar corridas de prueba con datos ajenos a los conjuntos de datos usados en el entrenamiento y la verificación para establecer la aplicabilidad de la Red desarrollada.



# **1. PROCESO DE HIDRODESULFURACIÓN**

## **1.1. ANTECEDENTES**

Las primeras investigaciones sobre desulfuración de las fracciones del petróleo fueron realizadas en 1885 sobre un catalizador de óxido de cobre, pero no fue sino hasta los años '80 del siglo XX, con la aparición de la reformación catalítica, que los procesos de hidrodesulfuración se hicieron más atractivos, ya que el hidrógeno requerido en este proceso se obtenía a un precio razonablemente barato. En aquel entonces la idea de eliminar el azufre obedecía a la de eliminar un compuesto corrosivo cuya presencia incrementaba los costos de producción, disminuía la susceptibilidad de las gasolinas al plomo, y que al reducir su concentración mejoraba notablemente el color y el olor de los productos (Rodríguez, 1994).

Después de la segunda guerra mundial, con la introducción al mercado de los catalizadores de platino para la reformación catalítica es cuando la hidrodesulfuración de las naftas en particular y el hidrotratamiento en general se hizo indispensable para evitar el envenenamiento de los catalizadores. Por aquel entonces las unidades de hidrotratamiento utilizaban catalizadores a base de óxidos o sulfuros de metales, como el Co, Ni, Fe, Mo, W, soportado sobre bauxita o tierras de Fuller; pero poco a poco se fueron desarrollando los catalizadores que dominaron el proceso por más de 20 años, los catalizadores en base a “molibdato de cobalto” sulfurado sobre gama-alúmina, los cuales tenían una vida larga (varios años de operación) y podían soportar varias regeneraciones.

Es durante esta época cuando las primeras unidades de hidrotratamiento son instaladas en México, más precisamente en la Refinería de Minatitlán, la cual contaba con un paquete compuesto por tres unidades que entraron en operación en 1964: (Rodríguez, 1994)

1. Hidrodesulfuración de gasolina, con una capacidad de 21,000 BPD.
2. Hidrodesulfuración de kerosina, con una capacidad de 15,000 BPD, e
3. Hidrodesulfuración de diesel con una capacidad de 15,000 BPD.

Este fue sólo el principio, ya que antes de 5 años todas las refinerías del sistema PEMEX fueron dotadas con unidades de hidrotratamiento, e incluso a principio de los años '70 la

primera unidad de Hidrodesintegración de Residuos de América Latina entró en operación en la Refinería de Salamanca.

A pesar de que a nivel mundial el número de unidades de Hidrodesulfuración (HDS) se incrementaba, sólo el hidrotreamiento de naftas el cual proporciona la carga a refoamación era considerado como indispensable; el hidrotreamiento del resto de las fracciones del petróleo continuaban siendo concebidas y operadas bajo la idea de eliminar el azufre por corrosivo y para mejorar la apariencia de los productos.

Los datos actuales de las plantas hidrodesulfuradoras en México, se resumen en la tabla 1.1.

Tabla 1.1 Unidades Hidrodesulfuradoras de Petróleos Mexicanos

REFINERÍA	UNIDAD	CARGA	CAPACIDAD BLS/DIA
Minatitlán	HDG	Gasolina	21,000
Minatitlán	HDK	Kerosina	12,000
Minatitlán	HDD	Diesel	17,000
Salamanca	HDS-1	Nafta	8,000
Cd. Madero	UPH-500	Destilados Intermedios	15,000
Cd. Madero	UPH-600	Nafta	18,000
Salamanca	U-4	Lubricantes	10,000
Salamanca	U-7	Diesel-Sin	14,000
Salamanca	U-8	Diesel-Sin	14,000
Tula	U-400-I	Nafta	36,000
Tula	U-700-I	Diesel-Sin	25,000
Tula	U-800-I	Diesel-Sin	25,000
Cd. Madero	UPH-501	Diesel-Sin	25,000
Salamanca	HDS-2	Nafta	25,000
Minatitlán	U-400	Nafta	36,000
Cadereyta	U-400	Nafta	25,000
Cadereyta	U-700	Diesel-Sin	25,000
Cadereyta	U-800	Destilados Intermedios	25,000
Salina Cruz	U-400-I	Nafta	25,000
Salina Cruz	U-700-I	Destilados Intermedios	25,000
Cangrejera	HDS	Nafta Ligera	86,000
Minatitlán	U-100	Diesel-Sin	25,000
Salina Cruz	U-800-I	Diesel-Sin	25,000
Salina Cruz	U-400-II	Nafta	36,000
Salina Cruz	U-700-II	Destilados Intermedios	25,000
Salina Cruz	U-800-II	Diesel-Sin	25,000
Tula	U-400-II	Nafta	36,000
Tula	U-700-II	Diesel Sin	25,000
Tula	U-800-II	Diesel-Sin	25,000

## 1.2. FUNDAMENTOS

La hidrodesulfuración catalítica es el proceso mediante el cual se reduce el contenido de compuestos de azufre presentes en las diferentes fracciones del petróleo, por reacción de estos compuestos con hidrógeno, en presencia de un catalizador a condiciones de operación convenientes. Además de remover el azufre, se tienen importantes remociones de compuestos de nitrógeno, cloro, metales pesados, metales alcalinos y alcalinotérreos. Por otra parte, dado que se efectúan reacciones de hidrogenación, se obtienen substanciales incrementos en la calidad de los productos hidrotratados debido a la disminución en el contenido de compuestos insaturados tales como: aromáticos y olefinas.

El proceso de hidrotratamiento se mantiene en auge debido a las crecientes necesidades de eliminar el azufre y otros contaminantes presentes en diversas fracciones del petróleo, ya que originan problemas tales como:

- Corrosión del equipo de proceso.
- Disminución en la calidad del producto terminado.
- Envenenamiento de los catalizadores del proceso de Reformación Catalítica.
- Contaminación atmosférica, cuando se emplean como combustibles.

En el proceso de hidrodesulfuración, los compuestos orgánicos de azufre se convierten a ácido sulfhídrico e hidrocarburos saturados. Es un hecho que al aumentar el peso molecular de la fracción a tratar, aumenta su punto de ebullición y su contenido de azufre; esto implica que el hidrotratamiento deberá efectuarse a condiciones de operación más severas (Instituto Mexicano del Petróleo, 1994).

Modificando las condiciones de operación del proceso es posible hidrotratar una gran variedad de fracciones líquidas del petróleo, de acuerdo a su contenido de azufre y peso molecular, como se muestra en la tabla 1.2.

Tabla 1.2 Características físico-químicas de varias fracciones del petróleo

TIPO DE CARGA	TIE/TFE* (°C)	Peso específico 20/4 °C	Azufre (% peso)
Gasolina	37/195	0.708/0.730	0.03-0.13
Turbosina	155/270	0.780/0.810	0.2-0.4
Kerosina	160/320	0.790/0.820	0.5-0.8
Gasóleo ligero	180/380	0.840-0.860	0.9-1.6
Aceite cíclico ligero (Diesel)	220/380	0.950-0.960	2.4-2.6
Aceite lubricante	--	--	0.7-0.9
Gasóleo de vacío	240/530	0.890-0.910	2.0-2.3

\*TIE/TFE Temperatura inicial y final de ebullición.

El incremento de la calidad de los productos hidrotratados, se debe principalmente a la remoción de compuestos contaminantes y a la saturación de las moléculas de los compuestos aromáticos y olefínicos que los contienen.

Los contaminantes tales como: ácido sulfhídrico, amoníaco, cloruros de amonio, etc., pueden ser fácilmente separados en las etapas posteriores del proceso; además, el catalizador al actuar como filtro, retiene importantes cantidades de compuestos metálicos, los cuales pueden ser dañinos para los procesos subsecuentes.

Resumiendo, la hidrodesulfuración aplicada al tratamiento de naftas es un proceso de refinación catalítica en el que se emplea una combinación de un catalizador selectivo e hidrógeno, para modificar la estructura de los compuestos de azufre, oxígeno, nitrógeno, y compuestos metálicos, así como para la saturación de compuestos olefínicos.

## 1.3. DESCRIPCIÓN DEL PROCESO

Para la descripción de la Unidad Hidrodesulfuradora de Naftas se consideran las secciones de: Preparación de Carga, Reacción, Estabilización y Fraccionamiento. (Figura 1.1). (Instituto Mexicano del Petróleo, 1994)

Cabe mencionar que esta descripción aplica para aquellas plantas que operan entre 24 y 28 kg/cm<sup>2</sup> man, aun cuando existen plantas que operan a presiones de 40-50 kg/cm<sup>2</sup> man, como es el caso de la unidad UPH-600 de la Refinería de CD. Madero, Tamaulipas.

### SECCIÓN DE PREPARACIÓN DE CARGA.

La alimentación de Nafta proveniente del tanque de almacenamiento y/o de la Planta de Destilación Atmosférica a las condiciones de 35-40 °C y 4-7 kg/cm<sup>2</sup> man, pasa por los filtros de carga a fin de eliminar los sólidos presentes en esta corriente. (Ver figura 1.1)

La nafta llega al Tanque de Carga C1, a las condiciones de 35-40 °C y 25-30 kg/cm<sup>2</sup> man, en donde por una pierna de extracción se elimina el agua que pudiera ser arrastrada junto con la corriente de carga. La carga se envía a precalentamiento en los Intercambiadores de Calor E1 a través de las bombas de carga que opera a una presión de descarga de 35-40 kg/cm<sup>2</sup> man, en los cuales el efluente del Reactor cede parte de su carga térmica, para de esta forma incrementar la temperatura de la carga desde 40-50 °C, hasta 270-285 °C, con lo cual se vaporiza parcialmente, en este punto se une a esta corriente, otra compuesta de gas rico en hidrógeno a 85-90 °C, y 35-40 kg/cm<sup>2</sup> man; esta última corriente está formada por una mezcla de las corriente de hidrógeno de recirculación procedente del Compresor K1 y del hidrógeno fresco de reposición, el cual normalmente es proporcionado por la Planta Reformadora de Naftas a 60-65 °C y 35-40 kg/cm<sup>2</sup> man., la cual se inyecta en la línea de descarga del Compresor K1.

La carga vaporizada en un 60-70%, entra posteriormente al Calentador a Fuego Directo F1 donde recibe la carga térmica necesaria para alcanzar la temperatura de reacción en el Reactor C2, saliendo a una presión de 25-30 kg/cm<sup>2</sup> man.

### SECCIÓN DE REACCIÓN.

La mezcla de gasolina e hidrógeno que sale del Calentador a Fuego Directo F1, entra al Reactor de Hidrodesulfuración C2, donde se llevan a cabo todas las reacciones de hidrotratamiento, a temperaturas que se encuentran en el intervalo de 290-325 °C.

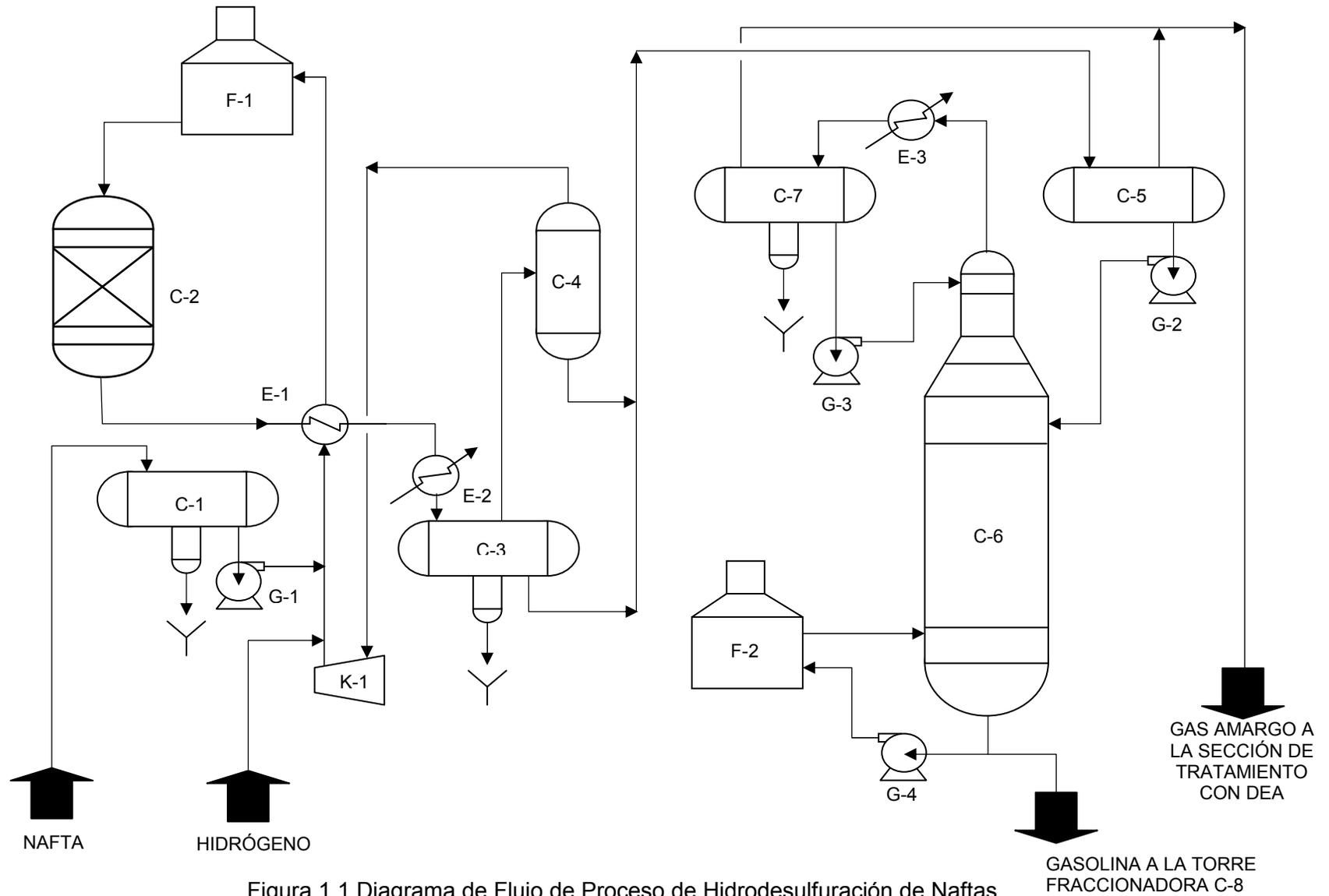


Figura 1.1 Diagrama de Flujo de Proceso de Hidrodesulfuración de Naftas

El Reactor C2, cuenta con un indicador de presión diferencial para conocer la caída de presión a través del lecho catalítico, el cual entre otros factores es una medida de la resistencia al flujo por acumulación excesiva de material extraño en la parte superior de Reactor.

Puesto que las reacciones de hidrodesulfuración son exotérmicas existe un incremento en la temperatura  $\Delta T$  del Reactor, que es función del tipo de carga y de la concentración de contaminantes presentes. Es recomendable alimentar la carga al Reactor a la mínima temperatura a la cual se obtenga el grado de hidrodesulfuración requerido. Esto con la finalidad de que en caso de que los productos no cumplan con la especificación del contenido de azufre, las condiciones de reacción se puedan incrementar.

El efluente del Reactor C2 sale a 305-330 °C, y 25-30 kg/cm<sup>2</sup> man., para posteriormente pasar a los Precalentadores de Carga E1 y por un banco de enfriadores E2, después de los cuales alcanza una temperatura de 40-45 °C, para posteriormente llegar al Tanque Separador de Alta Presión C3. De este tanque la fase líquida ligera se envía a la Torre Desbutanizadora C6, de la Sección de Estabilización. En este mismo recipiente se separa el agua a través de una pierna de extracción. La fase gaseosa rica en hidrógeno, se envía al Tanque de Succión C4, del Compresor de Recirculación de Hidrógeno K1, de donde se recircula a la línea de alimentación de carga a la planta.

## **SECCIÓN DE ESTABILIZACIÓN.**

La finalidad de esta sección es la separación de hidrocarburos ligeros del efluente del Reactor C2, por medio de una Torre Desbutanizadora C6, y una Torre Fraccionadora C8.

La fase líquida procedente del Tanque Separador de Alta Presión C3, a una temperatura de 40-45 °C y una presión de 20-25 kg/cm<sup>2</sup> man, se precalienta por medio de bancos de cambiadores de calor hasta una temperatura de 140-145 °C, para después alimentarse al Tanque Separador de Baja Presión C5, que alimenta a la Torre Desbutanizadora C6. En este recipiente también se reciben los líquidos procedentes del fondo del Tanque de Succión C4, del Compresor K1, y opera a una presión de 5-10 kg/cm<sup>2</sup> man, y a una temperatura de 140-145 °C. Debido a la disminución de presión, se efectúa la separación de hidrocarburos ligeros e hidrógeno en la fase gaseosa, e hidrocarburos pesados en la fase líquida. La fase líquida sirve de alimentación a la Torre Desbutanizadora C6, por medio de la bomba de carga. La fase gaseosa se enfría y condensa en el enfriador de gas amargo y se recibe a 35-40 °C en un segundo separador que opera a 5-10 kg/cm<sup>2</sup> man, y cuya función es recuperar los hidrocarburos de la fase gaseosa del primer separador y dejar libre el gas amargo que se envía a la sección de tratamiento con DEA.

Los hidrocarburos líquidos provenientes de los separadores se alimentan a la Torre Desbutanizadora C6, la cual opera a 90-95 °C y 15-20 kg/cm<sup>2</sup> man. en el domo, y a 235-245 °C y 15-20 kg/cm<sup>2</sup> man. en el fondo.

Los vapores del domo de la Torre C6 se condensan en el Condensador del Domo E3, recibándose en el Acumulador de Reflujo C7, a una temperatura de 50-55 °C y una presión de 15-20 kg/cm<sup>2</sup> man. La presión de este recipiente y por ende del domo de torre, se controla mediante el envío de hidrocarburos ligeros y ácido sulfhídrico (H<sub>2</sub>S) hacia la Sección de Tratamiento con DEA. La fase líquida de este acumulador se divide en dos corrientes: una se utiliza como reflujo a la torre para control de temperatura en el domo. La otra parte se envía previo enfriamiento en el enfriador de destilado líquido de la Torre Desbutanizadora, como producto GLP a tratamiento a la Sección de DEA.

El Tanque acumulador de Reflujo C7, tiene una línea de purga para recuperar el agua amarga, separada de la nafta desulfurada. El drenado de agua de esta línea de purga se realiza de forma manual, y se envía de forma intermitente al cabezal de aguas amargas.

El producto de fondos de la Torre Desbutanizadora C6, se envía por diferencia de presión y a control de nivel, hacia la Torre Fraccionadora C8, en donde se separa el iso-hexano y compuestos más ligeros, de la corriente que será la carga a la Planta Reformadora, la cual debe estar constituida principalmente por hexanos y más pesados.

## 1.4. VARIABLES DE OPERACIÓN

Las condiciones de operación dependen del tipo de carga a tratar, pero debe considerarse como regla general que la severidad de tratamiento debe incrementarse cuando aumenta la temperatura media de destilación de la carga, ya que también se incrementa la concentración de compuestos de azufre difíciles de remover (Ver tabla 1.2).

En general las condiciones menos severas se aplican a las naftas, seguidas por los destilados intermedios y las más severas para las fracciones pesadas.

Las variables de operación de mayor importancia, que influyen considerablemente en el buen funcionamiento tanto del catalizador como de la planta son:

1. Temperatura.
2. Presión.
3. Espacio-Velocidad (LHSV).
4. Relación Hidrógeno/Hidrocarburo(H<sub>2</sub> / HC).

las cuales se detallan a continuación:

---

---

\* Efecto de la temperatura.

La severidad del tratamiento se incrementa directamente con la temperatura, debido al aumento en las velocidades de reacción, incrementando el depósito de carbón sobre la superficie del catalizador y acortando el tiempo de vida del mismo. En las reacciones de hidrogenación el consumo de hidrógeno aumenta, pasando por un máximo que posteriormente declina como consecuencia del inicio de las reacciones de deshidrogenación. (Ver figura 2.2)

La temperatura debe mantenerse lo más baja posible pero manteniendo el nivel de actividad requerido, y así poder mantener la cantidad de carbón al mínimo y retardar la desactivación del catalizador. Sin embargo, la temperatura se debe incrementar progresivamente para compensar la pérdida de actividad por agotamiento del catalizador.

\* Efecto de la presión.

El efecto de la presión está directamente relacionado con el efecto de la composición del gas de recirculación y la relación hidrógeno/hidrocarburo ( $H_2 / HC$ ).

Al aumentar la presión se incrementa hasta cierto grado la remoción de azufre, nitrógeno, oxígeno, la conversión de aromáticos y la saturación de olefinas, además, se tendrá un efecto favorable para la disminución del depósito de carbón en el catalizador gracias a la alta presión parcial de hidrógeno, esencial para llevar a cabo las reacciones de hidrodesulfuración.

\* Efecto del espacio-velocidad (LHSV)

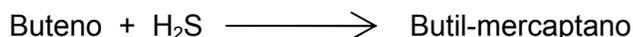
Al disminuir el espacio-velocidad y manteniendo las demás variables constantes, se incrementa el grado de hidrotreatmento hasta un cierto límite, debido al aumento del tiempo de residencia en el reactor. Sin embargo, al operar a un espacio-velocidad bajo se dificulta el control de la temperatura de reacción debido al carácter exotérmico de las reacciones, y se favorece el depósito de carbón.

Al aumentar únicamente el espacio-velocidad trae como consecuencia una disminución en el grado de hidrotreatmento y por lo tanto, en la intensidad de las reacciones catalíticas, de esta forma el consumo de hidrógeno disminuye. (Ver figura 1.2)

\* Efecto de la relación hidrógeno/hidrocarburo ( $H_2 / HC$ )

Se ha observado que a un mayor valor de relación se obtiene un menor depósito de carbón sobre el catalizador, aumentando con ello la vida de éste, hasta un cierto límite; dado por las condiciones de operación y contaminantes de la carga. La presión parcial de hidrógeno aumenta la velocidad de reacción, sobre todo cuando las cargas contienen compuestos pesados de azufre que incrementan el consumo de hidrógeno. En algunos diseños de reactores de hidrodesulfuración se puede usar la inyección de hidrógeno a lo largo del reactor para compensar el consumo del mismo y mantener la relación  $H_2 / HC$  en un nivel constante, además de usarse como medio de enfriamiento (quench).

La formación de olefinas se incrementa con la temperatura. Por tal motivo, para compuestos cuya especificación de contenido de azufre es de 0.2 ppm o menor, se ha tratado de alcanzar aumentando la severidad del hidrot ratamiento, a fin de disminuir aun más la concentración de azufre en el afluente del reactor, sin embargo el efecto es el contrario, ya que el contenido de azufre en el producto se incrementa, debido al aumento en la concentración de olefinas, las cuales se recombinan con el ácido sulfhídrico del gas de recirculación para formar mercaptanos. Por ejemplo:



En la figura 1.2 se muestra gráficamente el efecto de las variables de operación sobre la remoción de azufre en el proceso de hidrodesulfuración.

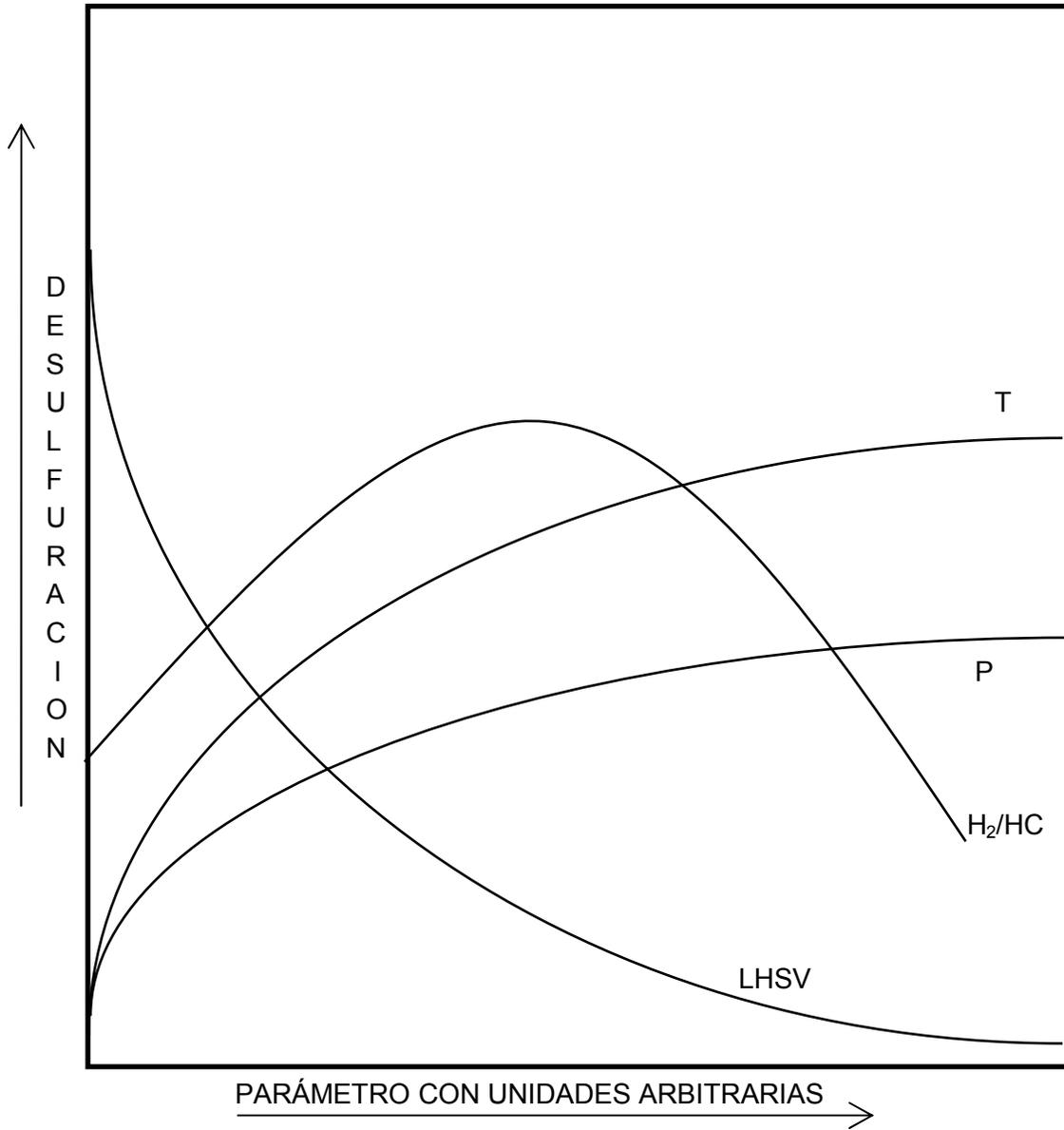


Figura 1.2 Efecto de las variables del proceso de hidrodesulfuración.



## **2. REDES NEURONALES** **ARTIFICIALES**

### **2.1. ANTECEDENTES**

Recientemente las RNA han venido a ser el centro de atención dentro de la Inteligencia Artificial, debido principalmente a su amplio campo de aplicabilidad y su habilidad para manejar problemas complejos y no lineales. La tecnología de las RNA es aplicable en la solución de problemas de la Industria Química, y se espera que tenga un impacto significativo en muchas áreas técnicas y comerciales. (Bhagat, P., 1990)

La aplicación de las RNA, algunas veces denominadas simplemente redes neuronales; permite obtener altas velocidades de procesamiento como resultado de las conexiones de sus elementos en forma masiva y en paralelo. Estas características se deben principalmente a la representación distribuida de los datos dentro de la red, y a su capacidad de adaptación y “aprendizaje”, para mejorar su desempeño. (Lippman, 1987)

Tomando como base estas cualidades, diversas áreas de la Ingeniería Química, tales como el diagnóstico de fallas, control, diseño y simulación de procesos; se han visto beneficiadas, ya que las RNA han permitido simular procesos complejos a partir de datos imprecisos y/o incompletos (Hoskins y Himmeblau, 1988).

La idea de usar las RNA como una herramienta alternativa para solucionar problemas que requieren un análisis complejo de datos, no es nueva. Durante los últimos 40 o 50 años, se ha intentado establecer la estructura neuronal “real” del cerebro, y desarrollar un algoritmo matemático equivalente al proceso de aprendizaje de los seres humanos (Willis, *et.al.*, 1991). La principal motivación detrás de estas investigaciones es emular el alto nivel de procesamiento de información que el cerebro humano es capaz de realizar. Sin embargo, la estructura cerebral es extremadamente compleja, con aproximadamente  $10^{11}$  neuronas y alrededor de  $10^{14}$  -  $10^{15}$  sinapsis (conexiones entre neuronas).

Mientras que el funcionamiento de una sola neurona está relativamente bien entendido, su contribución dentro del conjunto de los elementos cerebrales es menos clara y objeto de varias teorías (Willis, *et.al.*, 1991). Como consecuencia, la topología de las RNA está basada en un conocimiento limitado de las funciones del sistema biológico neuronal.

Aun si la neurofisiología fuera capaz de descifrar las complejidades del cerebro, sería extremadamente difícil, si no es que imposible simular su inmensa estructura distribuida,

debido principalmente a las limitaciones en la tecnología actual de los circuitos computacionales.

Así, en lugar de modelar de manera exacta las intrincadas funciones del cerebro humano, las RNA intentan utilizar la filosofía de interconexión en una escala más modesta y manejable. (Willis, *et.al.*, 1991)

El primer modelo formal de una RNA fue desarrollado por W.S. McCulloch y W. Pitts, de la Universidad de Chicago en 1943, y a partir de entonces surgieron 3 etapas bien definidas en este campo (Palomares, *et.al.*, 1991):

\* Etapa inicial (40's - 50's)

Esta etapa esta caracterizada por un gran entusiasmo y desarrollos significativos como los del GRUPO PERCEPTRON (1958) de la Universidad de Cornell, con un dispositivo de reconocimiento de formas capaz de "aprender" denominado Perceptrón Mark I, y los de B. Widrow y M. Hoff de la Universidad de Stanford diseñando un modelo alternativo al del Perceptrón, llamado red neuronal Adaline, la cual se aplicó con éxito en la predicción del clima.

\* Etapa pasiva (70's - 80's)

Iniciada a partir del desacredito del Perceptrón en 1969, por M.L. Minsky y S. Papert, quienes realizaron un análisis matemático-computacional del Perceptrón, encontrando fuertes limitaciones y poca fiabilidad, y las reacciones negativas de Mortimer Taume, las cuales plasmó en su libro "Computer and Common Sense, The Myth of Thinking Machine", en el cual compara el estado científico de la cibernética con el de la astrología.

\* Etapa actual (80's – a la fecha)

Desde 1991, se han realizado esfuerzos más serios por entender mejor los sistemas biológicos neuronales y su representación con fines computacionales. Se estima que de 1983 a 1988, se tuvo un incremento del 100% en la investigación sobre RNA respecto a la década anterior, y de ahí en adelante se han desarrollado cada día nuevos algoritmos y nuevas aplicaciones de las RNA.

Por otra parte cabe mencionar que Werbos, P.J. (1974) elaboró un algoritmo de adiestramiento para las Redes Neuronales denominado "RETROPROPAGACION", el cual quedó en el olvido por mucho tiempo hasta que fue reconsiderado en la década de los 80's; en la que Rumelhart (1986) reconoció la importancia del algoritmo de retropropagación el cual se ha convertido en uno de los principales métodos de entrenamiento de las RNA. (Palomares, 1991)

A partir de entonces, las RNA se han aplicado en diversos campos de la Ingeniería Química, y dentro de la bibliografía podemos encontrar una amplia gama de ejemplos, algunos de los cuales se mencionan a continuación:

En el control de procesos, podemos mencionar los trabajos de Psaltis, *et.al.* (1988), quienes aplicaron las redes neuronales para controlar una “planta” que transforma coordenadas polares a cilíndricas. Mientras que Bath y McAvoy (1999) lo hicieron para modelar la respuesta dinámica de un medidor de pH en un reactor CSTR. Asimismo, encontramos diversos trabajos adicionales, que a continuación se mencionan como referencia: Antsaklis, (1990); Psychogios y Ungar, (1991); Cooper, *et.al.*, (1992); Hoskins y Himmelblau, (1992); Kavchak y Budman, (1999); Hussain y Kershbaum, (2000); Martins y Coelho, (2000); y Kuttisupakorn, *et.al.*, (2001).

Por otra parte, también ha habido aplicaciones orientadas al control de calidad de productos, Sabharwal, *et.al.*, (1997); Jia y Yu, (2001). Al diagnóstico de fallas, Venkatasubramanian y Chan, (1989); Venkatasubramanian, *et.al.*, (1990); Ungar, *et.al.*, (1990). A la predicción de propiedades físico-químicas, como los trabajos de Bodor *et.al.*, (1991), Para la estimación de la solubilidad en agua de diversos compuestos orgánicos; Petersen, *et.al.*, (1994). Para la predicción del equilibrio vapor-líquido, Sablani, S.S. (2001). Otro campo en el cual las redes neuronales han encontrado bastante aceptación es el de la optimización de procesos, como se muestra en las siguientes publicaciones: Epping y Nitters, (1990); Savkovic-Stevanovic, (1993); Savkovic-Stevanovic, (1994); Zhao, *et.al.*, (2000); y Nascimento, *et.al.*, (2000).

Asimismo, también ha habido aplicaciones orientadas a la identificación de sistemas, lo cual debe ser entendido como la deducción de la funcionalidad que relaciona las variables de entrada, con las variables de salida de un proceso. Aquí podemos mencionar como pionero el trabajo de Hoskins y Himmelblau, (1988), y como relevantes los trabajos de Chu *et.al.*, (1990); Palomares *et.al.*, (1991); Pollard, *et.al.*, (1992); Savkovic-Stevanovic, *et.al.* (1993).

Algunos de los trabajos que permiten vislumbrar la posibilidad de aplicación de las Redes Neuronales a la modelación de procesos son aquellos desarrollados por Bath, N.V., *et.al.*, (1990); donde se analizan tres casos: un reactor en estado estacionario, el seguimiento dinámico del pH en un tanque agitado, y la interpretación de datos usando las redes neuronales como un biosensor. En Willis, *et.al.*, (1991); se utilizan para estimar la concentración de biomasa en sistemas de fermentación industrial, así como la composición de la corriente de domos de una torre de destilación. Mientras que Psychogios y Ungar, (1992); las usan para modelar un bio-reactor en modo batch. Lennox, *et.al.*, (1998); las aplicaron para modelar un proceso de vitrificación; Molga, y Cherbanski, (1999); modelaron un sistema reaccionante líquido-líquido, el cual presenta simultáneamente la transferencia de masa y reacción química; Nascimento, *et.al.*, (1999);

desarrollaron un modelo para la polimerización de Nylon 66 usando un reactor extrusor de tornillos gemelos; Molga, *et.al.*, (2000); modelaron sistemas reactivos cuya cinética es compleja de reproducir, en esta caso la oxidación del 2-octanol con ácido nítrico; Gontarski, C.A., *et.al.*, (2000); simularon el comportamiento de una planta de tratamiento de aguas; Choai, *et.al.*, (2000); modelaron una columna de extracción líquido-líquido; Zoretto, *et.al.*, (2000); las aplicaron para describir la dinámica de la producción de cerveza a través de un proceso batch; a Abilov y Zeybek, (2000); les permitió predecir exitosamente la mejor topología de una refinería de petróleo; Fullana, *et.al.*, (2000); simularon el comportamiento de extractores que utilizan fluidos supercríticos; Tholudur, *et.al.*, (2000); modelaron la dinámica compleja del crecimiento celular y la producción de proteínas en procesos biotecnológicos; Sha, *et.al.*, (2001); las aplicaron para la simulación del proceso de cristalización.

Sin embargo, los trabajos más relevantes en relación a la modelación del comportamiento de procesos químicos complejos se inició con el trabajo de McGreavy, *et. al.*, (1994); quienes aprovecharon las características de las redes neuronales para caracterizar el comportamiento y la distribución de productos en el proceso de desintegración catalítica fluidizada, el cual por su propia naturaleza es extremadamente complejo. Los resultados obtenidos demostraron que las redes neuronales pueden ser usadas como herramientas de soporte operacional.

Continuando con estos trabajos, posteriormente Michalopoulos, *et.al.*, (2001), utilizaron las redes neuronales para modelar el comportamiento de una planta FCC en estado estacionario. Los datos utilizados para el desarrollo, entrenamiento y validación de la red neuronal, fueron obtenidos de una planta industrial. El modelo propuesto es capaz de predecir el porcentaje volumétrico de conversión basado en seis variables de entrada. Los resultados obtenidos demuestran que la RNA puede predecir muy cercanamente los datos reales. También es relevante la aplicación de las RNA que realizó Aguilera, *et. al.* (2003), ya que desarrolló un modelo que predice los rendimientos de productos de una planta FCC y los compara con aquellos obtenidos de modelos estadísticos.

La lista de contribuciones de la RNA a la Ingeniería Química es muy extensa y en esta tesis no se pretende presentar una revisión exhaustiva, solo se incluyen los trabajos más relevantes.

## **2.2. DESCRIPCIÓN**

Una de las muchas definiciones que se pueden dar a las RNA, fue proporcionada por Robert Hecht-Nielsen de la siguiente forma (Ripley, 1996):

“.....es un sistema computacional que está constituido por cierto número de elementos simples, altamente interconectados, los cuales procesan información por medio de un estado dinámico de respuesta a los estímulos externos”

La tecnología de las RNA está basada en los estudios realizados sobre el cerebro de los mamíferos, específicamente de la corteza cerebral. Las RNA pretenden emular la forma en que el cerebro humano procesa información.

El entrenamiento de las RNA puede llevarse a cabo a través de dos formas: supervisada y no supervisada (Sarle, 1997). El aprendizaje supervisado significa que la RNA tiene alguna información de referencia durante el entrenamiento, de tal forma que está en posibilidades de saber cual debe ser la respuesta correcta y saber como aplicar sus reglas de aprendizaje para autoajustarse. Esto es similar al proceso de aprendizaje de formas y colores por el niño.

En contraste, en el aprendizaje no supervisado la RNA no tiene conocimiento respecto a las respuestas correctas y no puede saber en forma exacta cual debe ser dicho valor. Considere a manera de ejemplo, como un bebé aprende a enfocar sus ojos. Los recién nacidos no conocen esta habilidad, pero la adquieren poco tiempo después de nacer. En pocos días, el bebé ha aprendido con poca o ninguna ayuda externa a asociar el conjunto de estímulos visuales con las formas y colores. Las RNA no supervisadas operan de forma similar aprendiendo con poca o ninguna información acerca de la respuesta correcta a un patrón de datos de entrada.

Aunque ambas formas de aprendizaje son importantes para diferentes aplicaciones, en el ámbito de la Ingeniería Química el aprendizaje supervisado es el más adecuado, debido a las características de los procesos, en los cuales se puede obtener una gran cantidad de datos de entrada y salida de los mismos.

Las RNA están en posibilidades de aprender relaciones complejas y no lineales, aun cuando la información de entrada sea imprecisa o esté viciada. Las RNA tienen fuertes ventajas de aplicación en las áreas de síntesis y reconocimiento de palabras, reconocimiento de patrones, clasificación de datos, diagnóstico de fallas, estudios de mercado y modelación de procesos. Estas características hacen que las RNA sean de gran ayuda para resolver diversos problemas de la Industria Química.

Una RNA está formada por varios elementos de procesamiento (neuronas) interconectados, como se muestra en la figura 2.1. Cada elemento de procesamiento recibe un cierto número de entradas  $X_i$  a las cuales se les asigna un factor de ponderación -denominado “peso”-  $W_i$ . A partir del total ponderado de las entradas, el elemento de procesamiento calcula una sola señal de salida ( $Y$ ).

Las siguientes cuatro etapas, mostradas en la figura 2.2, se llevan a cabo cuando cada neurona es activada o procesada:

1. Se reciben varias señales de entrada a partir de otras neuronas ( $X_i$ ).
2. Se calcula la suma ponderada de esas señales.
3. La suma calculada se transforma por medio de una función de transferencia- que normalmente, aunque no siempre, se fija al momento que se diseña la RNA-.
4. El resultado de la transformación (señal de salida  $Y$  de la neurona) se envía a otras neuronas.

Estas etapas se repiten para cada una de las capas de la RNA hasta llegar a la capa de salida.

El proceso de aprendizaje implica que de algún modo la neurona cambia su relación entrada/salida en respuesta a los datos de entrada. Debido a que la función de transferencia normalmente está fija, la única forma en que la salida de la neurona puede variar es ajustando sus factores de ponderación. Así las neuronas de la RNA aprenden, ajustando los pesos de los datos de entrada, y el modelo interno de la RNA es asimilado por el conjunto de esos factores. Aunque existen varias configuraciones de RNA, en particular las denominadas “RNA de alimentación progresiva” son ampliamente usadas en la Ingeniería Química (Michaloupoulos, et. al., 2001; Rasamany, et. al., 1995).

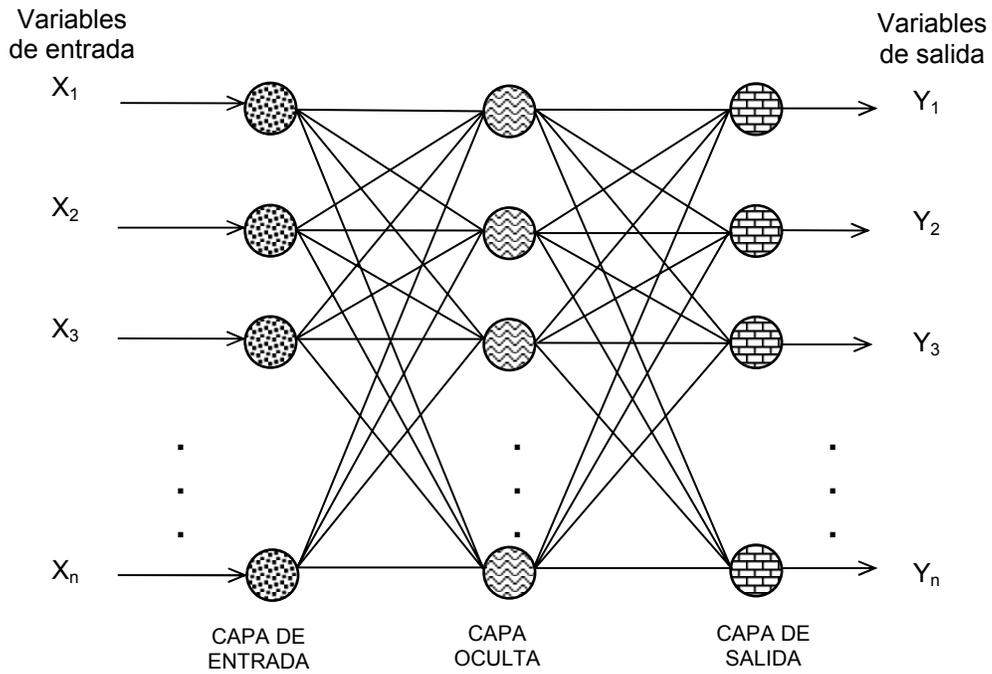


Figura 2.1 Configuración típica de una Red Neuronal Artificial

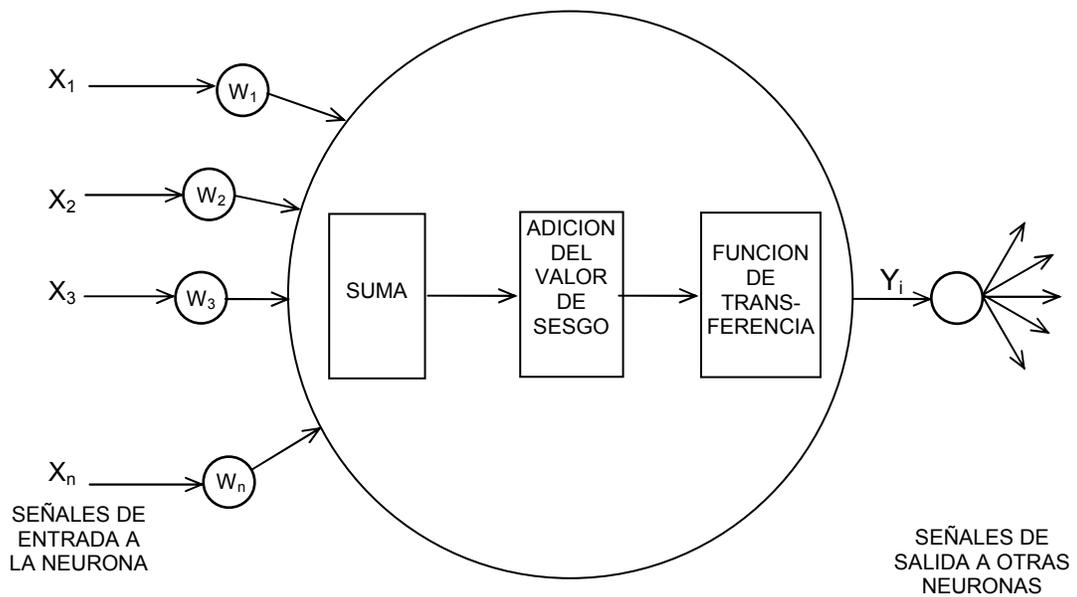


Figura 2.2 Etapas de procesamiento dentro de una neurona.

Las RNA de alimentación progresiva, siempre consisten de por lo menos tres capas de neuronas: una capa de entrada, una capa intermedia (algunas veces denominada capa oculta), y una capa de salida (ver figura 2.1).

La RNA se diseña para que cada elemento de cada capa esté totalmente conectado a cada elemento de la siguiente capa. En otras palabras, cada neurona de la capa de entrada manda su salida a cada una de las neuronas de la capa intermedia, las cuales a su vez mandan su salida a cada neurona de la capa de salida; para el caso de una RNA de tres capas.

El número de neuronas en la capa intermedia puede variar según sea la complejidad del problema de la información que deba procesar. Para un número dado de datos de entrada, si la capa intermedia es muy grande, puede suceder que la RNA no desarrolle un modelo útil. Por otra parte, una capa intermedia con pocos elementos, incrementará drásticamente el número de iteraciones requeridas para lograr la convergencia del entrenamiento de la RNA sin obtenerse una generalización adecuada (Michalopoulos, et. al., 2001).

## 2.3. MÉTODO DE APRENDIZAJE

El proceso de entrenamiento de una RNA, también conocido como algoritmo de retropropagación consiste de dos etapas:

a) Cálculo progresivo.

Como primer paso la RNA de alimentación progresiva, utiliza un conjunto aleatorio de factores de ponderación para actualizar el primer patrón de valores de la red. La capa de entrada recibe los datos y los envía a cada neurona de la capa intermedia. Cada neurona calcula una señal de salida o activación de la siguiente manera: primero se determina la suma de salida  $I_j$ , multiplicando cada señal de entrada por su factor de ponderación aleatorio y sumándole el valor de sesgo, - también llamado polarización-  $B_j$  (ver figura 2.2):

$$I_j = \sum W_{i,j-1} X_{i,j-1} + B_j \quad (3.1)$$

Esta suma ponderada es transformada usando una función de transferencia  $f(X)$  denominada función de activación de la neurona. Ésta determina la actividad generada por la neurona como resultado de una señal de entrada específica.

Para las RNA de retropropagación y en general para los problemas de la Ingeniería Química, la función de transferencia utilizada es la sigmoideal (Chouai, et. al., 2000). Dicha función, como se muestra en la figura No. 2.3, es continua y en forma de "S", se incrementa monotónicamente, y de manera asintótica se aproxima a valores fijos conforme los valores de la suma ponderada se aproxima a  $\pm \infty$ .

Generalmente, el límite superior de la sigmoide se establece como +1 y el límite inferior como 0 o -1. La siguiente función sigmoideal expresada como una función de  $I_j$ , es ampliamente usada:

$$X_{i,j} = f(I_j) = \frac{1}{1 + \exp(-\beta I_j)} \quad (3.2)$$

donde  $I_j$  es el valor de la suma ponderada, y  $X_{i,j}$  es la señal transformada. Este valor se convierte en la activación de las neuronas de la etapa intermedia, la cual es propiamente el valor de salida de dichas neuronas. Siguiendo a través de la RNA, esas salidas son tratadas como las entradas a las neuronas de la capa de salida. Enseguida se calcula la activación para cada una de las neuronas de la capa de salida, usando las señales provenientes de la capa intermedia y la función sigmoideal. Estas activaciones son los valores de salida de la RNA.

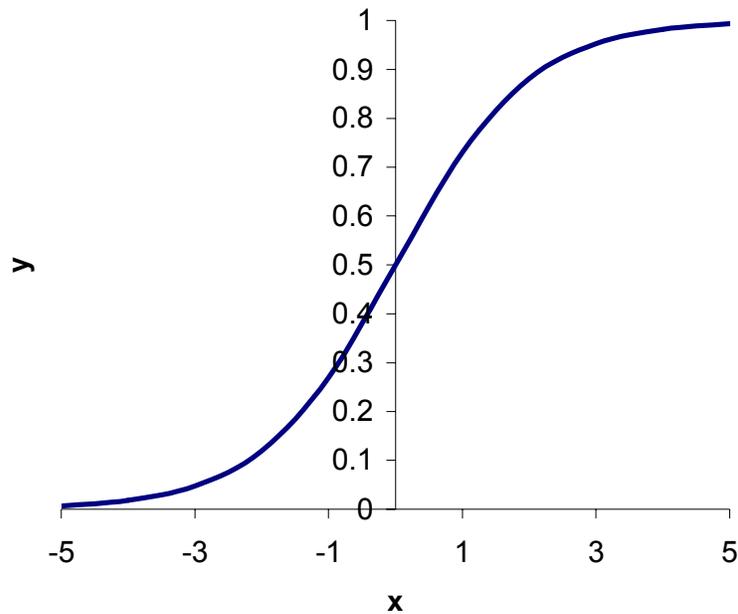


Figura 2.3 Función sigmoideal.

b) Cálculo regresivo.

La salida calculada a partir de la activación progresiva se compara con los valores reales de salida. La diferencia (error) es la base para modificar los factores de ponderación, generalmente para alcanzar los valores deseados son necesarias varias iteraciones. El principio aplicado por el método de retropropagación es la denominada "regla delta generalizada"; dicha regla, minimiza en forma iterativa el cuadrado promedio del error entre los valores de salida de la neurona y los datos conocidos. Esto se lleva a cabo calculando el gradiente del error  $\delta_j$  para cada neurona de la capa de salida:

$$\delta_j = X_j(1 - X_j)(T_j - X_j) \quad (3.3)$$

donde  $T_j$  es el valor correcto de la salida de la neurona  $j$ . El gradiente del error de las neuronas de la capa intermedia entonces se determina calculando la suma ponderada del error de la capa anterior (los detalles del cálculo de los gradientes se dan en el [apéndice A](#)):

$$\delta_j = (1 - X_j) \sum_k \delta_k W_{kj} \quad (3.4)$$

donde  $k$  es la neurona en la capa anterior. Así, el error se propaga regresivamente una capa. Este procedimiento se aplica en forma recursiva hasta la capa de entrada.

Los gradientes de error son usados para actualizar los factores de ponderación de la RNA:

$$\Delta W_{ji}(n) = \beta \delta_j X_i \quad (3.5)$$

y

$$W_{ji}(n+1) = W_{ji}(n) + \Delta W_{ji}(n) \quad (3.6)$$

donde  $(n)$  indica el número de iteración durante el entrenamiento y  $\beta$  es la velocidad de aprendizaje, la cual proporciona el tamaño de paso durante la etapa de disminución del gradiente.

Para acelerar la convergencia generalmente se incluye un término de momentum, el cual determina el efecto del cambio previo de los factores de ponderación, sobre los cambios en el espacio de los pesos.

El cambio de los factores de ponderación después de la  $n$ -ésima iteración es:

$$\Delta W_{ji}(n) = \beta \delta_j X_i + \alpha \Delta W_{ji}(n-1) \quad (3.7)$$

donde  $\alpha$  es el momentum y tiene un valor entre 0 y 1.

Así, el método de retropropagación es un algoritmo de disminución de gradiente que minimiza el promedio cuadrado del error (RMS) de la RNA, disminuyendo el gradiente de la curva de error. En un sistema simple, la curva de error es un paraboloide y la trayectoria de solución de la RNA generalmente tiende a la parte inferior de la curva. Sin embargo, en aplicaciones reales de la Ingeniería Química, la RNA esta formada por un sistema multidimensional, y la curva de error no es cóncava y continua pues presenta varios valles y picos. Como resultado de esta situación, el entrenamiento de una RNA para encontrar el mínimo global se vuelve mucho más complicado.

Los fundamentos teóricos de la capacidad predictiva de las RNA están dados por el teorema de la superposición de Kolmogorov, el cual en términos sencillos establece que (Bishop, 1998): *“cada función continua multivariable (para un dominio cerrado y limitado) puede ser representado como la superposición de un pequeño número de funciones”*. En términos de las RNA este teorema establece que: *“cualquier mapeo continuo  $y(x)$  de ‘d’ variables de entrada  $x_i$  respecto a una variable de salida ‘y’ se puede representar exactamente por una RNA de tres capas, teniendo  $d(2d+1)$  neuronas en la primera capa oculta y  $(2d+1)$  neuronas en la segunda capa oculta.”*

En la demostración que sigue, se usa la siguiente notación (Ayyagari, 2003):

$\mathfrak{R}$  = Línea real.

$\mathfrak{R}^n$  = Espacio Euclidiano n-dimensional.

$C(X)$  = Conjunto de funciones reales y continuas definidas para el conjunto X.

$N(X)$  = Conjunto de funciones de finidas para X que pueden ser expresadas como una RNA de tres capas.

$I$  = Intervalo cerrado  $[0, 1]$ .

$I^n$  = Hipercubo unitario cerrado en  $\mathfrak{R}^n$ .

Debido a que una RNA con más de una salida se puede caracterizar como varias RNA con una salida cada una, es suficiente enfocar nuestro análisis en RNA con una sola salida. Una RNA con una sola salida, donde cada nodo tiene una función de activación, matemáticamente se puede caracterizar como la función:

$$f(x) = \sigma \left( a + \sum_{i=1}^q b_i \sigma(c_i \cdot x + d_i) \right) \quad (3.8)$$

donde:

$\sigma : \mathfrak{R} \rightarrow I$	es la función de activación (normalmente sigmoideal) de los nodos.
$a$	es el umbral del nodo de salida.
$b_i$	es el peso de la conexiones del i-ésimo nodo oculto al nodo de salida.
$q$	es el número de nodos ocultos.
$c_i$	es el vector n-dimensional de los pesos de las n entradas al i-ésimo nodo oculto.
$x$	es el vector n-dimensional de entrada, y
$d_i$	es el umbral del i-ésimo nodo oculto.

Ahora bien  $\sigma$  es típicamente una función (uno a uno) monotónica y la adición de una constante  $a$  sólo representa una translación. Por lo tanto, sin perder generalidad, se puede escribir la ecuación anterior como:

$$f(x) = \sum_{i=1}^q b_i \sigma(c_i \cdot x + d_i) \quad (3.9)$$

Esta ecuación caracteriza una RNA de tres capas definida en  $I^n$ .

El teorema de la superposición de Kolmogorov da respuesta a la pregunta: Dada cualquier función continua  $f : I^n \rightarrow \mathfrak{R}$ , ¿Existe una función continua  $\phi : \mathfrak{R} \rightarrow \mathfrak{R}$  que caracterice a  $f$ ?

Es de gran ayuda analizar que sucede cuando se relaja la pregunta. Es fácil ver que existe una función (posiblemente *discontinua*)  $v : \mathfrak{R} \rightarrow \mathfrak{R}$  que caracteriza a  $f$ . La prueba de esta afirmación es como sigue. Es bien conocido hebreo que los espacios  $I^n$  y  $\mathfrak{R}$  tienen la misma cardinalidad. En otras palabras, existe una relación uno a uno, dentro del mapeo, digamos  $\eta = \mathfrak{R} \rightarrow \mathfrak{R}^n$ , la cual es independiente de  $f$ . Definiendo  $v = \mathfrak{R} \rightarrow \mathfrak{R}$  como:

$$v(x) \equiv f(\eta(x)) \quad (3.10)$$

dando la representación requerida. Por supuesto,  $v$  está lejos de ser continua.

El teorema de la superposición de Kolmogorov establece que  $f$  puede ser caracterizada por  $2n+1$  funciones continuas  $\phi_1, \dots, \phi_{2n+1}$  proporciona una forma específica para la

representación en lugar de solo demostrar su existencia. Kolmogorov demostró que  $f$  se puede representar como:

$$f(x_1, \dots, x_n) = \sum_{i=1}^{2n+1} \phi_i \left( \sum_{p=1}^n \psi_{ip}(x_p) \right) \quad (3.11)$$

donde  $\{\psi_{pq}\}, q=1, \dots, 2n+1, p=1, \dots, n$  son funciones continuas de  $\mathfrak{R}$  a  $\mathfrak{R}$  (llamadas funciones universales ya que sólo dependen de la dimensión  $n$  y no de  $f$ ). Las  $2n+1$  funciones continuas  $\phi_q$  en conjunto caracterizan  $f$ .

Haciendo una comparación entre las ecuaciones 3.8 y 3.11 se puede ver la correspondencia entre sus variables: Se puede establecer que  $q = 2n+1$ ,  $\phi_i(z) = b_i \sigma(z)$ , y  $\psi_{ip}(z) = c_{ip}z + d_i$ , donde  $c_{ip}$  es la  $p$ -ésima coordenada de  $c_i$ .

## 2.4. APLICACIONES EN LA INGENIERÍA QUÍMICA.

Existe una gran variedad de áreas donde se pueden aplicar las RNA en forma conveniente en la Industria Química. Un área de aplicación es la determinación de la composición química, en la cual los datos obtenidos del análisis por espectrometría infrarroja, cuasi-infrarroja o de resonancia magnética nuclear, se utilizan como datos de entrada para identificar los componentes y poder predecir las propiedades de una mezcla. (Hecht-Nielsen, 1990)

Usando las RNA con los datos del monitoreo de las variables de operación de una planta de proceso, puede realizarse un mejor control de la calidad. Debido a que la RNA es actualizada continuamente con nuevos datos, a fin de incrementar su conocimiento del proceso, sus datos de salida pueden ser usados por el sistema de control de la planta para establecer las condiciones de operación, que permitan optimizar el desempeño de esta. (Hecht-Nielsen, 1990)

Las RNA también pueden ser usadas en el diagnóstico de fallas. La continua degradación del funcionamiento de los equipos de proceso debido a su deterioro, así como los disturbios del proceso provocados por cuestiones ambientales, ocasiona que la variable de interés sufra una desviación de su valor normal. Se han usado varias

aproximaciones basadas en las RNA para diagnosticar las causas que originan tales desviaciones. (Venkatasubramanian y Chan; 1989; Ungar, *et.al.*; 1990)

En conclusión, como se ha visto, las RNA pueden apoyar a la Industria Química de muchas formas, desde la investigación básica hasta la investigación de mercados. Sin embargo para usar las RNA de una forma confiable, es necesario que los datos de entrada y salida incluyan todas las condiciones significativas, así como sus características de funcionamiento, de tal forma que el modelo desarrollado internamente por la RNA describa adecuadamente el proceso físico.

A fin de establecer la aplicabilidad de las RNA a los diversos problemas que surgen en la Ingeniería Química a continuación se presentan a detalle tres casos:

-  Simulación de una planta de tratamiento de aguas residuales.
-  Modelación de una columna de extracción líquido-líquido.
-  Modelación de una planta industrial de Craqueo Catalítico Fluidizado.

## **SIMULACIÓN DE UNA PLANTA DE TRATAMIENTO DE AGUAS RESIDUALES.**

El trabajo descrito en este apartado (Gontarski, C.A., *et. al.*, 2000) presenta una forma de predecir las propiedades ambientales de una corriente de salida de una planta de tratamiento de aguas residuales localizada en Rhodiaco Ltda. una de las mayores plantas químicas de Brasil. La planta industrial produce ácido tereftálico y genera aguas residuales que deben ser tratadas en un sistema de lodos activados. Se analizó la influencia de las variables de entrada, y se obtuvieron predicciones satisfactorias.

La modelación y simulación de procesos químicos ha sido desarrollada usando modelos determinísticos cada vez más complejos, debido a la reciente evolución de las computadoras personales. Sin embargo, algunas circunstancias hacen que estos modelos no sean aplicables. Por ejemplo, este es el caso cuando algunos de los datos usados en el modelo son difíciles de obtener o cuando el modelo es extremadamente complejo que requiere una gran cantidad de simplificaciones (Bath, *et. al.*, 1990).

En esta aplicación, se usan RNA para predecir la eliminación del carbón orgánico total (TOC) en la planta de tratamiento, usando el algoritmo delta-bar-delta para la estimación de los pesos y la función sigmoideal como la función de transferencia entre neuronas.

La base de datos original obtenida del sistema de control de planta y del laboratorio, contiene muchos datos inservibles y otros correspondientes a situaciones inestables de la

planta, que hacen necesaria su eliminación. Se utiliza el tiempo de residencia hidráulico promedio en cada equipo para establecer la relación entrada-salida de los datos.

Se usaron siete RNA para simular el sistema, una red para cada reactor y otra para predecir el contenido final de TOC en el efluente, basada en las condiciones de los últimos reactores. El entrenamiento de las RNA usa las siguientes variables de la planta de tratamiento de aguas residuales:

- 🌈 El TOC en el agua residual de entrada a la planta.
- 🌈 La relación de entrada entre el líquido y el lodo reciclado.
- 🌈 La concentración de sólidos suspendidos (lodos) en los reactores.
- 🌈 La concentración del oxígeno disuelto en los reactores.
- 🌈 El tiempo de residencia promedio de los lodos.
- 🌈 Los parámetros relacionados con la cinética de las reacciones, etc.

La salida de la RNA es la predicción de los valores de TOC en las corrientes efluentes de la planta.

El sobre-entrenamiento típico en las RNA de retropropagación se elimina a través del seguimiento del error de predicción de los datos de validación. Cuando este tipo de RNA se entrena con suficientes conexiones (lo cual implica tener el número adecuado de nodos en la capa oculta), el error RMS en los datos de validación decrece al inicio del entrenamiento y entonces alcanza un punto mínimo. En este punto mínimo se considera que la RNA es la adecuada.

Los errores RMS de los resultados obtenidos en esta planta aplicando el modelo de RNA, se muestra en la tabla 2.1.

Tabla 2.1 Predicción de valores de TOC en una planta de tratamiento de aguas residuales.

Equipo	RMS entrenamiento	RMS validación	Índice de correlación entrenamiento	Índice de correlación validación
Reactor 1A	0.0155	0.0265	0.9928	0.9779
Reactor 2A	0.0445	0.0463	0.9709	0.9615
Reactor 3A	0.0244	0.0307	0.9825	0.9724
Reactor 1B	0.0187	0.0229	0.9879	0.9807
Reactor 2B	0.0285	0.0339	0.9825	0.9699
Reactor 3B	0.0309	0.0311	0.9769	0.9812
Tanque final	0.0396	0.0401	0.9730	0.9684

También se desarrollo un análisis de sensibilidad a fin de determinar que variables son las que más afectan al proceso, y se determinó que el flujo de líquido y el pH de la corriente de entrada son las variables más importantes que controlan la planta, siempre y cuando todas las variables restantes se encuentran dentro del intervalo de los datos estudiados.

La principal conclusión es que las RNA pueden ser usadas para establecer mejores condiciones de operación, lo cual se establece a través de variables tales como la relación de división de la corriente de entrada a cada reactor.

También se establece que las RNA representan una posible ayuda en la operación, ya que son capaces de predecir perturbaciones y actuar pro-activamente para minimizar las fluctuaciones de la salida.

## **MODELACIÓN DE UNA COLUMNA DE EXTRACCIÓN LÍQUIDO-LÍQUIDO.**

Se presentó una aplicación de las RNA a la modelación de una Planta Piloto de una columna de extracción líquido-líquido (Chouai, A., et. al., 2000). Este proceso de separación presenta un comportamiento altamente no-lineal y una dinámica que cambia con el tiempo. Normalmente, los modelos de simulación de plantas químicas que describen aspectos de la hidrodinámica y la transferencia de masa son de estado estacionario, o si son dinámicos son muy complejos y necesitan un tiempo de cómputo excesivo para su convergencia. Este trabajo propone que se pueden obtener mejores predicciones usando RNA en lugar de modelos físicos. Los resultados obtenidos muestran la aplicación exitosa de un modelo de RNA a este proceso.

La extracción por solventes es uno de los procesos de separación más importantes dentro de la Ingeniería Química. Los procesos de extracción tienden a ganar aceptación, por ejemplo dentro de la producción de químicos finos, la producción de fármacos, etc. En la industria de producción de papel, la ozonización del aserrín de álamo usada para mejorar su digestibilidad enzimática, produce sustancias solubles en agua especialmente ácidos carboxílicos. La extracción por solventes se utiliza para recuperar esos ácidos de las aguas residuales a fin de limpiar los efluentes del proceso y con esto reducir la contaminación ambiental. Aquí se trata la recuperación de uno de esos ácidos carboxílicos, el ácido oxálico, a partir de soluciones acuosas sintéticas por un proceso de extracción líquido-líquido, usando como solvente selectivo una mezcla de tributil-fosfato (60 % vol.) + dodecano (40 % vol.). Este tratamiento es especialmente ventajoso cuando los efluentes acuosos contienen cantidades considerables de químicos valiosos recuperables.

Los experimentos de la recuperación del ácido oxálico de las aguas residuales se llevaron a cabo en una columna de disco y orificio, continua y agitada. Después de la etapa de extracción, la regeneración del solvente se realiza fácilmente por agotamiento del soluto de la fase orgánica usando una solución acuosa básica. En este estudio, el objetivo es minimizar la concentración del ácido oxálico en el refinado.

En general, el comportamiento de las columnas de extracción, en especial las columnas de pulsos, pueden ser representadas considerando que son un grupo de gotas fluyendo a contracorriente en una fase continua. La eficiencia de transferencia de masa de la columna se puede relacionar con las principales propiedades del grupo de gotas, tales como su diámetro promedio, la distribución del tamaño de gota, y la retención de la fase dispersa.

Las características del conjunto pueden ser analizadas en términos del transporte de las gotas debido a la fuerza de flotación, y al rompimiento y coalescencia debido a la intensidad del mezclado. La mayoría de estos mecanismos se deben a la turbulencia, provocada por la agitación mecánica (por ejemplo la intensidad de la pulsación en este caso). El rompimiento de la gota se puede ver como un mecanismo aleatorio debido a la turbulencia y coalescencia y puede ser considerado como el resultado de contactos aleatorios entre las gotas en un flujo agitado.

Ese mecanismo de distribución espacio-temporal da como resultado la formación de un conjunto de gotas que continuamente se separan y se juntan a lo largo de toda la columna. Estas características dan lugar a una operación cuasi-estacionaria donde la población de gotas exhibe una distribución local probabilística en relación al tamaño de las gotas. Se desarrolló un modelo dinámico que describe el comportamiento de distribución de las gotas a lo largo de toda la columna. Este modelo dinámico es complejo y a la fecha no ha podido ser integrado en línea para ser usado en un esquema de control predictivo.

Este trabajo muestra que varios aspectos de la dinámica de la columna de extracción (transferencia de masa y retención) se pueden representar con suficiente exactitud usando una aproximación simple basada en las RNA

Para obtener el modelo de proceso, antes que nada la RNA se debe entrenar fuera de línea. Se utilizan dos conjuntos de datos, usando validación cruzada para enseñar a la RNA el comportamiento dinámico de la columna de extracción líquido-líquido.

El primer conjunto de datos, llamado conjunto de aprendizaje, el cual consta de 6000 patrones de entrada-salida, se utiliza para calcular el error y para actualizar los factores de ponderación. El segundo conjunto, llamado conjunto de validación, conteniendo 2292

patrones, se usa para determinar los factores de ponderación óptimos para los que se obtiene el mínimo error de predicción.

Las entradas de la RNA están formadas por los valores pasados y presentes del pH y la conductividad, los valores pasados y presentes de los flujos y la intensidad de los pulsos, y por el valor presente de la concentración del ácido oxálico en la corriente de alimentación. Las salidas son los valores futuros del pH en la corriente del refinado y la conductividad en el fondo de la columna. Se consideró que el proceso de extracción líquido-líquido se puede aproximar por un sistema de primer orden. Se probaron algunos valores de entradas anteriores al  $(t-1)$ , pero dan una exactitud similar de la predicción de la RNA.

La topología de la RNA que proporciona la mejor predicción después del proceso de entrenamiento con validación cruzada fue la red formada por nueve nodos en la capa oculta.

El número de épocas fue de 3000 y el error final para los 6000 datos durante la fase de entrenamiento es de 0.36929, mientras que para los 2992 datos de validación en de 0.204567.

La RNA con esta configuración, es capaz de predecir la concentración final de ácido oxálico y el estado hidrodinámico de la columna de extracción líquido-líquido etapa por etapa, proporcionando la información deseada a intervalos sucesivos de 10 s.

Los valores calculados por la RNA están de acuerdo con los valores experimentales.

A fin de confirmar la validez de la metodología usada en el desarrollo del modelo, se presento un experimento usando un 2.0 % peso de ácido oxálico en la corriente de carga. Comparando los valores experimentales del pH en el refinado y la conductividad en el fondo de la columna, con las predicciones de salida de la RNA se observa que estos valores tienen un error menor al  $\pm 5\%$ .

La capacidad predictiva de la RNA es muy buena a pesar de las relativamente grandes y repentinas fluctuaciones del flujo y la intensidad del pulso lo cual afecta el funcionamiento de la columna. Es importante remarcar que el modelo RNA funciona adecuadamente ya que predice los valores de pH y conductividad muy cercanamente a los valores experimentales. Esto es muy importante para la descripción de comportamiento dinámico de una columna de pulsos.

Cerca de las condiciones de inundación; por ejemplo cuando ocurre una inundación cíclica, la cual es un fenómeno inestable, la RNA describe las variaciones repentinas de

los valores promedio de la conductividad, justo debajo del distribuidor donde se localiza la cama de la fase dispersa.

El valor óptimo de la conductividad es de 0.8 mS/cm (mili-Siemens / centímetro) el cual se obtiene al inicio de la inundación. La RNA puede seguir fácilmente cada incremento y decremento de la conductividad durante los experimentos de extracción.

Para demostrar la validez de la RNA obtenida, esta topología se probó con datos experimentales no usados durante el entrenamiento. De acuerdo a los resultados obtenidos, es claro que la RNA proporciona una representación muy cercana de los datos experimentales en un amplio rango de condiciones de operación de la columna de extracción líquido-líquido. Se obtienen valores de error menores al 10%.

Para muchos procesos químicos es difícil obtener un modelo de simulación, sin embargo las RNA pueden “aprender” la dinámica de los procesos a partir de datos de entrada y salida del proceso. Aquí se desarrolló una RNA para calcular el pH de una fase continua de salida y la conductividad del fondo de una columna. Los resultados presentados muestran que la RNA trabaja adecuadamente para una columna de extracción líquido-líquido la cual presenta un comportamiento altamente no-lineal y con parámetros que cambian a través del tiempo.

## **MODELACIÓN DE UNA PLANTA INDUSTRIAL FCC.**

En este trabajo se presentó un modelo basado en RNA usado para determinar el comportamiento en estado estacionario de una planta de Desintegración Catalítica Fluida (FCC) (Michalopoulos, et. al., 2001). Para desarrollar, entrenar y validar el modelo se usaron datos procedentes de una Refinería localizada en Grecia. Debido a la tremenda complejidad del proceso FCC, el desarrollo de un modelo representa un gran reto. El modelo propuesto es capaz de predecir el porcentaje de conversión volumétrica basada en seis variables de entrada del proceso. Este trabajo se enfocó a la determinación de una topología óptima del modelo, a fin de tener buenas propiedades de generalización. Los resultados obtenidos muestran que la RNA es capaz de predecir en forma precisa los datos experimentales. Los errores de predicción para los conjuntos de datos de entrenamiento y validación son similares, lo cual indica la capacidad del modelo para generalizar cuando se le presentan datos que no ha visto. El modelo de RNA también se compara con un modelo estadístico no-lineal existente. La comparación muestra que el modelo de RNA es superior al modelo estadístico.

El proceso FCC es muy importante dentro de la industria de refinación del petróleo, ya que convierte aceites de alto peso molecular en hidrocarburos ligeros. Las unidades

industriales FCC son diseñadas para procesar varios tipos de cargas, incluyendo destilados primarios, residuos atmosféricos y de vacío y gasóleos de vacío. Debido a que las unidades FCC son capaces de convertir grandes cantidades de cargas pesadas en productos ligeros más valiosos, cualquier mejora en el diseño, operación o control por pequeño que sea, da como resultado un incremento en los beneficios económicos de la planta.

La región de operación económicamente atractiva se determina tanto por las características de la carga y del catalizador, como de la distribución de productos requerida. En la práctica, la optimización de un planta FCC se realiza a través de un procedimiento de prueba y error. La desventaja de este procedimiento es que la transición de un estado a otro debe ser gradual y no siempre resulta en éxito, debido a la compleja interacción entre el reactor y el regenerador. Como resultado de este procedimiento se pueden disminuir la producción con la consecuente pérdida económica.

El modelo del proceso se puede utilizar para proponer una trayectoria óptima para un cambio de condiciones de operación de la planta, el cual minimice la pérdida de productos durante el cambio. Los modelos se pueden desarrollar a partir de principios fundamentales, tales como las leyes de la conservación de materia, energía y momentum, y otros principios de la ingeniería química. Tales modelos son capaces de explicar los procesos físicos implícitos y son llamados modelos fenomenológicos. Se han presentado varios modelos fenomenológicos para el proceso FCC. Sin embargo, debido a la complejidad de las unidades industriales del proceso FCC, es muy difícil obtener un modelo fenomenológico adecuado. La complejidad surge de la íntima interacción entre las variables de operación del reactor (Riser) y el regenerador. Además, surge un alto grado de incertidumbre en la cinética de las reacciones de craqueo y la desactivación del catalizador por deposición de carbón en el riser y el proceso de quemado del carbón en el regenerador. Aun si se obtiene un modelo fenomenológico adecuado, este puede ser muy complejo y normalmente requiere de simplificaciones para poder ser resuelto.

Otro método práctico para desarrollar modelos es una aproximación denominada “caja negra”, en la cual los modelos se obtienen exclusivamente a partir de datos experimentales de planta. Tales modelo no proporcionan un conocimiento detallado de los aspectos físicos implícitos, pero proveen una descripción de la relación dinámica entre las variables de entrada y salida del proceso. Un ejemplo de este tipo de herramientas, son los modelos estadísticos obtenidos a través de análisis de regresión. Sin embargo, la mayoría de las aproximaciones estadísticas están basadas en modelos de identificación de sistemas lineales. Desafortunadamente, la mayoría de los procesos de la Industria Química son no-lineales y para esos casos el desempeño de los modelos lineales no es adecuado.

En este trabajo se desarrolla y entrena un modelo basado en las RNA, a partir de datos operacionales de una unidad FCC industrial. El modelo es capaz de predecir el porcentaje de conversión volumétrica para un estado estacionario, a partir de seis variables de entrada. Se puso especial atención en obtener un modelo con buenas propiedades de generalización. Este modelo puede ser una herramienta valiosa para la evaluación de catalizadores y cargas, optimización de las condiciones de operación y también durante el diseño de la unidad FCC.

Los datos de entrenamiento y validación se tomaron de la operación industrial de la Refinería Aspropyrgos perteneciente a la compañía Hellenic Petroleum, S.A. (localizada en Atenas). Los conjuntos de datos se recolectaron una vez por semana durante dos años. Se excluyeron los meses cuando la planta estuvo en paro, obteniéndose 92 observaciones. Una vez de eliminar los datos que estaban fuera de los rangos normales de operación, quedaron 50 datos disponibles, los cuales son representativos de la operación típica de la planta. La salida del modelo es el porcentaje volumétrico de conversión. Las variables de entrada incluyen propiedades de la carga, propiedades del catalizador y condiciones de operación. Las variables del modelo y sus intervalos se muestran en la tabla 2.2.

Tabla 2.2 Variables del modelo del proceso FCC y sus intervalos.

Variable	Intervalo
Conversión, % vol.	70.74-85.18
Peso específico de la carga	0.90-0.92
Contenido de azufre en la carga, % peso	0.22-2.08
Relación catalizador/aceite	4.48-6.33
Área específica del catalizador, $m^2/g^2$	114-154
Flujo de circulación del catalizador, ton/min	14.18-20.36
T de salida del raiser, °C	514.7-518.7

El entrenamiento de la RNA se realizó usando un subconjunto de 40 datos aleatorios que incluyeron los máximos y mínimos cada variable. Los 10 datos restantes se usaron para la validación del modelo.

De acuerdo a los resultados obtenidos se puede observar que el modelo de RNA predice adecuadamente tanto los datos de entrenamiento como los de validación, no solo reproduciendo las trayectorias de las tendencias de las variables de proceso, sino prediciendo valores muy cercanos a la mayoría de los datos experimentales.

También se puede observar que las predicciones de los valores de entrenamiento son mejores que los valores de validación. Esto es de esperarse, ya que los valores de entrenamiento son responsables de la definición de los factores de ponderación del modelo. Así, la RNA “aprende” las características de los datos de entrenamiento mejor

que las de otro tipo de datos. El hecho de que los parámetros de análisis de los datos de entrenamiento sean muy similares a sus correspondientes de validación, significa que el modelo puede generalizar suficientemente bien el conocimiento adquirido de los datos de entrenamiento. La relación entre el error RMS de los datos de validación y entrenamiento es 1.09, el cual puede ser considerado como una mejora significativa sobre los modelos desarrollados con anterioridad (McGreavy, C., et. al., 1994).

## **3. DESARROLLO DEL MODELO**

### **3.1. SELECCIÓN DEL ALGORITMO.**

Desde las primeras Redes Neuronales, en los años 60's, los investigadores se han preguntado continuamente cual es el mejor procedimiento para determinar los parámetros óptimos de una Red Neuronal entrenada a través del algoritmo de retro-propagación.

En 1988, Kung, S.Y., y Hwang, J.N., propusieron el Análisis de Proyección Algebraica para determinar los valores óptimos del algoritmo de retropropagación tales como: la velocidad de aprendizaje y el número de neuronas en la capa oculta. Sin embargo, aun cuando en su trabajo mencionan que su metodología es superior al algoritmo de retro-propagación puro, en el sentido de que el algoritmo auto ajusta los parámetros antes mencionados y de una manera más rápida, en la práctica dicho algoritmo cayó en desuso rápidamente debido a su poca eficiencia.

Por otra parte, aun cuando ha habido un avance muy importante en el campo de la inteligencia artificial, a la fecha se puede afirmar que no existe ningún procedimiento totalmente válido para determinar los parámetros involucrados de estos novedosos paradigmas, excepto que el de prueba y error (Chouai, et. al., 2000; McGreavy, et. al., 1994; Michalopoulos, et. al., 2001).

Por esta razón en el presente trabajo se realizará un análisis exhaustivo de los datos disponibles, así como de los resultados obtenidos, para establecer el mejor modelo que represente el comportamiento de la operación de un Reactor de Hidrodesulfuración de Naftas. En primera instancia se realiza un análisis estadístico de los datos de operación del Proceso de Hidrodesulfuración, el cual se detalla más adelante, logrando con esto tener solamente los datos que representen más adecuadamente el comportamiento de dicho proceso.

Por otra parte, se realizan varias corridas de simulación con varias topologías y diferentes valores de los parámetros del algoritmo de retropropagación a fin de obtener los valores óptimos de los mismos; lo que implica una determinación por prueba y error, a falta de algún método analítico que permita hacer dicha determinación de una forma sistemática.

Para el desarrollo de esta tesis se aplica el procedimiento propuesto por Michalopoulos, et. al. (2001) el cual se resume enseguida, ya que prácticamente es el único trabajo que establece un procedimiento secuencial para desarrollar un modelo de RNA.

- ✚ **Recolección de datos.** A fin de modelar adecuadamente el proceso, los datos deben contener información suficiente sobre un amplio intervalo de operación. Debido a que las plantas químicas tardan varios días en alcanzar el estado estacionario, se recomienda tomar datos durante varios días.
  
- ✚ **Pre-procesamiento de los datos.** Los datos obtenidos directamente de la planta (datos crudos) no se pueden utilizar en estudios de identificación. Se tienen que filtrar para remover “ruido”, y aquellos que estén fuera del intervalo de operación debido a errores de medición. Así se obtiene un conjunto de datos selectos.
  
- ✚ **Selección del modelo.** Se requiere especificar la topología de la red. Dicha especificación incluye el número de capas ocultas, y el número de nodos en cada capa oculta. El número de nodos de entrada y salida está determinado por los datos de entrada y salida de la planta.
  
- ✚ **Entrenamiento y validación.** Durante el entrenamiento supervisado, la red aprende adaptando los pesos en base a los datos de entrenamiento proporcionados en las capas de entrada y salida. Las RNA usadas en este trabajo fueron entrenadas usando el algoritmo de retropropagación, el cual se implementó en el programa denominado QwikNet V2.23. Este simulador de bajo costo permite que las redes neuronales se diseñen y entrenen fácilmente y está disponible en Internet en la dirección (<http://qwiknet.home.comcast.net/>). Los pesos óptimos se determinan en forma iterativa hasta que se obtenga el error mínimo de los valores de las variables de salida respecto a las medidas, lo cual se determina a través del error RMS (Root Mean Squared).

$$RMS = \sqrt{\frac{\sum_p \sum_m (y_m^p - d_m^p)^2}{pm}} \quad (3.1)$$

donde  $p$  representa las muestras del conjunto de entrenamiento,  $d_m^p$  es el valor de salida deseado en el nodo  $m$ -ésimo sobre la muestra  $p$ -ésima y  $y_m^p$  es la salida calculada. Los parámetros tales como la velocidad de aprendizaje y el momentum usados en el algoritmo de retropropagación, así como otros detalles matemáticos se discuten en el [apéndice A](#). Para construir el modelo, los datos disponibles se dividen en dos conjuntos diferentes: entrenamiento y validación. La Red Neuronal se “entrena” usando los datos de entrenamiento. A fin de evitar la memorización de la red, se monitorea el comportamiento de la misma respecto a los datos de validación, después de cada iteración. De esta forma el conjunto de validación se

usa para medir la capacidad de generalización del modelo. El entrenamiento se detiene en el punto donde el error RMS del conjunto de validación empieza a incrementarse. Debido a que el comportamiento del algoritmo de retropropagación depende de los valores iniciales de los pesos, es necesario realizar varias corridas con diferentes pesos iniciales, seleccionados de forma aleatoria. La selección de la topología óptima de la Red Neuronal se realiza a través de un proceso iterativo, en el cual se prueban diferentes topologías (diferente número de capas ocultas, y/o diferente número de nodos en las capas ocultas) en cada iteración. Se considera que la red con el menor error de validación, es la más adecuada para la representación del proceso en cuestión.

## 3.2. DATOS EXPERIMENTALES

Los datos experimentales utilizados en el desarrollo del presente trabajo se obtuvieron de la Planta Hidrodesulfuradora de Naftas UPH-600 de la Refinería “Francisco I. Madero” localizada en Ciudad Madero, Tamaulipas. Se seleccionó dicha planta debido a que es la única Hidrodesulfuradora de Naftas en México que opera a alta presión, y en la que además se presenta un comportamiento errático y no muy estable que la hace adecuada para su análisis a través de un modelado usando RNA.

Para este fin se cuenta con información del Reactor de Hidrodesulfuración de Naftas que comprende las siguientes variables de operación:

- Flujo de carga, BPD.
- Temperatura de operación, °C.
- Presión de operación, kg/cm<sup>2</sup> man.
- Relación Hidrógeno/Hidrocarburo, m<sup>3</sup>/Barril.
- Pureza del Hidrógeno, % mol.
- Consumo de Hidrógeno, m<sup>3</sup>/día.
- Temperatura inicial de ebullición (TIE) de la carga, °C.
- Temperatura final de ebullición (TFE) de la carga, °C.
- Peso específico de la carga.
- Contenido de azufre de la carga, % peso.
- $\Delta T$  del lecho catalítico, °C.
- $\Delta P$  del lecho catalítico, kg/cm<sup>2</sup> man.
- Flujo de producto, BPD.
- TIE del producto, °C.
- TFE del producto, °C.
- Peso específico del producto.
- Contenido de azufre del producto, ppm peso.

las cuales se clasifican como “entradas” y “salidas” de la RNA, como se indica en la tabla 3.1 junto con sus intervalos correspondientes.

Tabla 3.1 Clasificación de los datos experimentales

ENTRADA	SALIDA
Flujo de carga, BPD (17840-25104)	Flujo de producto, BPD (16057-24437)
Temperatura de operación, °C (275-298.2)	TIE producto, °C (58.9-79.7)
Presión de operación, kg/cm <sup>2</sup> (43.2-63.2)	TFE producto, °C (173.3-195.3)
Relación H <sub>2</sub> /HC, m <sup>3</sup> /Bbl (23.58-47.86)	Peso específico producto, (0.725-0.741)
Pureza de H <sub>2</sub> , % mol (75.8-100)	Azufre producto, ppm (0.28-0.47)
Consumo de H <sub>2</sub> , m <sup>3</sup> /día (28658-126142)	ΔT, °C (0-9)
TIE carga, °C (53.7-72.5)	ΔP, kg/cm <sup>2</sup> (0.5-4.6)
TFE carga, °C (174.3-194.7)	
Peso específico de la carga, (0.726-0.738)	
Azufre de la carga, ppm (618.4-1591.2)	

Inicialmente se contaba con 286 juegos de datos correspondientes a un año de operación de la planta. Debido a la confidencialidad de los mismos no se presentan los valores originales, sólo sus gráficas de variación y los datos escalados en unidades arbitrarias requeridos para el desarrollo del modelo de RNA.

A continuación se muestran las gráficas de variación de los parámetros de operación del Reactor de Hidrodesulfuración correspondientes a un año de operación de la planta.

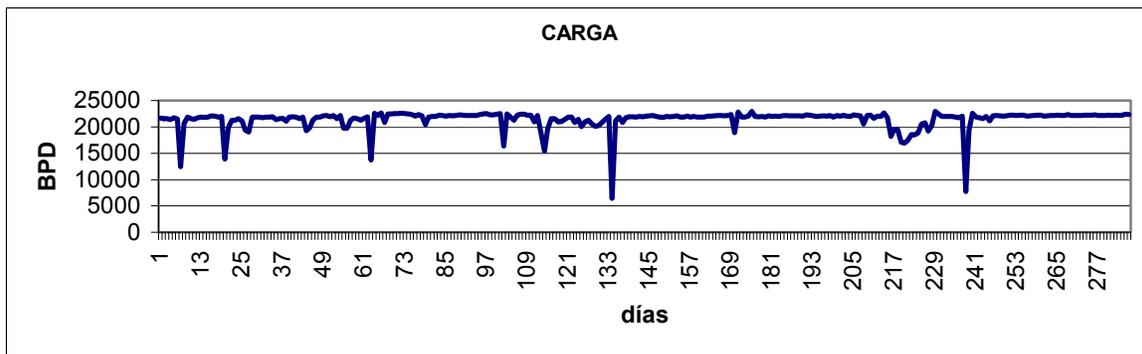


Figura 3.1 Variación de la carga.

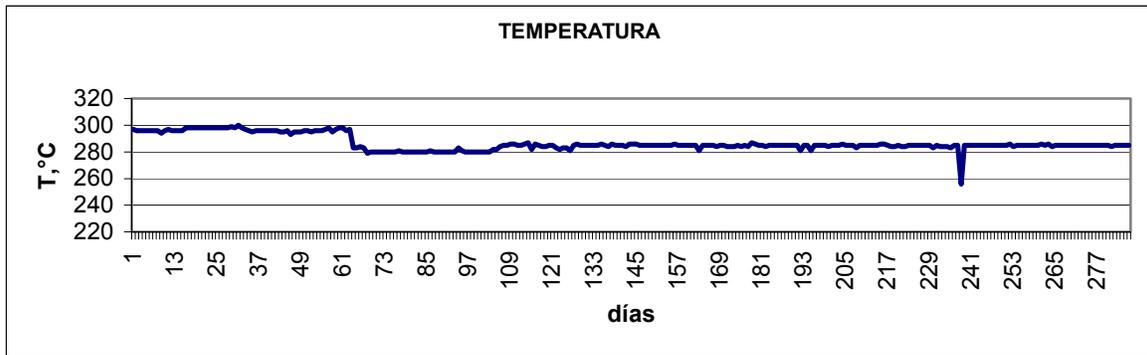


Figura 3.2 Variación de la temperatura de operación.

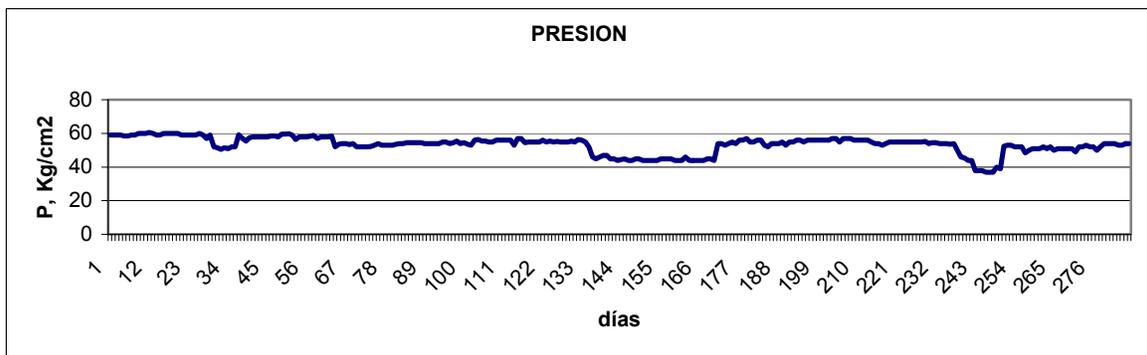


Figura 3.3 Variación de la presión de operación.

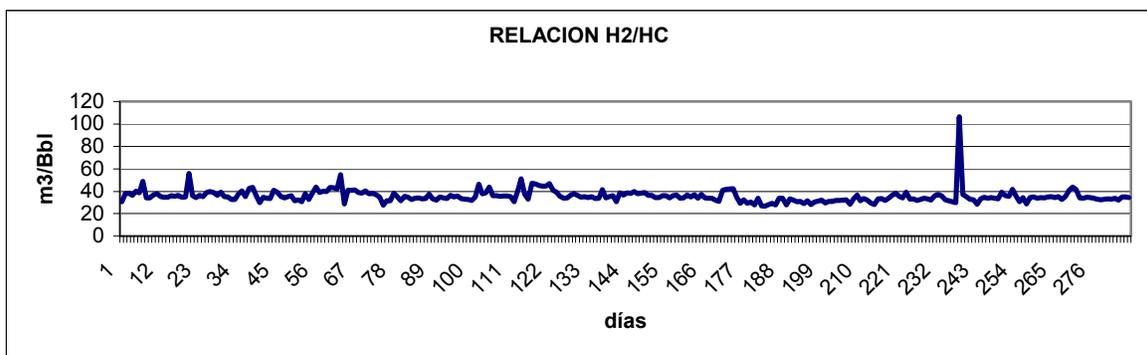


Figura 3.4 Variación de la relación H<sub>2</sub>/HC

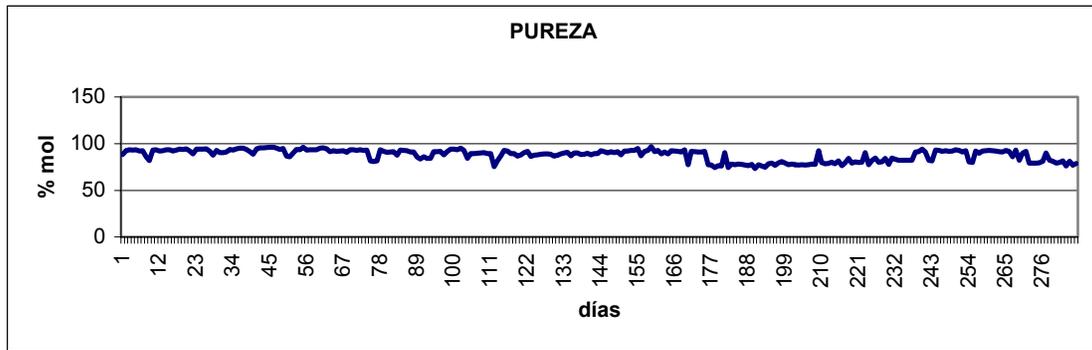


Figura 3.5 Variación de la pureza de H<sub>2</sub>.

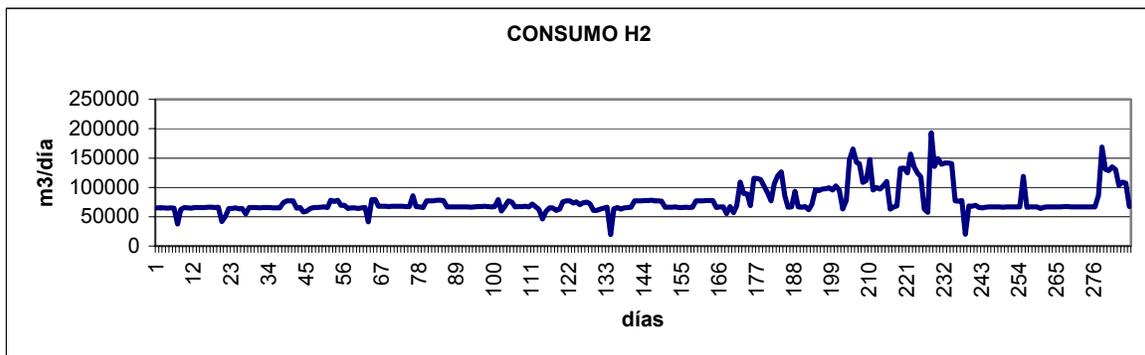


Figura 3.6 Variación del consumo de H<sub>2</sub>.

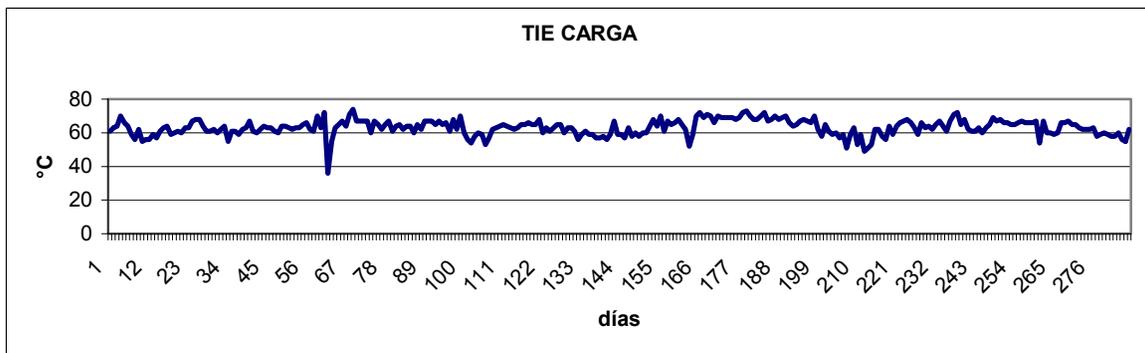


Figura 3.7 Variación de la TIE de carga.

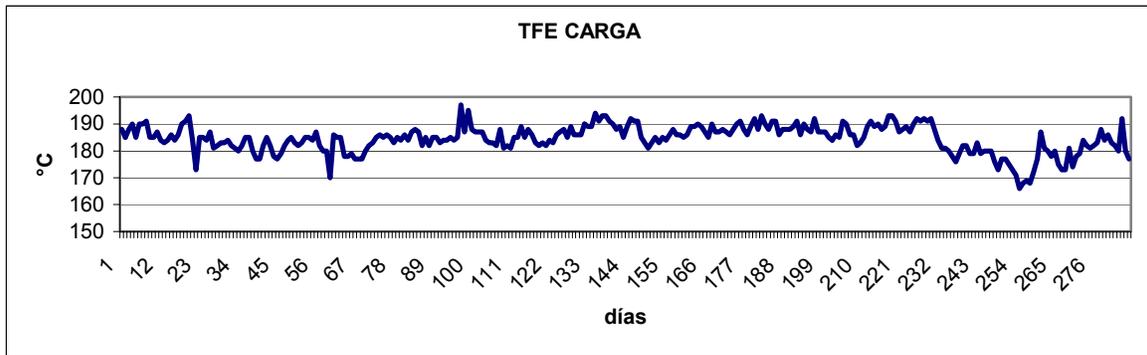


Figura 3.8 Variación de la TFE de carga.

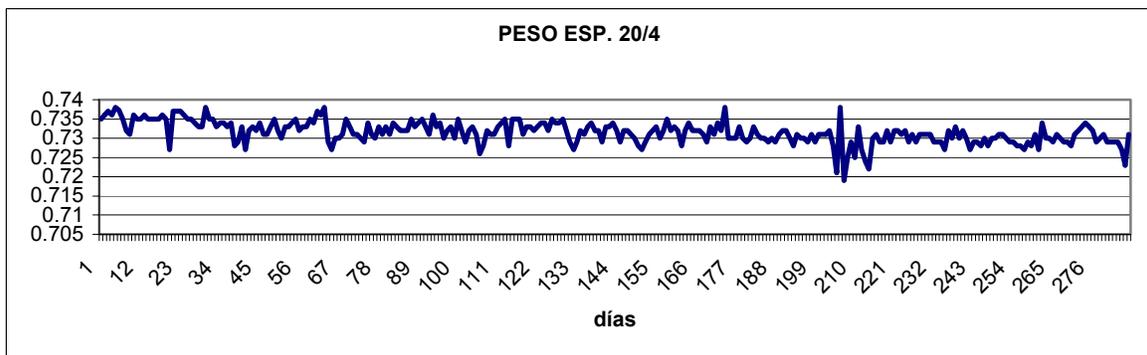


Figura 3.9 Variación del peso específico de carga

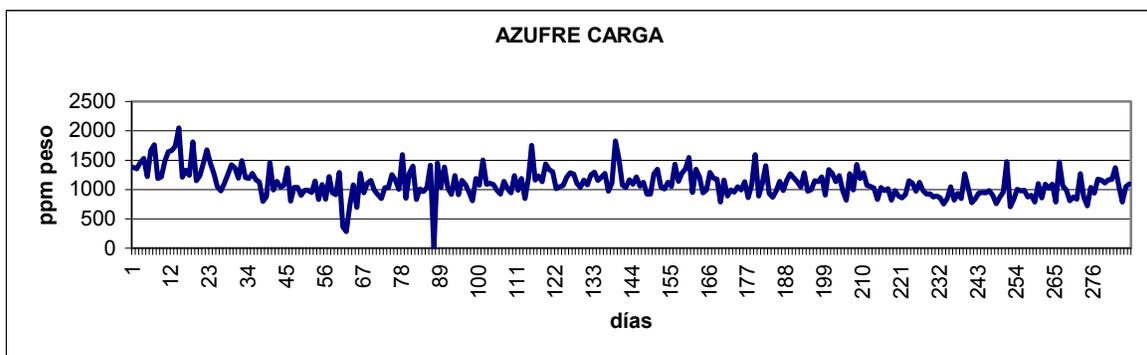


Figura 3.10 Variación del azufre de carga.



Figura 3.11 Variación del flujo de producto



Figura 3.12 Variación de la TIE de producto.

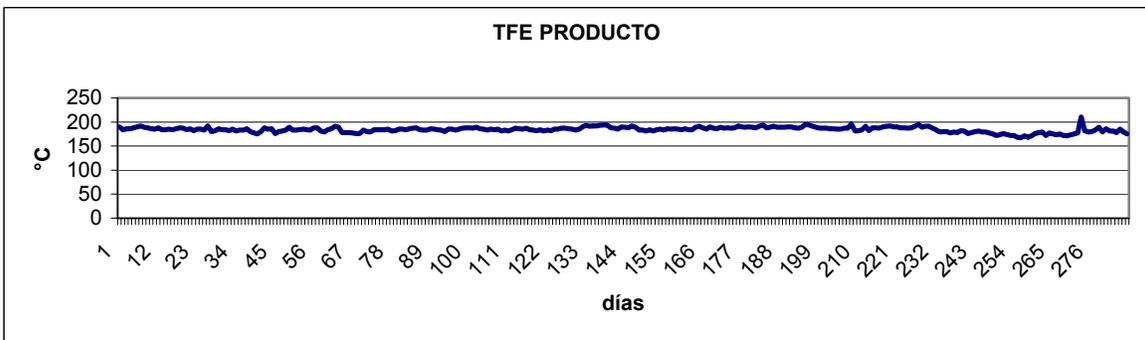


Figura 3.13 Variación de la TFE de producto.

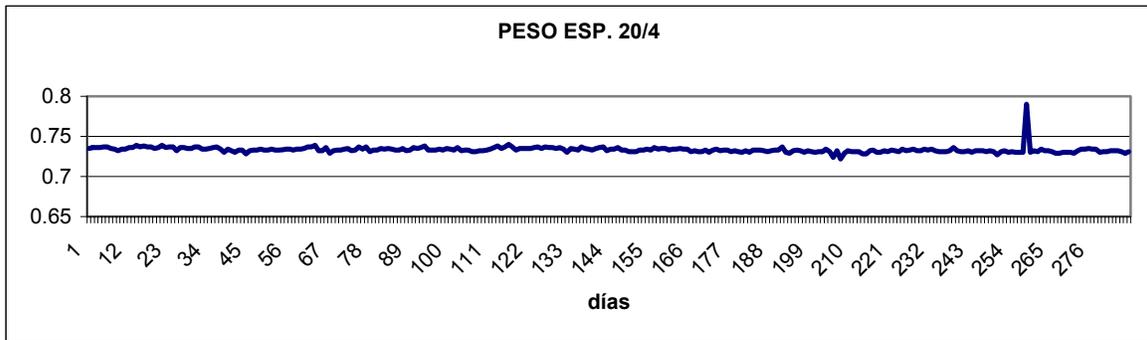


Figura 3.14 Variación del peso específico del producto.

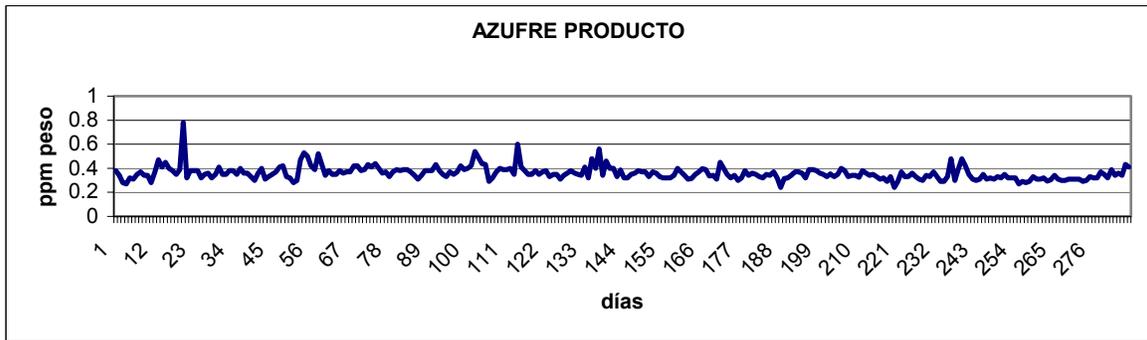


Figura 3.15 Variación de azufre en el producto.

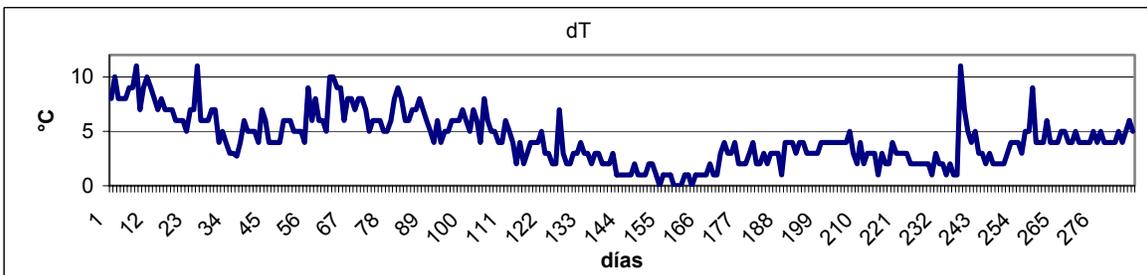


Figura 3.16 Variación de la  $\Delta T$ .

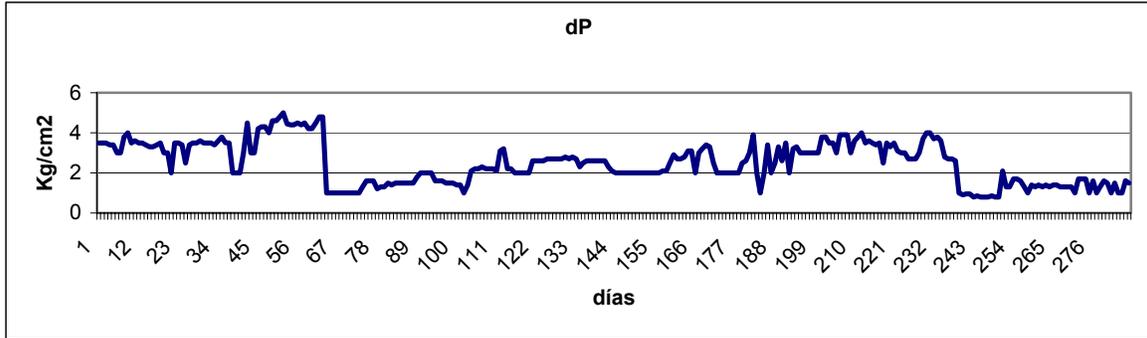


Figura 3.17 Variación de la  $\Delta P$ .

A fin de asegurar la validez de los datos se realizó un análisis estadístico de los mismos eliminando todo juego de datos en el que cualquiera de los valores de sus variables estuviera fuera del intervalo  $\pm 2\sigma$  respecto a la media (Michalopoulos, et. al., 2001); siendo  $\sigma$  la desviación estándar.

Este criterio esta basado en la consideración del intervalo de confianza que se maneja en el análisis estadístico de datos, el cual esta definido como el intervalo en el que cae una medición para una probabilidad dada.

Para el propósito de este trabajo se propone un intervalo de confianza de cada una de las variables involucradas en el desarrollo del modelo del 95.4 %, lo que en términos estadísticos equivale a un intervalo de  $\pm 2\sigma$  respecto a la media. Estas equivalencias están definidas considerando una distribución normal de los datos.

Para la distribución normal, la probabilidad de que una medición caiga dentro de  $n$  desviaciones estándar ( $n\sigma$ ) respecto a la media  $\mu$ , dentro del intervalo  $[\mu - n\sigma, \mu + n\sigma]$  esta dada por (Weisstein, W.E., 2004):

$$P(\mu - n\sigma < x < \mu + n\sigma) \equiv \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu - n\sigma}^{\mu + n\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \quad (3.2)$$

$$\equiv \frac{2}{\sigma\sqrt{2\pi}} \int_{\mu}^{\mu + n\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \quad (3.3)$$

definiendo,  $u \equiv \frac{(x - \mu)}{\sqrt{2\sigma}}$ , de tal forma que  $du = \frac{dx}{\sqrt{2\sigma}}$ , entonces:

$$P(\mu - n\sigma < x < \mu + n\sigma) = \frac{2}{\sigma\sqrt{2\pi}} \sqrt{2}\sigma \int_0^{\frac{n}{\sqrt{2}}} \exp(-u^2) du \quad (3.4)$$

$$= \frac{2}{\sqrt{\pi}} \int_0^{\frac{n}{\sqrt{2}}} \exp(-u^2) du = \operatorname{erf}\left(\frac{n}{\sqrt{2}}\right) \quad (3.5)$$

donde  $\operatorname{erf}(x)$  es la función error. En la tabla 3.2 se resumen los valores de probabilidad  $P(\mu - x_n < x < \mu + x_n)$  de las mediciones de una distribución normal caen dentro del intervalo  $[\mu - x_n, \mu + x_n]$  para  $x_n = n\sigma$  para valores pequeños de  $n$ .

Tabla 3.2 Valores de Intervalos de Confianza

$x_n$	$P(\mu - x_n < x < \mu + x_n)$
$\sigma$	0.6826895
$2\sigma$	0.9544997
$3\sigma$	0.9973002
$4\sigma$	0.9999366
$5\sigma$	0.9999994

En la tabla 3.3 se muestra la media y la desviación estándar resultantes del análisis estadístico de cada una de las variables de proceso de hidrodesulfuración.

Tabla 3.3 Análisis estadístico de los datos experimentales.

VARIABLE	MEDIA	DESVIACIÓN ESTÁNDAR
Carga, BPD	21471.7	1815.9
T operación, °C	286.7	5.81
P operación, kg/cm <sup>2</sup> man	53.2	5.0
Relación H <sub>2</sub> /HC, m <sup>3</sup> /barril	35.7	6.07
Pureza de H <sub>2</sub> , %mol	87.9	6.1
Consumo de H <sub>2</sub> , m <sup>3</sup> /día	77400	24371
TIE de la carga, °C	63.0	4.7
TFE de la carga, °C	184.5	5.1
Peso específico de la carga	0.7315	0.0029
Azufre de la carga, ppm peso	1104.8	243.2
Producción, BPD	20247.2	2095.1
TIE del producto, °C	69.3	5.2
TFE del producto, °C	184.3	5.5
Peso específico del producto	0.7333	0.0041
Azufre del producto, ppm peso	0.3593	0.056
$\Delta T$ , °C	4.3	2.34
$\Delta P$ , kg/cm <sup>2</sup> man	2.5	1.03

De la Tabla 3.3 podemos observar que las variables que más contribuyeron a la eliminación de datos fueron:

- Consumo de hidrógeno,
- $\Delta T$ , y
- $\Delta P$ .

ya que mientras que las variables restantes tienen una desviación estándar que representa un promedio del 15% de la media, las variables en cuestión muestran los siguientes porcentajes: Consumo de hidrógeno, 31.5%;  $\Delta T$ , 54.4%; y  $\Delta P$ , 41.2%. Una vez realizado el análisis estadístico de los datos se eliminaron 86 y quedaron 197 datos disponibles para el desarrollo del modelo de RNA.

Debido a que para el desarrollo de un modelo de redes neuronales –como ya se explicó en el capítulo anterior- se requiere tener tres tipos de datos: entrenamiento, validación y prueba, los 197 datos resultantes del análisis estadístico se distribuyeron en tres conjuntos (Ramasamy, et. al., 1995):

- 100 datos para entrenamiento,
- 60 datos para validación, y
- 37 datos para prueba.

El número de datos considerados para cada conjunto anterior, obedece a las recomendaciones establecidas en trabajos previos en los que se desarrollaron modelos de RNA para procesos químicos (Gontarsky, et. al., 2000; Chouai, et. al., 2000; Ramasamy, et. al., 1995)

Todos los datos fueron normalizados en el intervalo de 0 a 1. Esto se llevó a cabo determinando los valores mínimos y máximos de cada variable y usando la siguiente fórmula:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3.2)$$

Para obtener los valores originales de las variables se utilizó la siguiente fórmula:

$$x = x_{min} + x_{norm} (x_{max} - x_{min}) \quad (3.3)$$

Cabe aclarar que a pesar de que el programa QwikNet<sup>®</sup> realiza la normalización de las variables en forma automática, esta se realizó en forma manual a fin de mantener la secrecía de los datos experimentales usados en el desarrollo del modelo.

### 3.3. DISEÑO DE LA RED NEURONAL.

Aun cuando la inteligencia artificial surgió hace casi 40 años, y el algoritmo de retro-propagación hace aproximadamente 25 años; y a pesar de los considerables avances en el campo de la informática, podemos decir que a la fecha no existe ninguna metodología 100% confiable (Michalopoulos, et. al., 2001) que permita determinar los principales parámetros del algoritmo, tales como: el número óptimo de neuronas en la capa oculta, el intervalo de pesos iniciales, la velocidad de aprendizaje, y el momentum. (Ver [apéndice A](#))

En trabajos recientes como los presentados por McGreavy (1994); y Michalopoulos, (2001); en los cuales ambos analizan el comportamiento de una planta FCC, establecen que la mejor metodología de determinación de los parámetros del algoritmo de retro-propagación es la de prueba y error, tomando como criterio de selección de la mejor topología, aquella que arroje el menor error de predicción para los datos de validación.

Ahora bien, existen también otros criterios que a juicio de McGreavy, (1994), deben ser tomados en cuenta:

1. El número de neuronas en la capa oculta se debe minimizar, ya que entre más neuronas haya en esta capa, el número de datos de entrenamiento requeridos se incrementa considerablemente (Bath y McAvoy, 1992; Morris, et. al., 1994).
2. La exactitud de las predicciones depende de las características del conjunto completo de datos de entrenamiento.
3. La conectividad de las neuronas determina el alcance de la capacidad de predicción de la red neuronal.
4. En la selección de los datos de entrenamiento se debe tener cuidado en escoger datos que representen condiciones de operación normal a fin de evitar predicciones erróneas que puedan surgir de condiciones de operación inusuales (McGregor, 1994). En nuestro caso, esto se aseguró a través de la depuración de los datos usando el análisis estadístico.

Sólo se considerará el caso de una capa oculta, ya que de acuerdo a las diversas aplicaciones a la identificación de procesos, se ha podido establecer que esta topología proporciona resultados adecuados (Michalopoulos, et. al., 2001; Ramasamy, et. al., 1995; Chouai, et. al., 2000).

El número de nodos en las capas de entrada y salida dependerá del caso que se esté analizando de acuerdo a la siguiente descripción:

Lo ideal es contar con una RNA que modele la correlación entre todas las variables de entrada y todas las variables de salida, como se enlistan en la tabla 3.1. Sin embargo, como primer paso se analizará la red que contenga 10 variables en la capa de entrada y 7 variables en la capa de salida, tal y como se muestra en la figura 3.18.

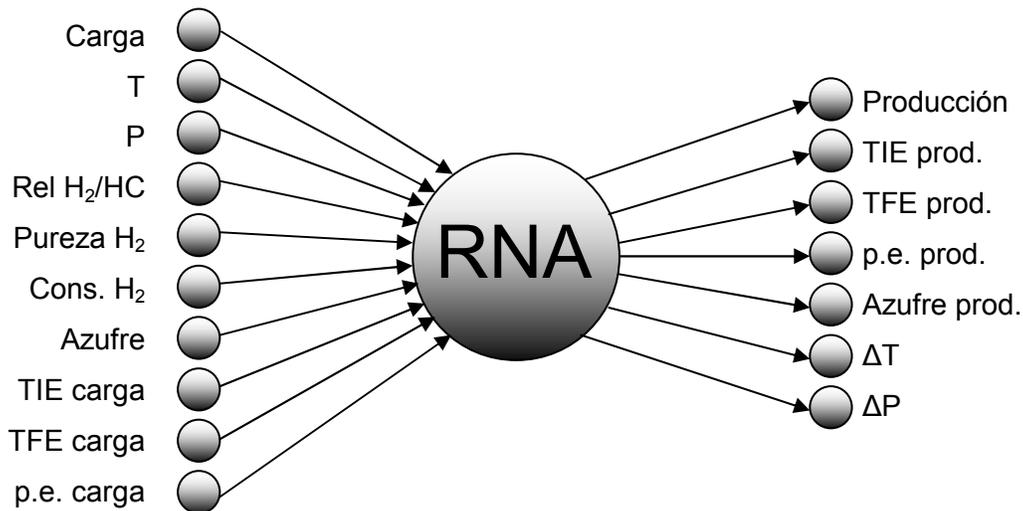


Figura 3.18 Esquema de variables de entrada-salida para el caso 10-X-7.

Cabe aclarar que este es un esquema simplificado con el único propósito de mostrar las variables de entrada y salida consideradas. Las siglas **RNA** en la capa intermedia significan que precisamente esa es la información que se va a determinar como parte de este trabajo.

La mayoría de los modelos RNA desarrollados en la literatura (Michalopoulos, et. al., 2001; Chouai, et. al., 2000; Gontarsky, et. al., 2000; Nascimento, et. al., 1999; Ramasamy, et. al., 1995) han probado una sola topología que incluye todas las variables de proceso, y según lo reportan ha sido suficiente para obtener buenos resultados. En el caso estudiado en esta tesis, la Hidrodesulfuración es un proceso que tiene características hidráulicas, termodinámicas y cinéticas muy complejas, que a la fecha no están bien comprendidas. Esto hace necesario realizar un estudio más profundo a fin de determinar el modelo RNA que mejor represente este proceso. Para este fin también se analizarán varias topologías que surgen de la red total antes mostrada, las cuales se muestran a continuación.

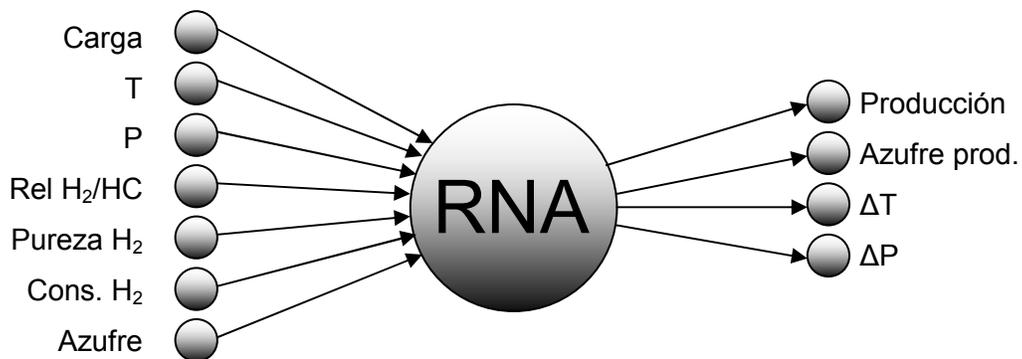


Figura 3.19 Esquema de la topología "Proceso".

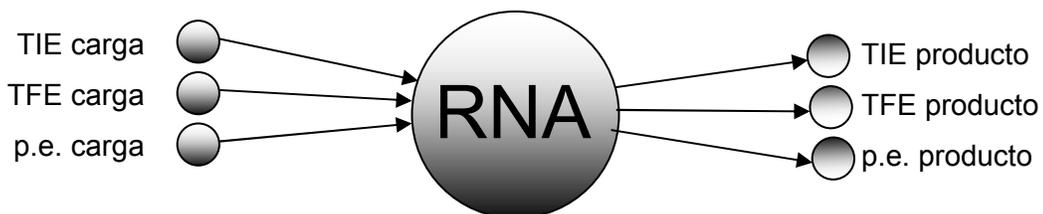


Figura 3.20 Esquema de la topología "Propiedades".

Una vez construidos los modelos correspondientes a las tres topologías y probándolas, se determinará cuales representan mejor el comportamiento de las variables de Reactor, seleccionándolas como el modelo de RNA del proceso.

Para todos los casos el procedimiento para obtener los parámetros óptimos será el siguiente.

Debe quedar claro que el objetivo del procedimiento es determinar las topologías que den el error mínimo en la predicción del conjunto de datos de validación.

Las dimensiones de las capas de entrada y salida se resumen en la siguiente tabla, y serán los casos que se analizan en el presente trabajo.

Tabla 3.4 Número de nodos de entrada y salida de los esquemas propuestos

ESQUEMA	NODOS DE ENTRADA	NODOS DE SALIDA
Global	10	7
Proceso	7	4
Propiedades	3	3

El número de nodos de entrada y salida establecidos en la tabla anterior, corresponden a las variables de entrada y salida en cada topología, de acuerdo a las figuras 3.18–3.20, en las cuales las “X” intermedias en la nomenclatura de la topología se refiere a que no se conoce el número de nodos presentes en la capa oculta de la red; y entonces el objetivo de procedimiento es encontrar el número óptimo de nodos en la capa oculta.

Existen varios métodos para la determinación de los nodos ocultos. Esos métodos se basan ya sea en consideraciones heurísticas (Bishop, 1998), tales como “regularización”, “poda” (pruning), y “paro del entrenamiento”; o estadísticos (Sarle, W.S., 1997) tales como: “prueba de hipótesis”, “criterio de información”, y “validación cruzada”.

Aunque la determinación del número óptimo de nodos ocultos aun está bajo investigación intensiva, parece ser que los métodos estadísticos tienen mayores beneficios que los métodos heurísticos (Anders, y Korn; 1999).

En este trabajo se utilizará un procedimiento de prueba y error basado en el principio estadístico de validación cruzada, cuyo diagrama de flujo se muestra en la figura 3.21 (Michalopoulos, et. al., 2001).

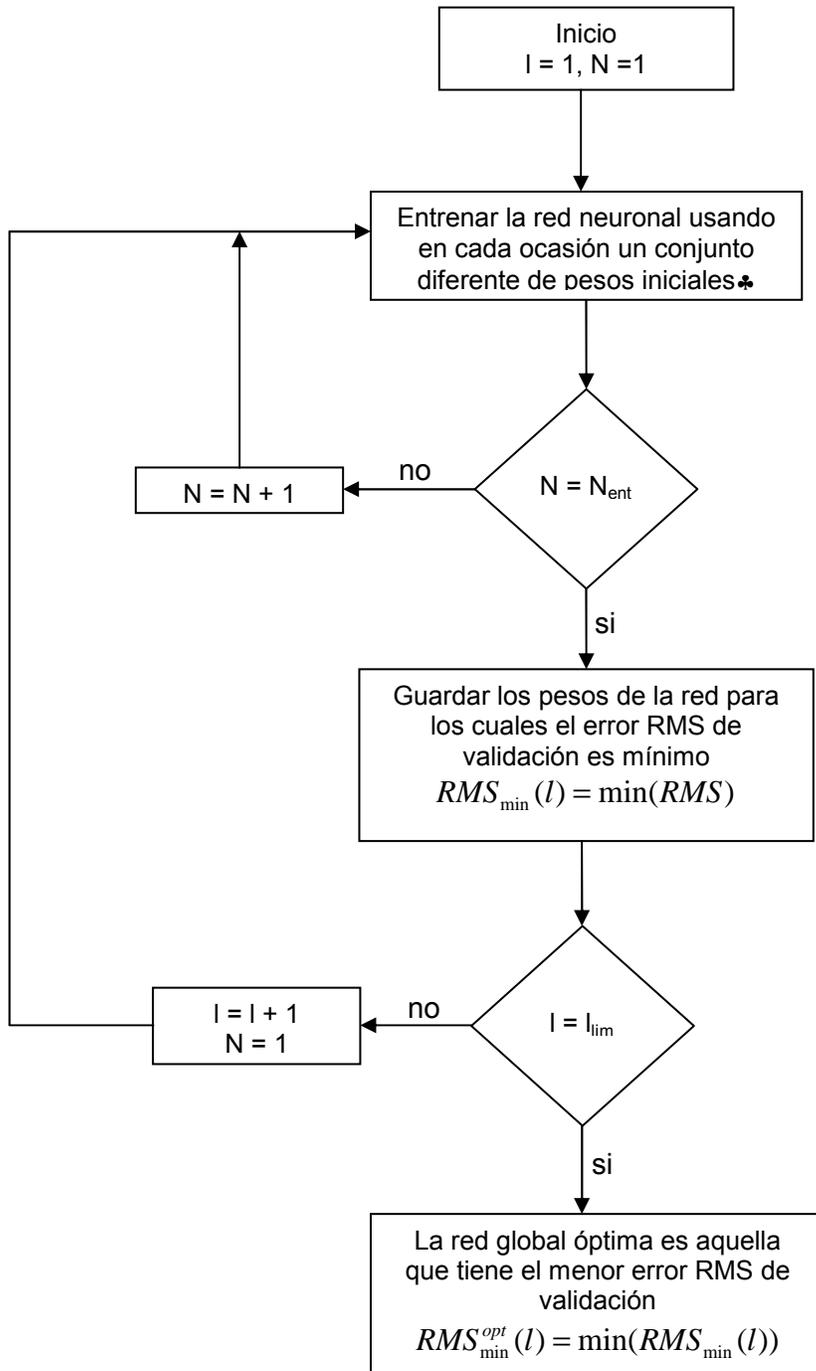


Figura 3.21 Diagrama de flujo de la determinación de una red neuronal óptima.

♣ Experimentando sobre todos los datos de entrada-salida del conjunto de validación.

La descripción del procedimiento se da enseguida:

Primero se establece el valor del número de nodos de la capa intermedia a su valor mínimo ( $l = 1$ ). Entonces, la red se entrena usando el algoritmo de retro-propagación. El procedimiento de entrenamiento se repite varias veces ( $N_{ent}$ ) usando un conjunto de pesos iniciales diferentes cada vez. La red que proporcione el menor error RMS ( $R_{min}$ ) en la predicción de los datos de validación, es la red óptima para el valor particular de  $l$ , y los valores de los pesos correspondientes a esta corrida se guardan. Posteriormente se incrementa  $l$ , y se obtienen los nuevos pesos óptimos para esta nueva topología. El procedimiento es iterativo hasta que  $l$  alcanza el valor de  $l_{lim}$ , esto es el límite superior de número de nodos ocultos.

Cabe aclarar que los pesos iniciales se establecen de manera aleatoria dentro del intervalo propuesto, tal y como se describe en el siguiente capítulo.

Generalmente,  $l_{lim}$  depende de la complejidad del problema a resolver, y en varios modelos desarrollados (Michalopoulos, et. al., 2001; Choauí, et. al., 2000; Gontarsky, et. al., 2000; Nascimento, et. al., 1999; Ramasamy, et. al., 1995; McGreavy, et. al., 1994) se puede observar que la topología seleccionada tiene en la capa intermedia, no más nodos que la capa de entrada, y no menos nodos que la capa de salida; es por eso que en nuestro caso, en el cual la topología global tiene 10 nodos de entrada y 7 nodos de salida, el número de nodos en la capa intermedia debe estar entre estos valores; sin embargo a fin de extender la búsqueda del número óptimo de nodos ocultos más allá de los valores recomendados, en este trabajo se evalúan topologías con hasta 20 nodos en la capa intermedia.

El número óptimo de nodos ocultos corresponde al valor de  $l$ , para el cual la red alcanza el menor  $RMS_{min}$ . Sin embargo, se debe aclarar que debido a que solo se puede realizar un número limitado de corridas para explorar la superficie de error, el conjunto óptimo de pesos obtenido puede representar un mínimo local profundo de la superficie de error, y no el mínimo global (Michalopoulos, et. al., 2001).

Los resultados obtenidos, y la RNA propuesta se presentan en el siguiente capítulo.

## **4. DETERMINACIÓN DE LA TOPOLOGÍA ÓPTIMA**

### **4.1. TOPOLOGÍA GLOBAL.**

En primer término se analizó la topología que incluye todas las variables de entrada y de salida, tal y como se describe en la figura 3.18 y en la cual las variables se clasifican como:

Tabla 4.1 Clasificación de las variables, topología 10-X-7.

ENTRADA	SALIDA
Flujo de carga	Flujo de producto
Temperatura de operación	TIE producto
Presión de operación	TFE producto
Relación H <sub>2</sub> /HC	Peso específico producto
Pureza de H <sub>2</sub>	Azufre producto
Consumo de H <sub>2</sub>	$\Delta T$
TIE carga	$\Delta P$
TFE carga	
Peso específico de la carga	
Azufre de la carga	

Las topologías que se probaron son las siguientes, de acuerdo al criterio establecido en el algoritmo presentado en la figura 3.21:\*

- 10-6-7
- 10-7-7
- 10-8-7
- 10-9-7
- 10-10-7
- 10-12-7

- 10-14-7
- 10-16-7
- 10-18-7, y
- 10-20-7

y para cada una de las topologías se realizaron entrenamientos con los siguientes intervalos de pesos iniciales:

- -1,1
- -2,2
- -4,4
- -6,6
- -8,8
- -10,10
- -15,15
- -20,20

Es importante establecer que en el algoritmo sólo se proponen los intervalos de los pesos iniciales, pero su asignación final se realiza de forma aleatoria pudiendo tomar cualquier valor dentro de dicho intervalo (Chouai, et. al., 2000). La función de los pesos es modificar la señal transportada de un nodo a otro, ya sea aumentando o disminuyendo la influencia de la conexión involucrada (Michalopoulos, et. al., 2001).

El entrenamiento se realiza usando la técnica de validación cruzada, y se detuvo en el momento que se detectó que el error de validación comenzó a incrementarse (Sarle, 1999). La velocidad de aprendizaje del algoritmo de retropropagación se estableció en 0.1, mientras que el momentum se fijó en 0.0 ([ver apéndice A](#)).

Cabe aclarar que estas últimas consideraciones se aplicaron a todas las topologías analizadas.

Una vez realizado el análisis, se obtuvieron los siguientes resultados del error RMS para la fase de entrenamiento. En las tablas 5.2 a 5.11, el asterisco representa el conjunto de pesos para los cuales se obtuvo el menor error promedio.

---

\*La nomenclatura de las topologías de las redes neuronales es la siguiente: por ejemplo la topología 10-6-7, indica que la red neuronal tiene 10 nodos en la capa de entrada, 6 nodos en la capa intermedia u oculta, y 7 nodos en la capa de salida.

---

---

Tabla 4.2 Errores promedio de la topología 10-6-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.125394
-2,2	0.109064
-4,4	0.105547
-6,6	0.103036*
-8,8	0.103913
-10,10	0.104923

Tabla 4.3 Errores promedio de la topología 10-7-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.122573
-2,2	0.103109
-4,4	0.099404
-6,6	0.0965815
-8,8	0.0968514
-10,10	0.0949922*
-12,12	0.0979935

Tabla 4.4 Errores promedio de la topología 10-8-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.118466
-2,2	0.100433
-4,4	0.094780
-6,6	0.091538
-8,8	0.091532
-10,10	0.091364*
-12,12	0.0915701

Tabla 4.5 Errores promedio de la topología 10-9-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.116240
-2,2	0.0977512
-4,4	0.0901805
-6,6	0.0884574
-8,8	0.0855697*
-10,10	0.0880639
-12,12	0.0877055

Tabla 4.6 Errores promedio de la topología 10-10-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.114334
-2,2	0.944434
-4,4	0.852436
-6,6	0.844689
-8,8	0.851554
-10,10	0.833626
-12,12	0.832635
-14,14	0.831143
-16,16	0.0825216
-18,18	0.0820554*
-20,20	0.831068
-22,22	0.0841973

Tabla 4.7 Errores promedio de la topología 10-12-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.109628
-2,2	0.0862991
-4,4	0.0814135
-5,5	0.0768315
-6,6	0.0773471
-8,8	0.0759774*
-10,10	0.0763062
-12,12	0.0771444
-14,14	0.0773744

Tabla 4.8 Errores promedio de la topología 10-14-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.107892
-2,2	0.084043
-4,4	0.0702977
-6,6	0.0699524
-8,8	0.067928
-10,10	0.0686253
-12,12	0.0703367
-14,14	0.0684334*
-16,16	0.0722059

Tabla 4.9 Errores promedio de la topología 10-16-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.104821
-2,2	0.0793449
-4,4	0.0670125
-6,6	0.0639351*
-8,8	0.0646759
-10,10	0.0649743
-12,12	0.0651453
-14,14	0.0647859
-16,16	0.0655935

Tabla 4.10 Errores promedio de la topología 10-18-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.101843
-2,2	0.0732126
-4,4	0.0587184
-6,6	0.0566874*
-8,8	0.0577578
-10,10	0.0611793
-12,12	0.0586645

Tabla 4.11 Errores promedio de la topología 10-20-7, entrenamiento (Caso "Global").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.100187
-2,2	0.068882
-4,4	0.0548013
-6,6	0.0521632*
-8,8	0.0556018
-10,10	0.0559977
-12,12	0.053476

Para cada uno de los procesos de entrenamiento se obtuvieron las gráficas correspondientes a las variables de salida pronosticadas, sin embargo por simplicidad sólo se presentan las gráficas correspondientes a la topología 10-8-7.

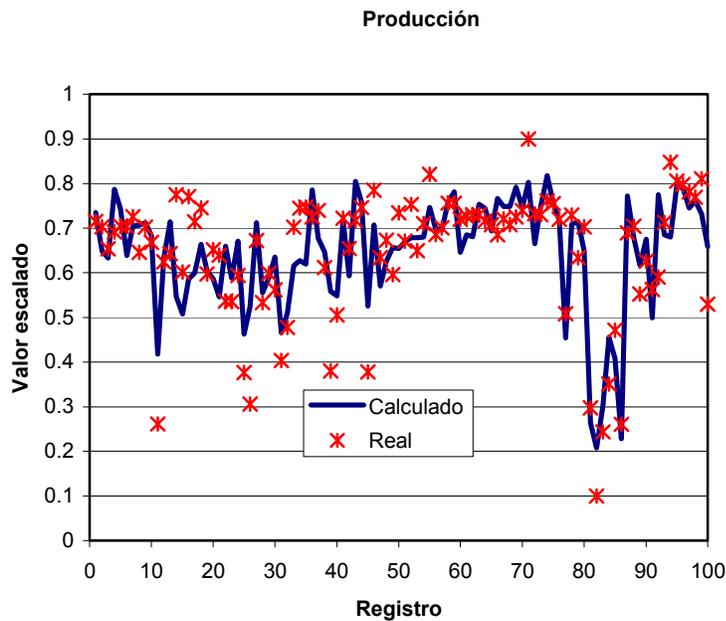


Figura 4.1 Producción, fase de entrenamiento.

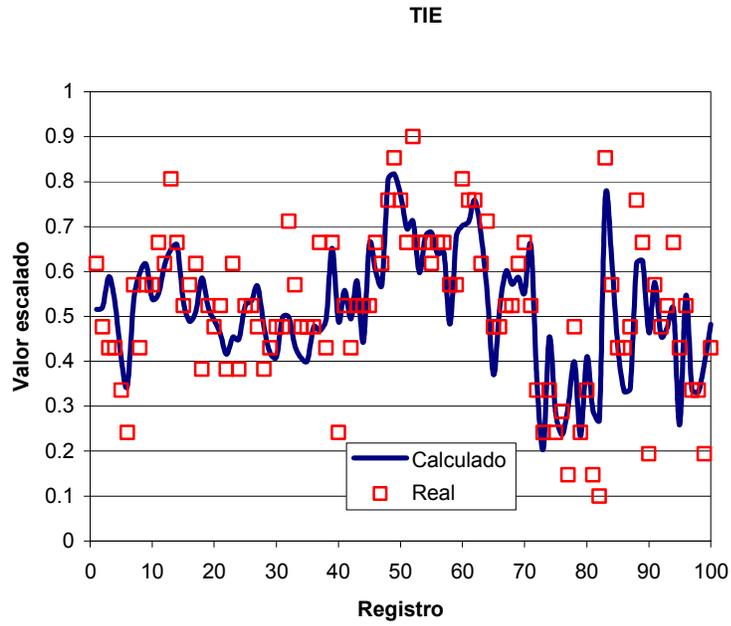


Figura 4.2 TIE del producto, fase de entrenamiento.

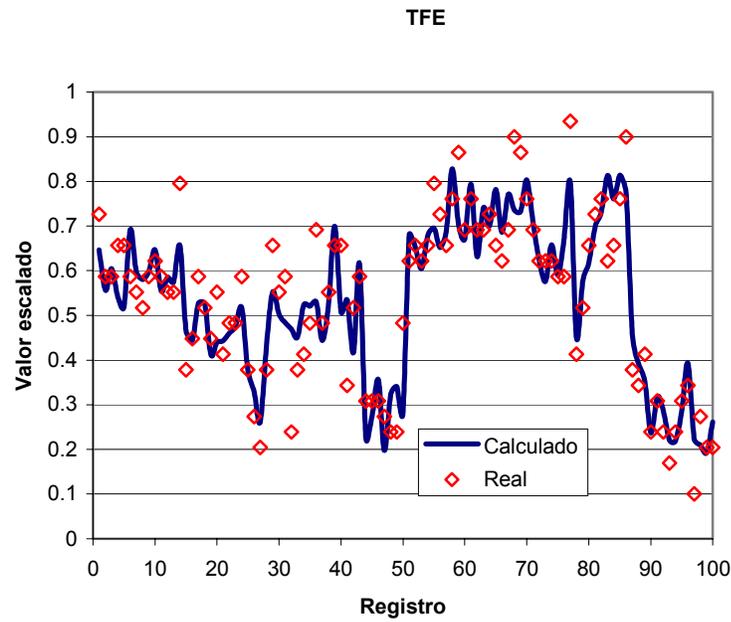


Figura 4.3 TFE del producto, fase de entrenamiento

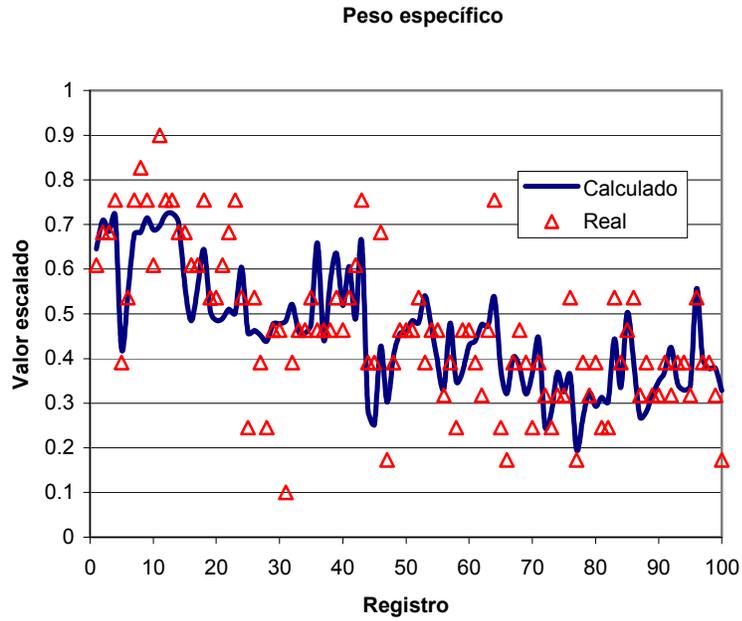


Figura 4.4 Peso específico del producto, fase de entrenamiento

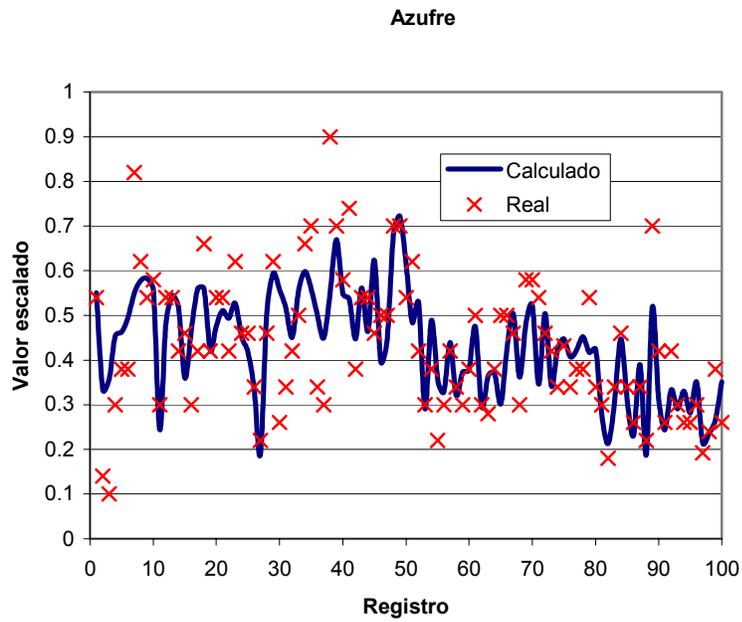


Figura 4.5 Azufre en el producto, fase de entrenamiento

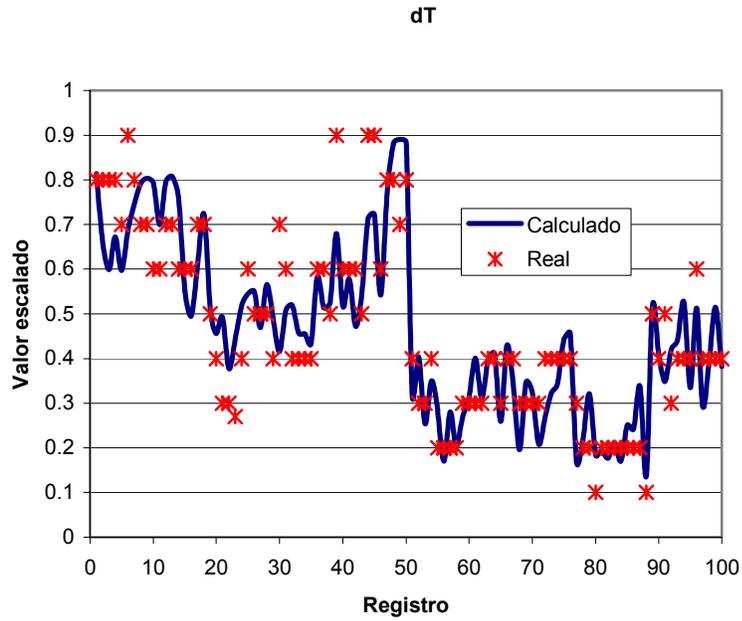


Figura 4.6  $\Delta T$  del lecho catalítico, fase de entrenamiento

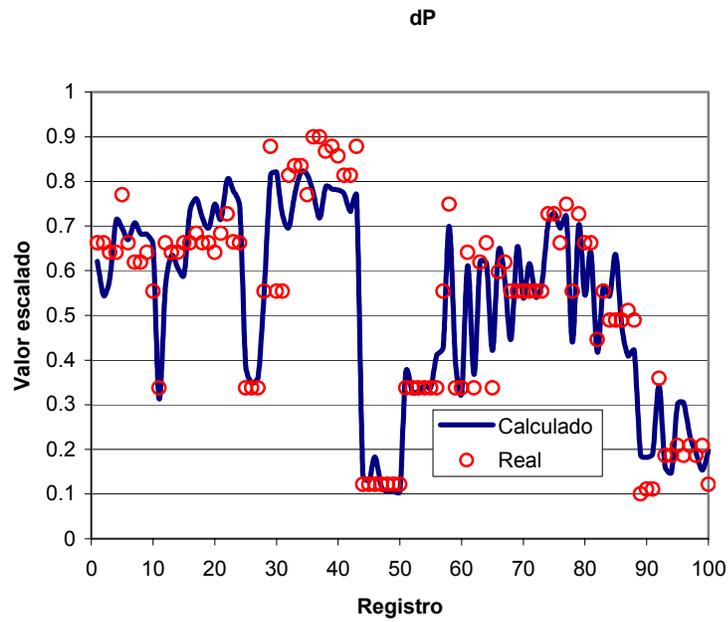


Figura 4.7 Caída de presión del lecho catalítico, fase de entrenamiento.

Las figuras 4.1 a 4.7 muestran los valores calculados y reales de las variables de salida de la topología "Global": producción, TIE de producto, TFE de producto, peso específico del producto, azufre en el producto,  $\Delta T$ , y  $\Delta P$ .

Como se puede observar existen variables que son mejor ajustadas tales como: TIE, TFE, y peso específico; mientras que la producción, el azufre, y las deltas, tienen un ajuste de menor calidad. Esto pone de manifiesto que existen valores para estas últimas variables que provocan que el ajuste no sea el adecuado, aun a pesar de haber eliminado valores conflictivos a través del análisis estadístico.

Además se puede observar que no en todos los casos se tuvo como límite máximo el conjunto de pesos de -12 a 12; esto se explica debido a que es necesario variar el conjunto de pesos iniciales hasta encontrar el intervalo en el cual se presenta un mínimo en el error promedio, tal es el caso para las topologías: 10-10-7, 10-12-7, 10-14-7, y 10-16-7.

Por otra parte, es evidente que los errores disminuyen conforme se aumentan los nodos de la capa oculta, esto podría llevar a pensar que entre más nodos ocultos tenga la red neuronal mejores serán las predicciones de dicha red. Sin embargo, esto es sólo una falsa estimación, ya que al aumentar indiscriminadamente los nodos de la capa oculta, lo que se está haciendo es que la red neuronal "memorice" los patrones de datos, y no que desarrolle la capacidad de "generalizar", es decir predecir adecuadamente aún para datos que no se hayan presentado a la red con anterioridad.

En caso de que la red neuronal "memorice", será totalmente ineficiente al predecir las variables correspondientes a los datos que no le hayan sido presentados, lo cual es una característica no deseable.

Por esta razón, es necesario tener otro criterio de selección y esto lo proporciona la metodología denominada "validación cruzada", ya descrita en la [sección 3.3](#).

Para nuestro caso, a la red neuronal en cuestión se le presentaron simultáneamente datos de entrenamiento y datos de validación, estos últimos son datos que la red no utiliza para su entrenamiento y por lo tanto no ha visto. Se monitorea la variación del error de validación y en cuanto se empieza a incrementar, se detiene el proceso de entrenamiento, esto se hace para diferentes conjuntos de pesos iniciales como se vio anteriormente, obteniéndose los siguientes resultados de la tabla 4.12.

Tabla 4.12 Errores de validación de la Red Neuronal "Global".

TOPOLOGÍA	PESOS	ERROR RMS DE VALIDACIÓN
10-6-7	-6,6	0.249778
10-7-7	-10,10	0.294936
10-8-7*	-10,10*	0.234573*
10-9-7	-8,8	0.286564
10-10-7	-18,18	0.251978
10-12-7	-8,8	0.252247
10-14-7	-14,14	0.331277
10-16-7	-6,6	0.296585
10-18-7	-6,6	0.314288

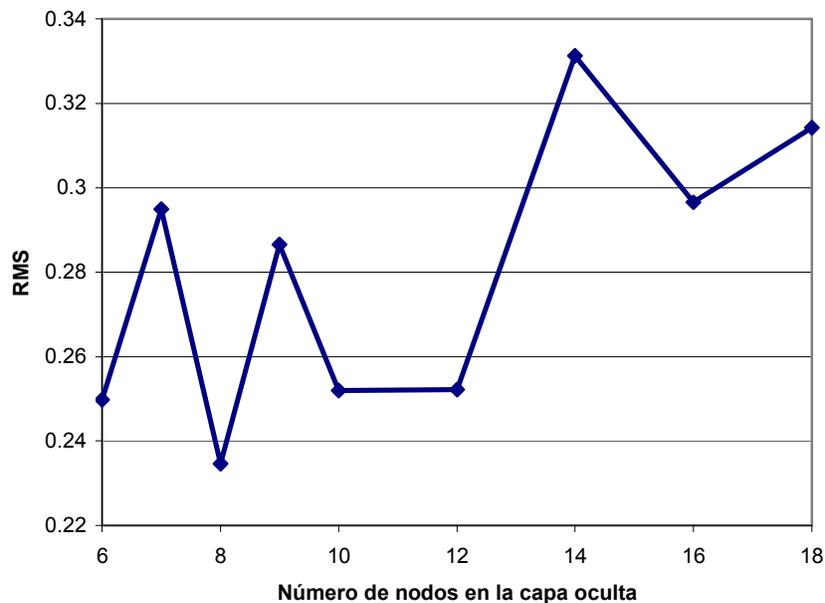


Figura 4.8 Error RMS de validación. Caso "Global"

Como se observa en la tabla 4.12 y la figura 4.8, la topología que proporciona el menor error RMS es la: 10-8-7, y es la que será utilizada para el diseño de la red neuronal que represente la operación del Reactor de Hidrodesulfuración de Naftas.

En las siguientes figuras se presentan los valores reales y calculados de las variables del Reactor de Hidrodesulfuración de Naftas, usando la topología "Global" (10-8-7). Los valores de entrenamiento y validación se muestran en la misma gráfica. Los 100

primeros puntos corresponden a los valores de entrenamiento, mientras que los 61 restantes corresponden a los valores de validación.

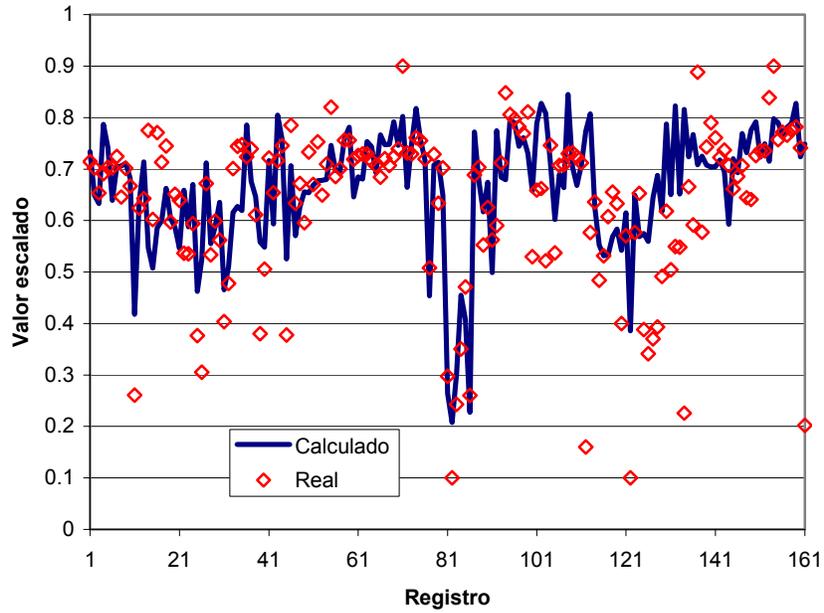


Figura 4.9 Comparación de valores de producción. Caso "Global".

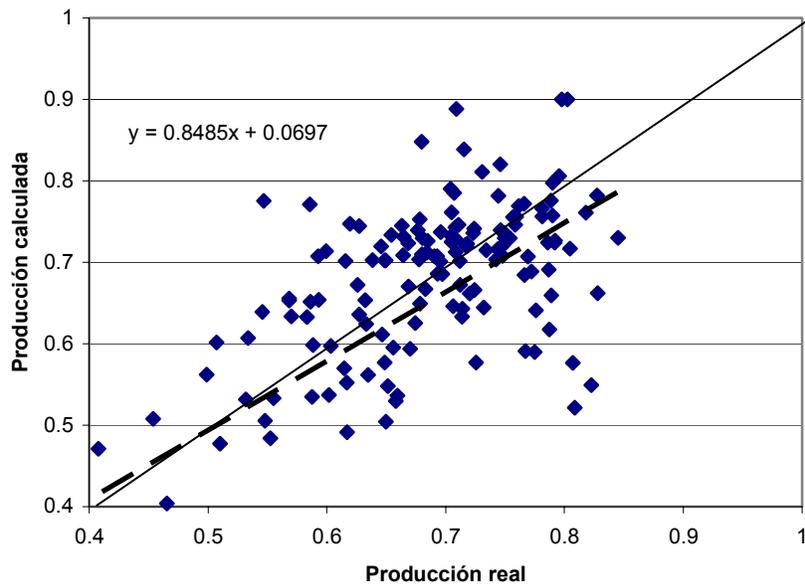


Figura 4.10 Dispersión de los valores de producción. Caso "Global".

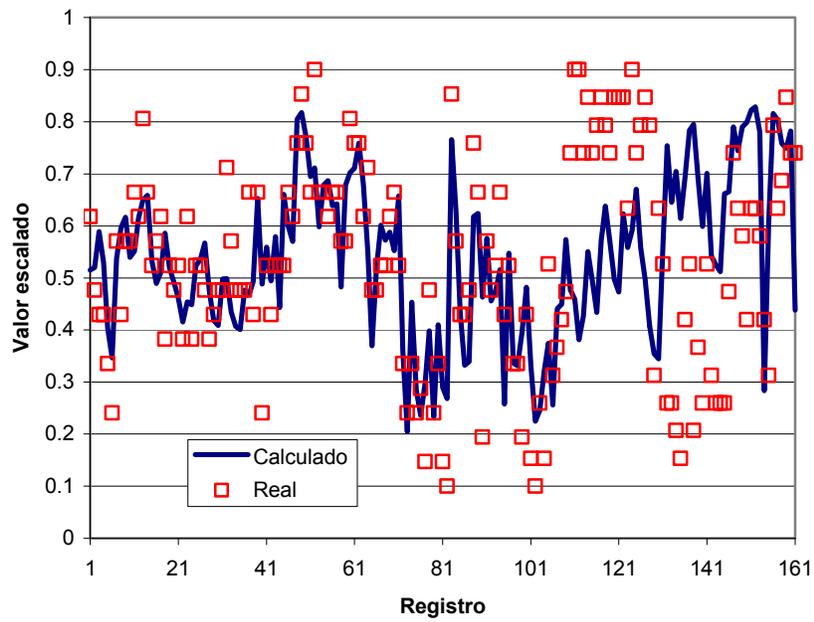


Figura 4.11 Comparación de valores de TIE producto. Caso "Global".

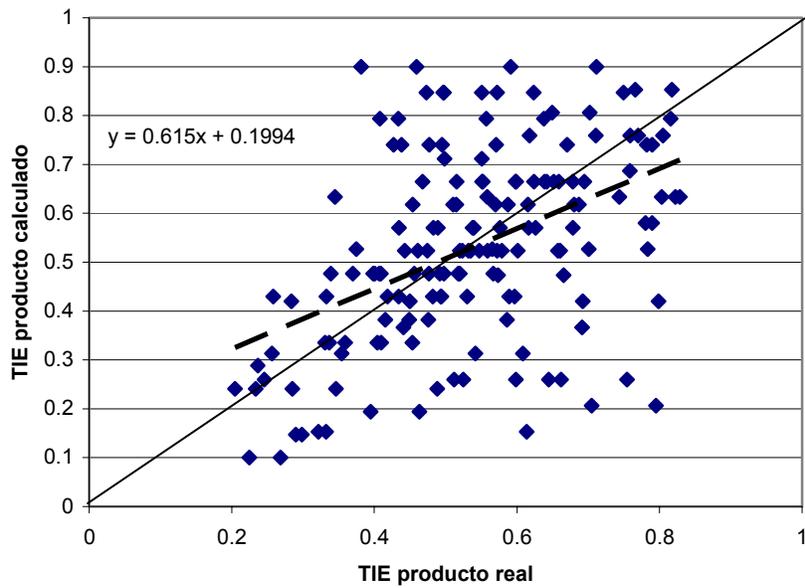


Figura 4.12 Dispersión de los valores de TIE producto. Caso "Global".

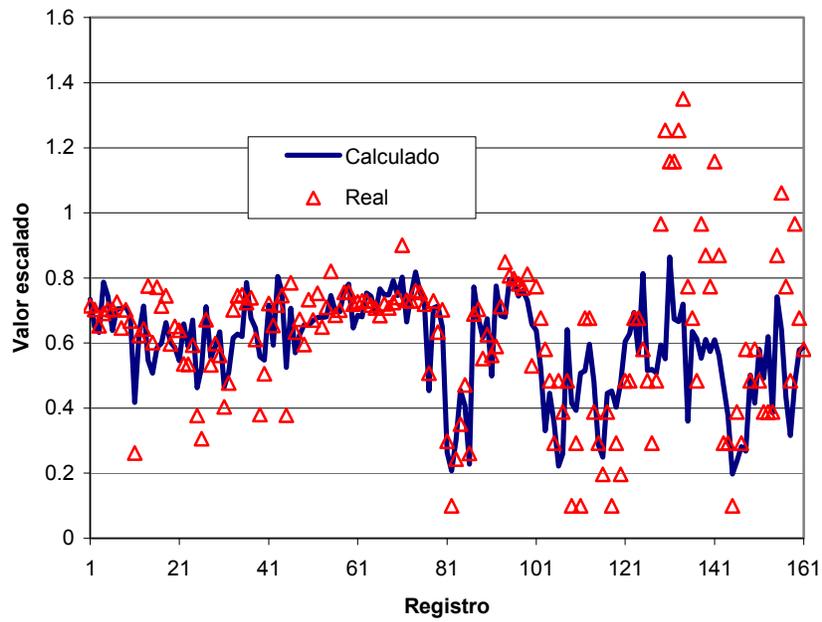


Figura 4.13 Comparación de valores de TFE producto. Caso "Global".

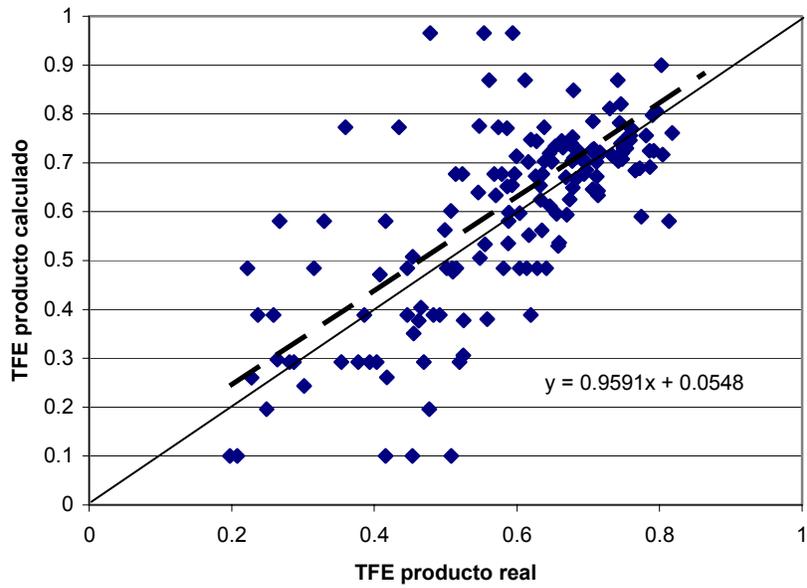


Figura 4.14 Dispersión de los valores de TFE producto. Caso "Global".

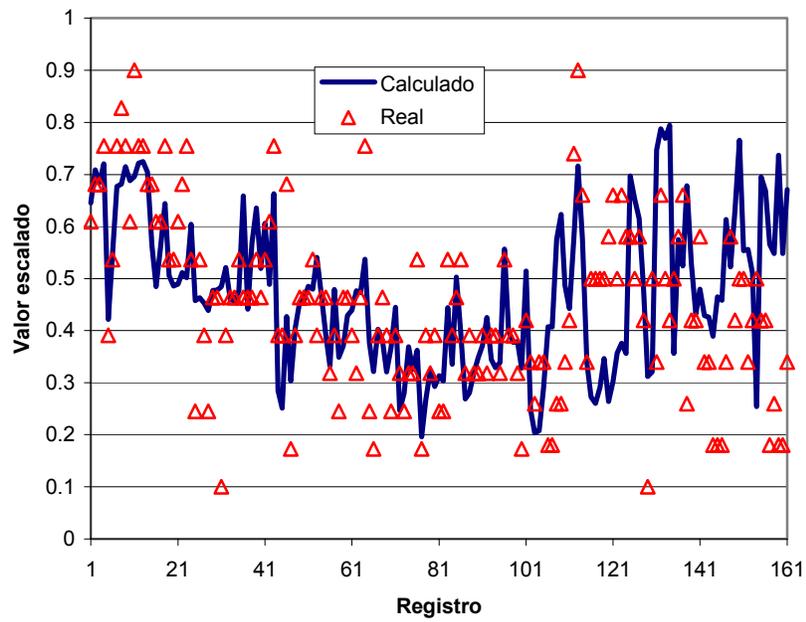


Figura 4.15 Comparación de valores de peso específico producto. Caso "Global".

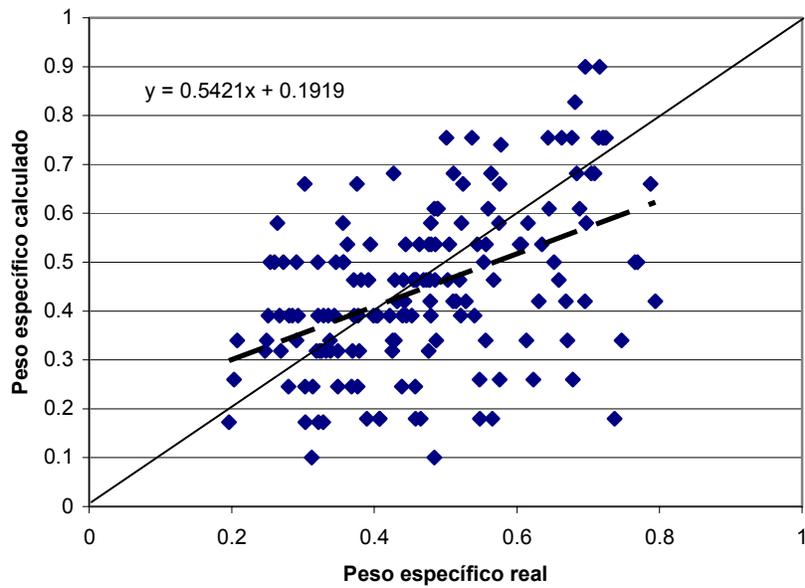


Figura 4.16 Dispersión de los valores de peso específico producto. Caso "Global".

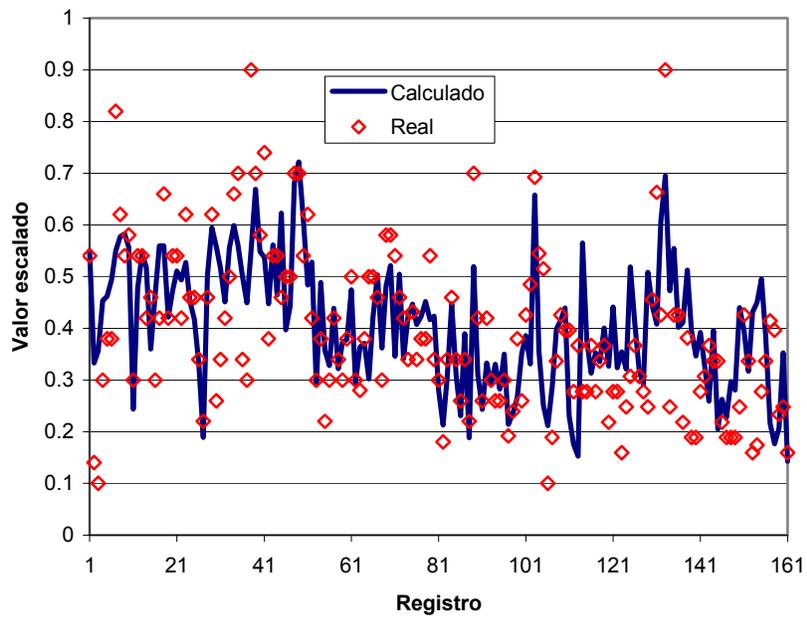


Figura 4.17 Comparación de valores de azufre producto. Caso "Global".

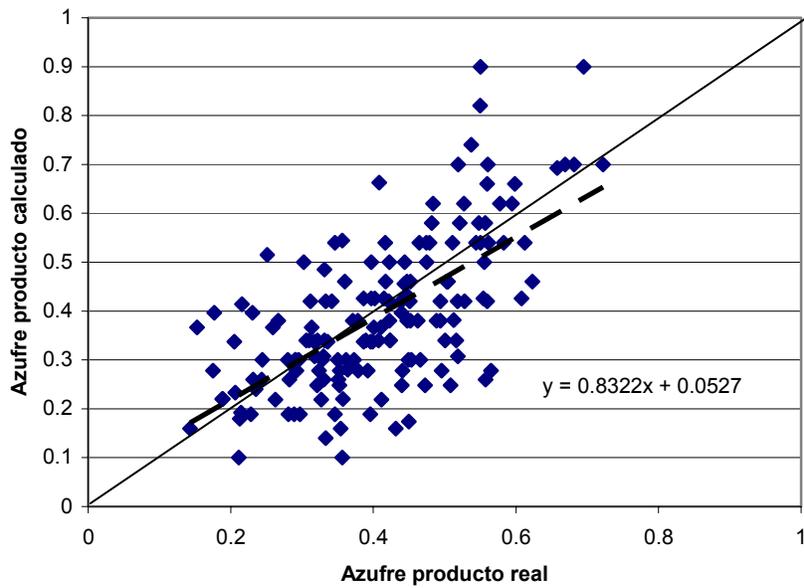


Figura 4.18 Dispersión de datos de azufre producto. Caso "Global".

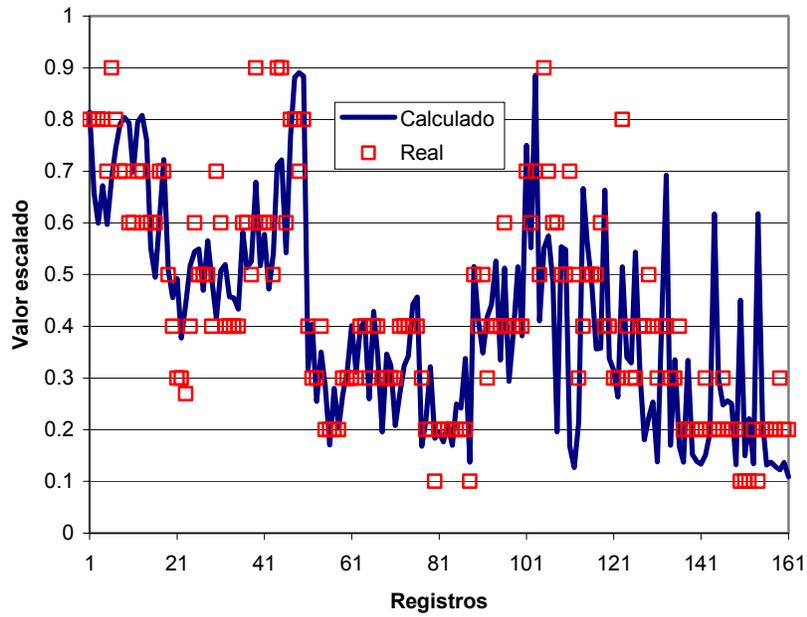


Figura 4.19 Comparación de valores de  $\Delta T$ . Caso "Global".

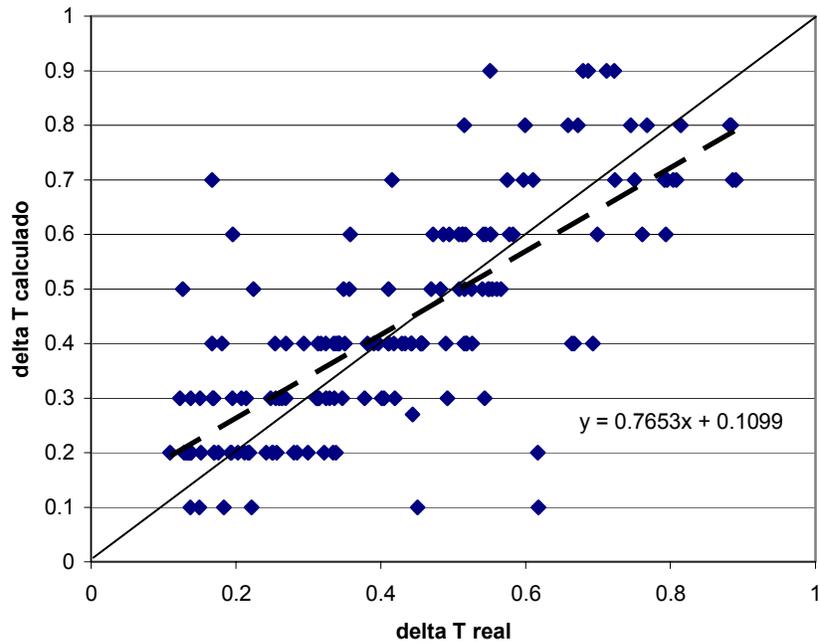


Figura 4.20 Dispersión de datos de  $\Delta T$ . Caso "Global".

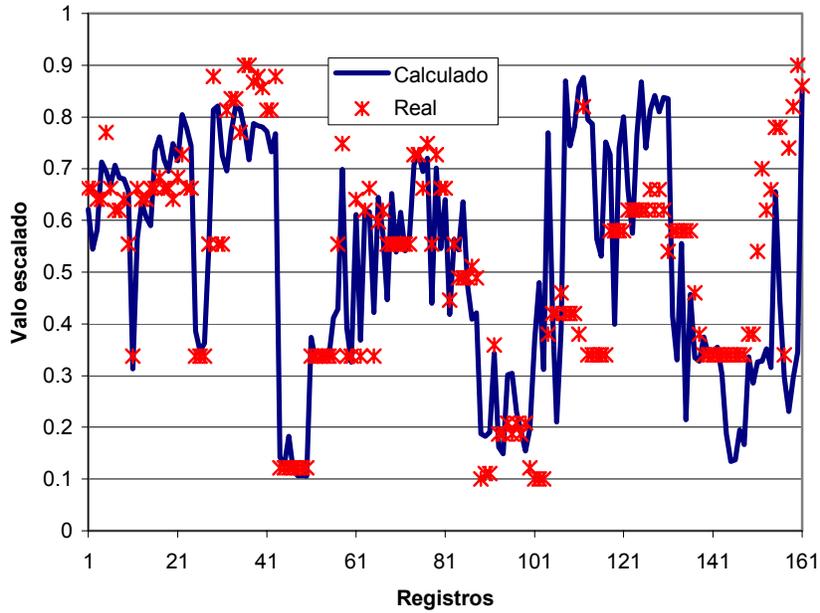


Figura 4.21 Comparación de datos de  $\Delta P$ . Caso "Global".

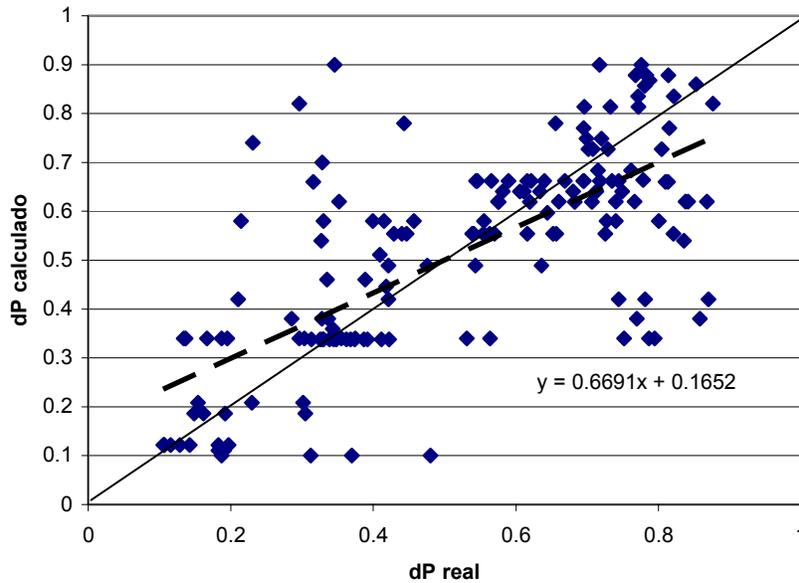


Figura 4.22 Dispersión de datos de  $\Delta P$ . Caso "Global".

En cada par de figuras de la 4.9 a la 4.22 se muestra en primer lugar la comparación entre los valores calculados y los valores reales, con lo que se tiene una idea de la bondad de la capacidad predictiva del modelo de redes neuronales. Por otra parte las

figuras 4.10, 4.12, 4.14, 4.16, 4.18, 4.20 y 4.22, muestran las gráficas de dispersión correspondientes a cada una de las variables de salida.

Una vez analizadas las gráficas antes mencionadas podemos concluir que las variables que presentan mayor dispersión de los valores calculados son: TFE producto, peso específico producto,  $\Delta T$ , y  $\Delta P$ .

Por esta razón y en un intento de obtener un modelo que proporcione mejores predicciones, se divide la topología “Global” en dos topologías alternativas denominadas “Proceso” y “Propiedades”. En la primera se han eliminado como variables de entrada y salida aquellas correspondientes a las propiedades (TIE, TFE, y peso específico) de la carga y producto, ya que al parecer dichas variables son las que provocan que la predicción del modelo sea adecuada. En la topología “Propiedades” sólo se analizan las variables TIE, TFE y peso específico para la carga y producto.

La topología “Proceso” se describe en el apartado 4.2, mientras que la topología “Propiedades” se presenta en el 4.3.

## 4.2. TOPOLOGÍA PROCESO.

Las variables de entrada y salida que se consideraron para esta topología se muestran en la tabla 4.13.

Tabla 4.13 Clasificación de las variables, topología 7-X-4.

ENTRADA	SALIDA
Flujo de carga	Flujo de producto
Temperatura de operación	Azufre producto
Presión de operación	$\Delta T$
Relación $H_2/HC$	$\Delta P$
Pureza de $H_2$	
Consumo de $H_2$	
Azufre de la carga	

Las topologías que se estudiaron de acuerdo al número de nodos en la capa oculta son:

- 7-4-4
- 7-5-4
- 7-6-4
- 7-7-4
- 7-8-4
- 7-10-4
- 7-12-4, y
- 7-14-4.

Los intervalos de pesos iniciales que se investigaron son los siguientes:

- -1,1
- -2,2
- -4,4
- -6,6
- -8,8
- -10,10
- -12,12
- -14,14
- -16,16
- -18,18

Después de la fase de entrenamiento se obtuvieron los siguientes valores del error promedio RMS para cada juego de pesos iniciales:

Tabla 4.14 Errores promedio de la topología 7-4-4, entrenamiento (Caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.142829
-2,2	0.12572
-4,4	0.115266
-6,6	0.112271
-8,8	0.113203
-10,10	0.112148*
-12,12	0.119184
-14,14	0.113865

Tabla 4.15 Errores promedio de la topología 7-5-4, entrenamiento (Caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.139144
-2,2	0.120486
-4,4	0.109595
-6,6	0.105536
-8,8	0.110153
-10,10	0.109455
-12,12	0.105092*
-14,14	0.10584
-16,16	0.107243

Tabla 4.16 Errores promedio de la topología 7-6-4, entrenamiento (Caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.135004
-2,2	0.117268
-4,4	0.103752
-6,6	0.102368
-8,8	0.104286
-10,10	0.101795
-12,12	0.105866
-14,14	0.100913*
-16,16	0.10194
-18,18	0.103925

Tabla 4.17 Errores promedio de la topología 7-7-4, entrenamiento (Caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.132137
-2,2	0.112799
-4,4	0.102605
-6,6	0.0997066
-8,8	0.0993315
-10,10	0.099642
-12,12	0.0975986
-14,14	0.0970233
-16,16	0.0965694*
-18,18	0.0978888

Tabla 4.18 Errores promedio de la topología 7-8-4, entrenamiento (Caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.130945
-2,2	0.109876
-4,4	0.0986918
-6,6	0.0932456
-8,8	0.0916651*
-10,10	0.0939812
-12,12	0.0939437

Tabla 4.19 Errores promedio de la topología 7-10-4, entrenamiento (Caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.127043
-2,2	0.105391
-4,4	0.0890707
-6,6	0.0855472
-8,8	0.077791*
-10,10	0.084521
-12,12	0.0791162
-14,14	0.0835304

Tabla 4.20 Errores promedio de la topología 7-12-4, entrenamiento (Caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.123535
-2,2	0.0979916
-4,4	0.0811921
-6,6	0.0745149
-8,8	0.072374*
-10,10	0.0800731
-12,12	0.0744177
-14,14	0.0738009

Tabla 4.21 Errores promedio de la topología 7-14-4, entrenamiento (caso "Proceso").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.121952
-2,2	0.0952446
-4,4	0.0753456
-6,6	0.0681993
-8,8	0.0665019*
-10,10	0.0685809
-12,12	0.0701096
-14,14	0.0717397

Al igual que en el caso anterior el criterio para seleccionar la mejor topología es aquella que proporcione el menor error de validación. Para este caso dichos errores se muestran en la tabla 4.22, y gráficamente en la figura 4.23.

Tabla 4.22 Errores de validación de la Red Neuronal "Proceso"

TOPOLOGÍA	PESOS	ERROR RMS DE VALIDACIÓN
7-4-4	-10,10	0.199749
7-5-4*	-12,12*	0.169485*
7-6-4	-14,14	0.214154
7-7-4	-16,16	0.221837
7-8-4	-8,8	0.233079
7-10-4	-8,8	0.210391
7-12-4	-8,8	0.237859
7-14-4	-8,8	0.225085

(\*) Valor mínimo.

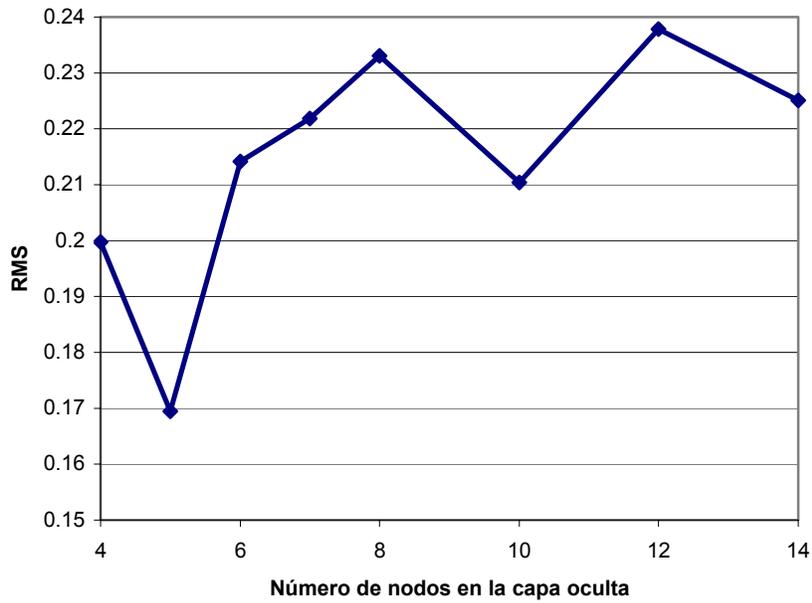


Figura 4.23 Error RMS de validación. Caso “Proceso”

Como se puede observar la topología que proporciona el menor error RMS es la 7-5-4, por lo que es la configuración seleccionada para efectuar el análisis de predicción.

En las siguientes figuras se presentan los valores medidos y calculados de las variables del Proceso de Hidrodesulfuración de Naftas, para la topología “Proceso” (7-5-4). Los valores de entrenamiento y validación se muestran en la misma gráfica. Los 100 primeros puntos corresponden a los valores de entrenamiento, mientras que los 61 restantes corresponden a los valores de validación.

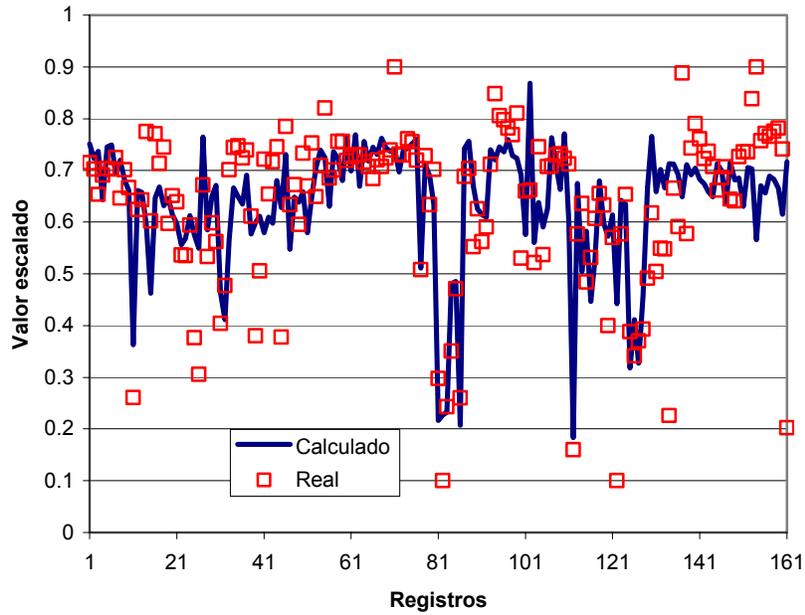


Figura 4.24 Comparación de datos de producción. Caso "Proceso".

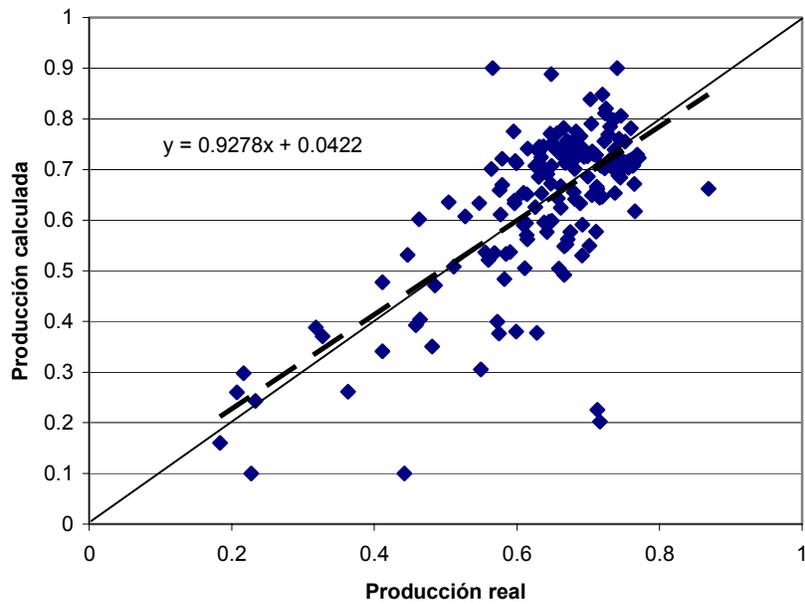


Figura 4.25 Dispersión de datos de producción. Caso "Proceso".

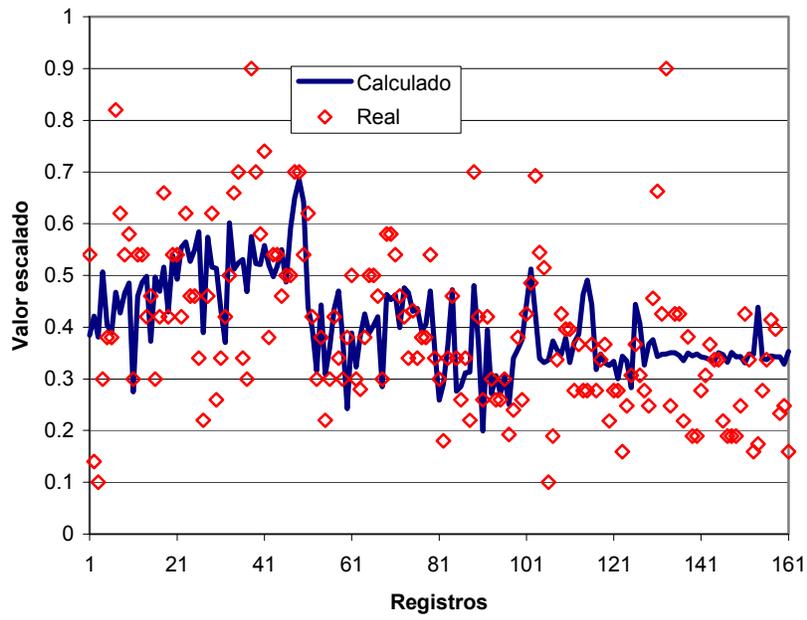


Figura 4.26 Comparación de datos de azufre. Caso "Proceso".

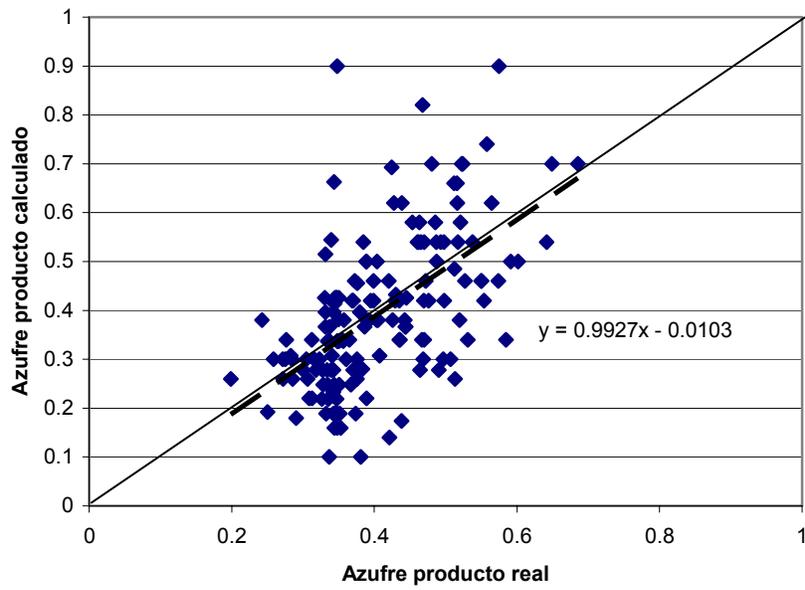


Figura 4.27 Dispersión de datos de azufre. Caso "Proceso".

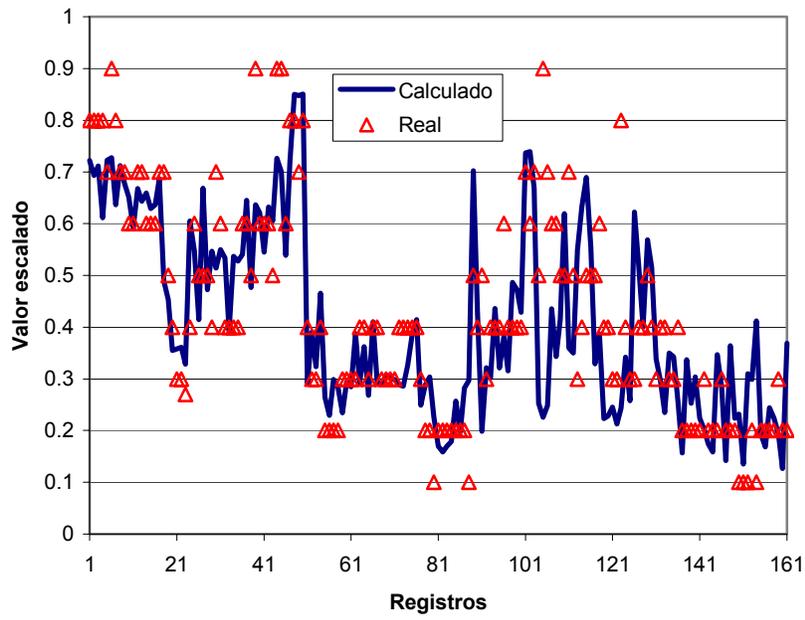


Figura 4.28 Comparación de datos de  $\Delta T$ . Caso "Proceso".

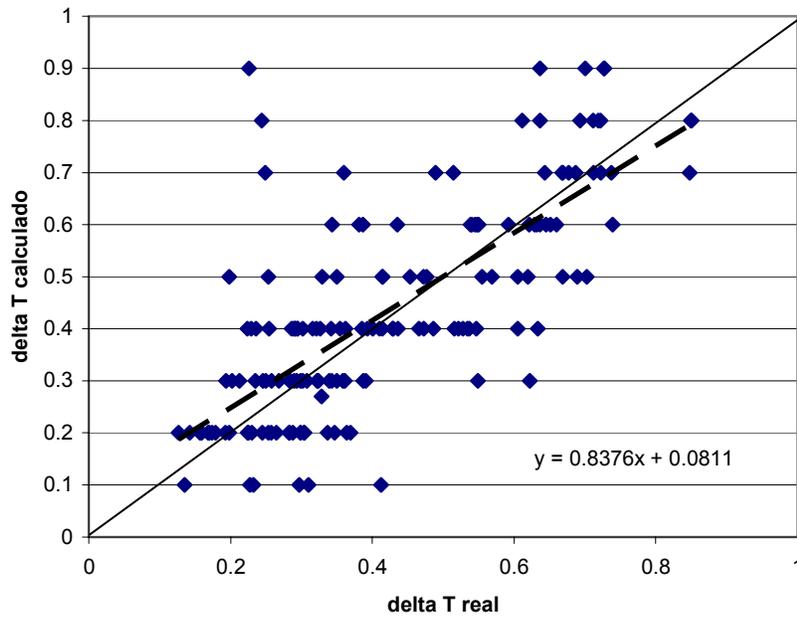


Figura 4.29 Dispersión de datos de  $\Delta T$ . Caso "Proceso".

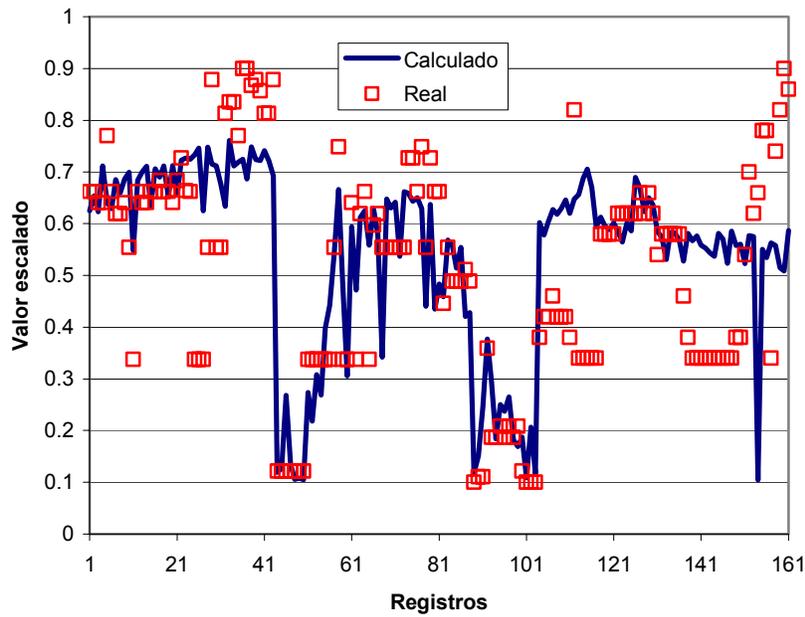


Figura 4.30 Comparación de datos de  $\Delta P$ . Caso "Proceso".

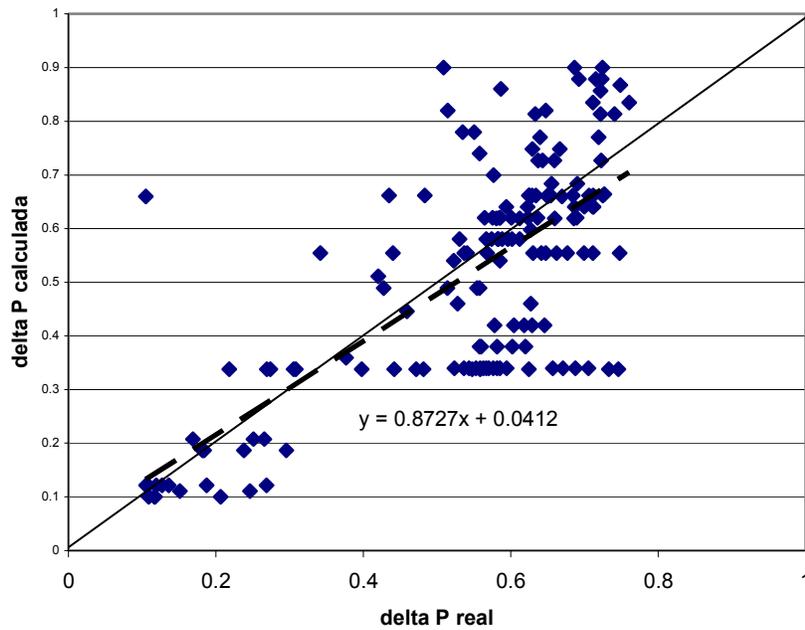


Figura 4.31 Dispersión de datos de  $\Delta P$ . Caso "Proceso".

Comparando las ecuaciones de regresión de los valores calculados por el modelo de RNA de las topologías "Global" y "Proceso", se puede observar que al dividir la topología

“Global” se mejora considerablemente la capacidad predictiva del modelo. Sólo por poner un ejemplo, la ecuación de regresión para el azufre pasa de:

$$y = 0.8322x + 0.0527$$

en la topología global, a:

$$y = 0.9927x + 0.0103$$

en la topología “Proceso”, y como se puede observar los valores numéricos en esta última ecuación están más cerca de los valores ideales, indicando así una mejor regresión.

### 4.3. TOPOLOGÍA PROPIEDADES

Debido a que en el caso “Global” el ajuste para la predicción de las propiedades del producto no fue del todo adecuado, se decidió que en el caso “Proceso” dichas variables no fueran incluidas. Sin embargo, es necesario contar con un modelo que permita predecir en forma adecuada las propiedades de la Nafta producto.

Por esta razón, en esta sección se presenta el análisis realizado con una topología que incluye exclusivamente la predicción de las propiedades de la Nafta producto, usando como variables de entrada las propiedades de la Nafta carga, quedando la distribución de variables de la siguiente manera:

Tabla 4.23 Clasificación de las variables. Topología “Propiedades”.

ENTRADA	SALIDA
TIE carga	TIE producto
TFE carga	TFE producto
Peso específico carga	Peso específico producto

Como se puede observar la propuesta de proporcionalidad es 1:1, buscando con esto obtener una mejor predicción de las propiedades del producto.

Para este caso se evaluarán las siguientes topologías:

- 3-2-3
- 3-3-3
- 3-4-3
- 3-5-3
- 3-6-3
- 3-8-3
- 3-10-3

Con los intervalos de pesos iniciales siguientes:

- -1,1
- -2,2
- -4,4
- -6,6
- -8,8
- -10,10
- -12,12

Una vez realizada la fase entrenamiento, se obtuvieron los siguientes valores de error RMS para cada una de las topologías propuestas.

Tabla 4.24 Errores promedio de la topología 3-2-3, entrenamiento (Caso “Propiedades”).

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.153819
-2,2	0.128375
-4,4	0.124492*
-6,6	0.124954
-8,8	0.124903
-10,10	0.124914

Tabla 4.25 Errores promedio de la topología 3-3-3, entrenamiento (caso “Propiedades”).

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.140903
-2,2	0.112272
-4,4	0.107748
-6,6	0.107596*
-8,8	0.107932
-10,10	0.107638

Tabla 4.26 Errores promedio de la topología 3-4-3, entrenamiento (Caso "Propiedades").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.130152
-2,2	0.109782
-4,4	0.107292
-6,6	0.104969*
-8,8	0.105059
-10,10	0.105026

Tabla 4.27 Errores promedio de la topología 3-5-3, entrenamiento (Caso "Propiedades").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.123875
-2,2	0.108869
-4,4	0.102246*
-6,6	0.103342
-8,8	0.103455

Tabla 4.28 Errores promedio de la topología 3-6-3, entrenamiento (Caso "Propiedades").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.119895
-2,2	0.107481
-4,4	0.101281
-6,6	0.0992013*
-8,8	0.100618
-10,10	0.100315

Tabla 4.29 Errores promedio de la topología 3-8-3, entrenamiento (Caso "Propiedades").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.114003
-2,2	0.105108
-4,4	0.0973941*
-6,6	0.0977677
-8,8	0.0994959

Tabla 4.30 Errores promedio de la topología 3-10-3, entrenamiento (Caso "Propiedades").

INTERVALO DE PESOS INICIALES	RMS Promedio
-1,1	0.112188
-2,2	0.104455
-4,4	0.0978345
-6,6	0.0947621
-8,8	0.0922965*
-10,10	0.0958134
-12,12	0.0950974

Los valores marcados con asterisco (\*) en las tablas anteriores indican que para esos conjuntos de pesos iniciales se obtuvieron los menores errores RMS para la fase de entrenamiento, y por lo tanto esos pesos fueron utilizados para realizar la validación de la red neuronal. Los resultados de la validación se muestran en la tabla 4.31 siguiente:

Tabla 4.31 Errores de validación para la topología "Propiedades".

TOPOLOGÍA	PESOS	ERROR RMS DE VALIDACIÓN
3-2-3	-4,4	0.220011
3-3-3	-6,6	0.215923*
3-4-3	-6,6	0.218684
3-5-3	-4,4	0.220737
3-6-3	-6,6	0.229017
3-8-3	-4,4	0.229940
3-10-3	-8,8	0.241463

La representación gráfica de estos valores se muestra en la figura 4.32.

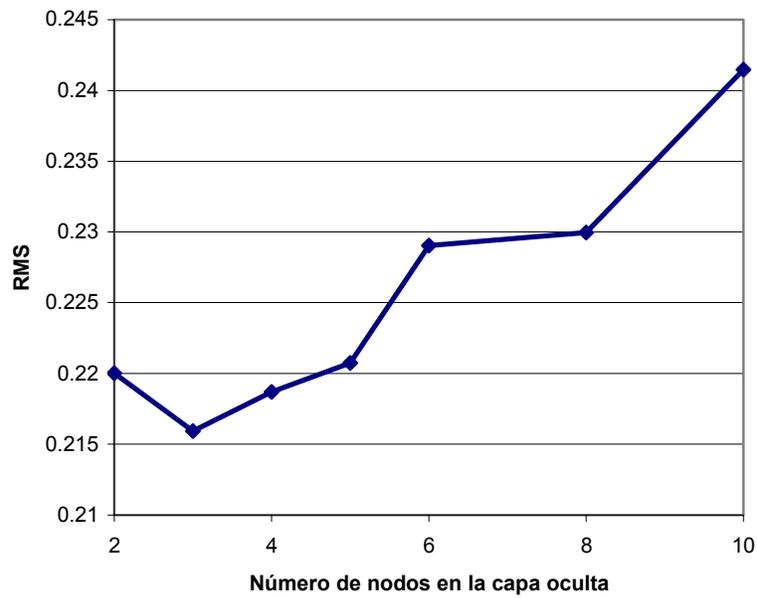


Figura 4.32 Error RMS de validación. Caso "Propiedades".

Como se puede observar claramente en la figura anterior, la topología que proporciona el menor valor de error RMS de validación es la 3-3-3, la cual será utilizada para realizar el análisis correspondiente.

En las siguientes figuras se muestra la comparación entre los valores de las propiedades obtenidos, con sus correspondientes valores reales; así como las gráficas de distribución para cada una de las propiedades.

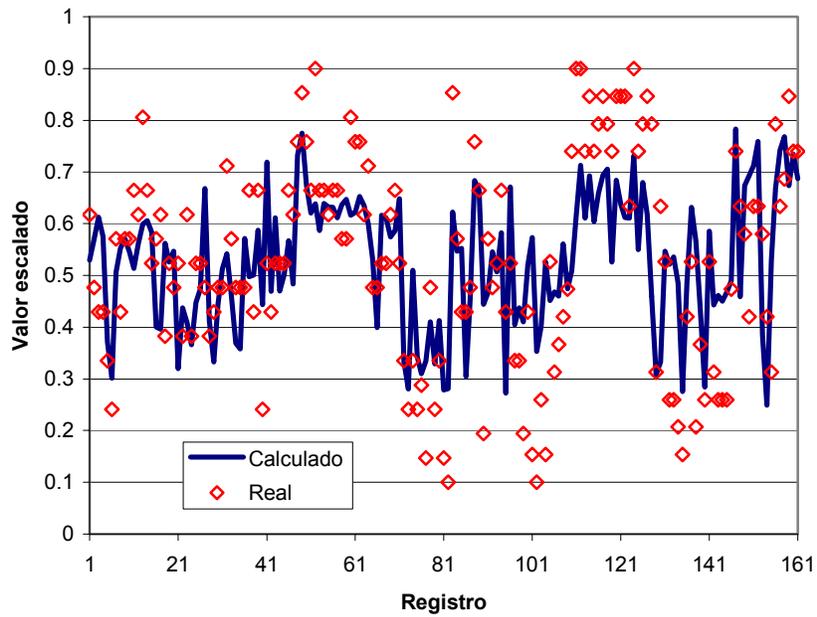


Figura 4.33 Comparación de datos TIE producto. Caso "Propiedades".

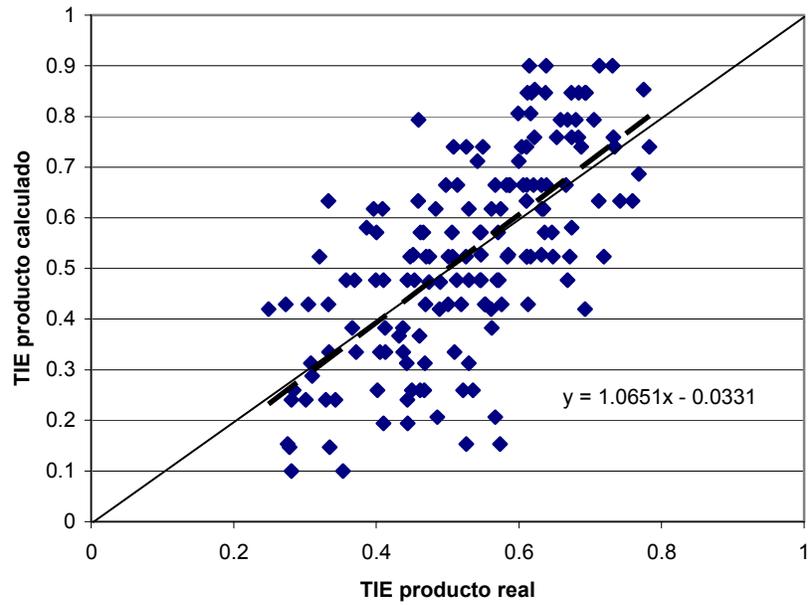


Figura 4.34 Dispersión de datos de TIE producto. Caso "Propiedades".

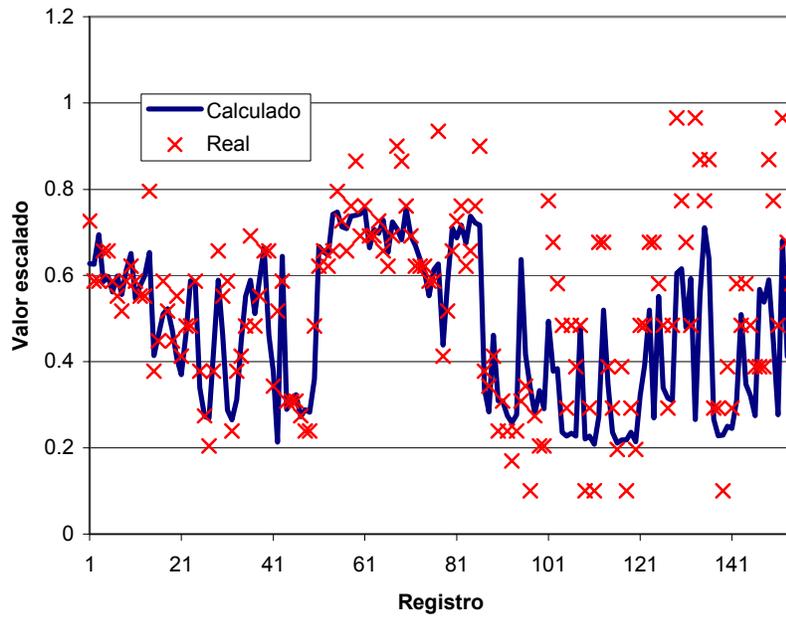


Figura 4.35 Comparación de datos TFE producto. Caso "Propiedades".

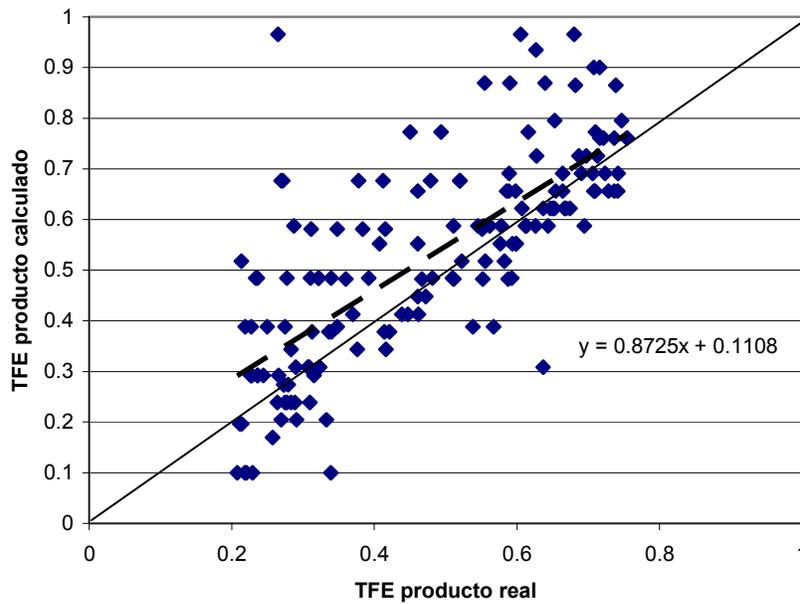


Figura 4.36 Dispersión de datos de TFE producto. Caso "Propiedades"

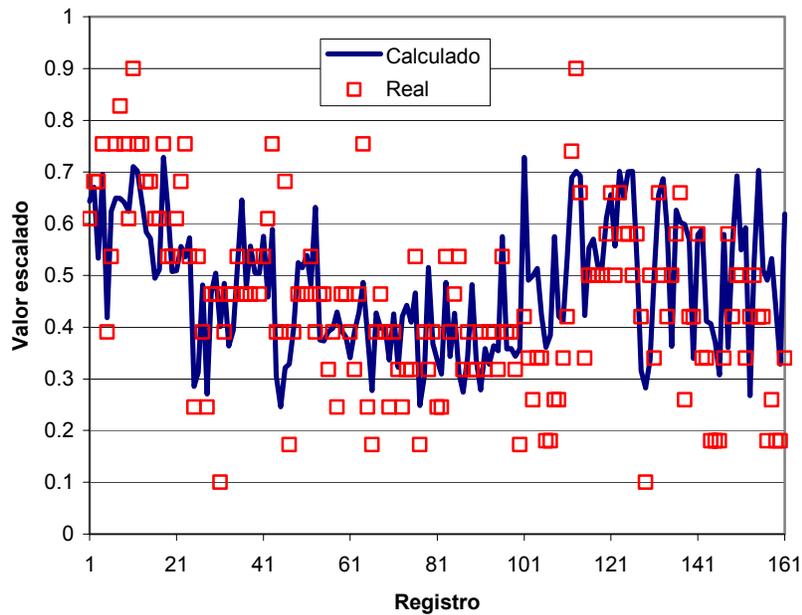


Figura 4.37 Comparación de datos peso específico producto. Caso “Propiedades”.

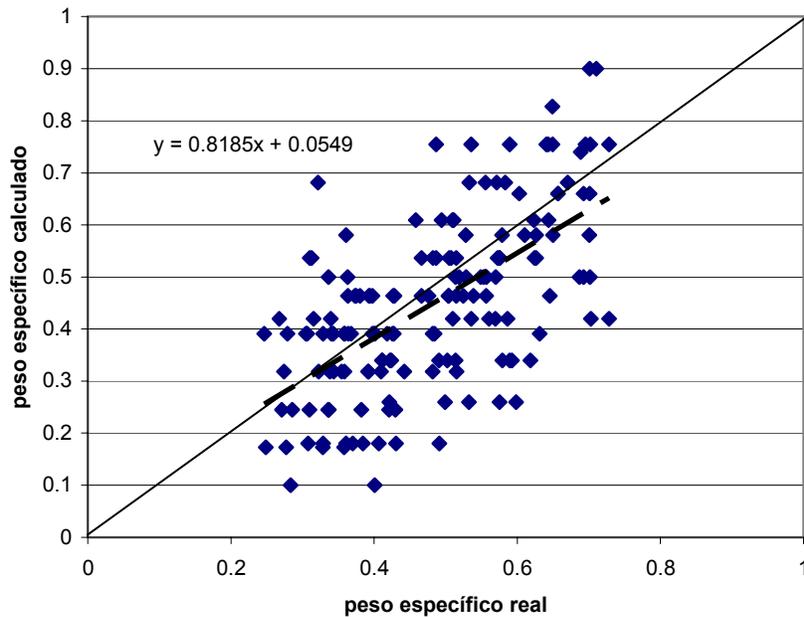


Figura 4.38 Dispersión de datos de peso específico producto. Caso “Propiedades”

Al igual que para el caso de la topología “Proceso” antes analizada, se puede ver que la topología “Propiedades” también ofrece una considerable mejora en la capacidad predictiva del modelo, lo cual también se puede demostrar comparando las ecuaciones de la regresión lineal de las propiedades del producto.

## 4.4. SELECCIÓN DE LA TOPOLOGÍA.

A fin de realizar la selección de la topología que mejor reproduce las variables de salida del Reactor de Hidrodesulfuración de Naftas, se toman como base las gráficas de dispersión mostradas en las figuras 4.10, 4.12, 4.14, 4.16, 4.18, 4.20, 4.22, 4.25, 4.27, 4.29, 4.31, 4.34, 4.36, y 4.38.

En cada una de las gráficas antes mencionadas, se muestra una línea de referencia cuya ecuación es:

$$y = mx + b \quad (6.1)$$

donde:

m = pendiente, y

b = ordenada al origen.

En las mismas gráficas también se muestra la línea de regresión que mejor ajusta los valores de salida calculados por cada una de las topologías, para cada una de las variables.

Si se tuviera un modelo ideal que pronosticara exactamente los valores reales del proceso, los parámetros de las líneas de ajuste antes mencionadas tendrían un valor de la pendiente igual a 1.0, y de la ordenada al origen igual a 0.0. Sin embargo, debido que el modelo propuesto como cualquier otro modelo sólo proporciona una aproximación a los valores reales del proceso, una forma de estimar la bondad del modelo propuesto es comparar los valores de los parámetros de su línea de regresión con los valores antes mencionados.

Así pues, en la siguiente tabla se muestran los parámetros de regresión para las tres topologías: "Global", "Proceso" y "Propiedades".

Tabla 4.32 Comparación de parámetros de regresión.

VARIABLES	TOPOLOGÍA					
	GLOBAL		ALTERNATIVA		PROPIEDADES	
Parámetros	m	b	m	b	m	B
<b>SALIDA</b>						
Producción	0.8485	0.0697	0.9278	0.0422	----	----
TIE producto	0.6150	0.1994	----	----	1.0651	0.0331
TFE producto	0.9591	0.0548	----	----	0.8725	0.1108
Peso esp. producto	0.5421	0.1919	----	----	0.8185	0.0549
Azufre producto	0.8322	0.0527	0.9927	0.0103	----	----
$\Delta T$	0.7653	0.1099	0.8376	0.0811	----	----
$\Delta P$	0.6691	0.1652	0.8727	0.0412	----	----

Como se observa en la tabla 4.32, la división de la topología “Global” en las topologías “Proceso” y “Propiedades” mejora considerablemente la capacidad predictiva del modelo ya que los valores correspondientes de la pendiente y la ordenada al origen están más cerca de los valores ideales; a excepción de los valores de TFE producto, cuyos valores disminuyen.

Por este motivo se seleccionan como las topologías que mejor representan el modelo de Hidrodesulfuración de Naftas, la topología “Proceso” y la topología “Propiedades”, aun cuando disminuye la calidad del modelo para la variable TFE producto, se toma como una consideración poco impactante en comparación con las considerables mejoras que se obtienen en la capacidad predictiva de las variables restantes.

En las siguientes figuras se muestran esquemáticamente las topologías seleccionadas.

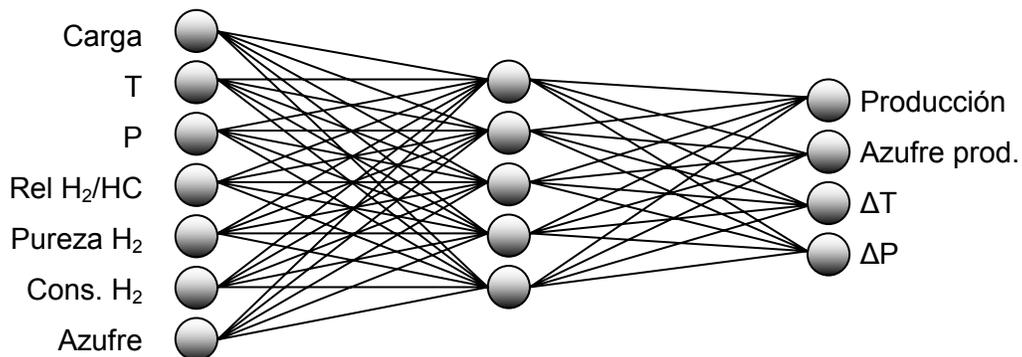


Figura 4.39 Topología “Proceso” seleccionada.

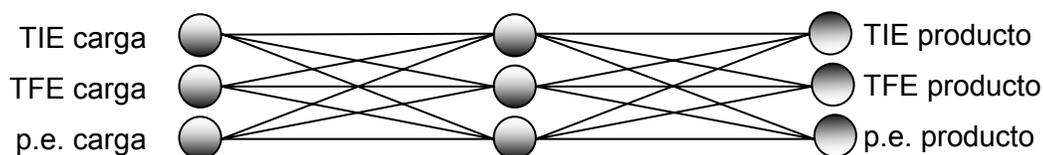


Figura 4.40 Topología “Propiedades” seleccionada.

Cada una de las líneas de interconexión entre los nodos de cada una de las capas de las redes neuronales representa un factor de ponderación (peso), y adicionalmente se debe considerar un nodo para la capa intermedia y la capa de salida denominado sesgo o polarización (bias).

Una vez realizado el procedimiento de entrenamiento-validación, se obtienen los valores tanto de los pesos como de los sesgos correspondientes. Las tablas 4.33 y 4.34 muestran estos valores para la topología “Proceso”, mientras que las tablas 4.35 y 4.36 muestran los correspondientes a la topología “Propiedades”.

Tabla 4.33 Valores de los pesos y sesgos para la topología “Proceso”. Capa oculta.

i/j	Pesos							Sesgo
	1	2	3	4	5	6	7	
1	1.5059	1.3920	0.8141	-1.0809	0.3824	-0.4959	0.6603	0.4779
2	-1.2621	1.9993	-6.0127	-1.5768	9.5844	-7.5290	-1.0450	-2.0176
3	-1.1503	-0.8137	-1.6417	3.9650	1.7947	2.0535	-1.9947	3.6232
4	-1.9201	1.7338	-11.036	-10.017	2.7998	-3.9106	2.2902	3.6828
5	-3.4480	3.2998	5.1820	-1.8530	-0.2446	-2.0292	2.6267	-0.2571

Tabla 4.34 Valores de los pesos y sesgos para la topología “Proceso”. Capa de salida.

i/j	Pesos					Sesgo
	1	2	3	4	5	
1	3.85194	-0.26152	-0.61319	-0.52496	-1.59257	-0.4241
2	0.64793	0.42798	-0.09543	-0.52980	0.137367	-0.8890
3	-1.73332	3.58732	-4.20251	-2.10412	-0.875547	3.4944
4	4.1864	-2.1257	2.38989	-0.10407	2.96715	-5.5664

Tabla 4.35 Valores de los pesos y sesgos para la topología “Propiedades”. Capa oculta

i/j	Pesos			Sesgo
	1	2	3	
1	-4.43867	-2.98272	-4.20848	0.610347
2	0.756819	0.938212	-0.834173	-0.856785
3	-0.379684	5.21926	1.15631	-0.0991506

Tabla 4.36 Valores de los pesos y sesgos para la topología “Propiedades”. Capa de salida.

i/j	Pesos			Sesgo
	1	2	3	
1	-4.43867	-2.98272	-4.20848	0.610347
2	0.756819	0.938212	-0.834173	-0.856785
3	-0.379684	5.21926	1.15631	-0.0991506

Los valores anteriores son los que se usan para generar el programa que realizará la simulación y estimará el comportamiento de un Reactor de Hidrodesulfuración de Naftas, y el cual será usado para evaluar el comportamiento del modelo.

Como parte complementaria de este trabajo se desarrollo el programa “**NeuroNaf**” el cual incluye el modelo de RNA desarrollado para predecir el comportamiento de las variables del Reactor de Hidrodesulfuración de Naftas. La descripción de dicho programa se puede ver en el [apéndice B](#), mientras que los resultados obtenidos se muestran en el siguiente capítulo.

## 5. ANÁLISIS DE RESULTADOS

### 5.1. PREDICCIÓN DE DATOS DE PRUEBA.

Una vez obtenidos los parámetros de la Red Neuronal, y haber realizado también la validación correspondiente lo que resta es probar la capacidad predictiva de la Red Neuronal presentándole juegos de datos que no hayan sido utilizados ni para la fase de entrenamiento, ni para la fase de validación. Con esta acción se podrá tener una idea verdadera acerca de la bondad de las predicciones hechas por el modelo desarrollado.

Se utilizaron 36 juegos de datos, los cuales fueron separados del conjunto original precisamente para la fase de prueba. Por cuestiones de confidencialidad no se presentan los juegos de datos completos, presentándose sólo las gráficas de los valores de las variables de salida, para nuestro caso: Producción, TIE del producto, TFE del producto, Peso específico del producto, Azufre del producto, delta T del proceso y delta P del lecho catalítico.

En las siguientes figuras se muestra en forma individual cada una de las variables antes mencionadas:

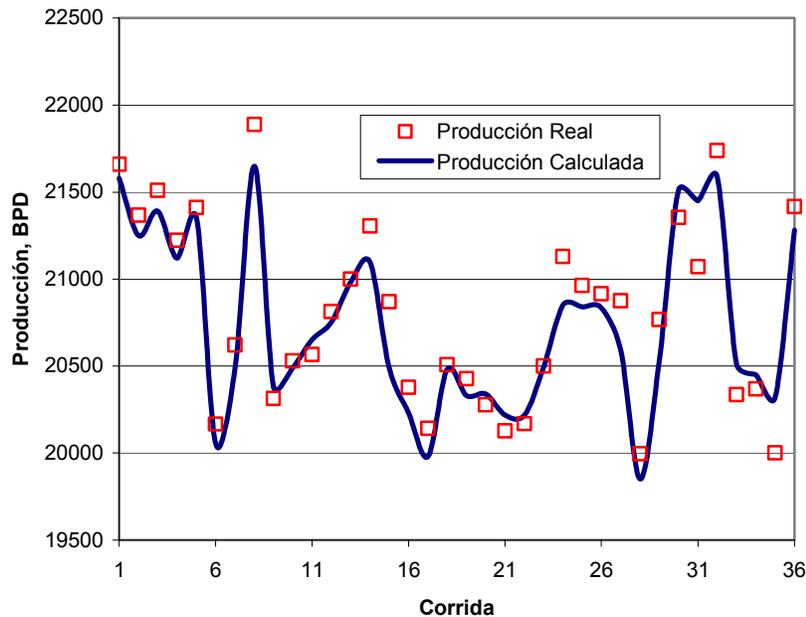


Figura 5.1 Comparación de datos de producción. Caso Prueba.

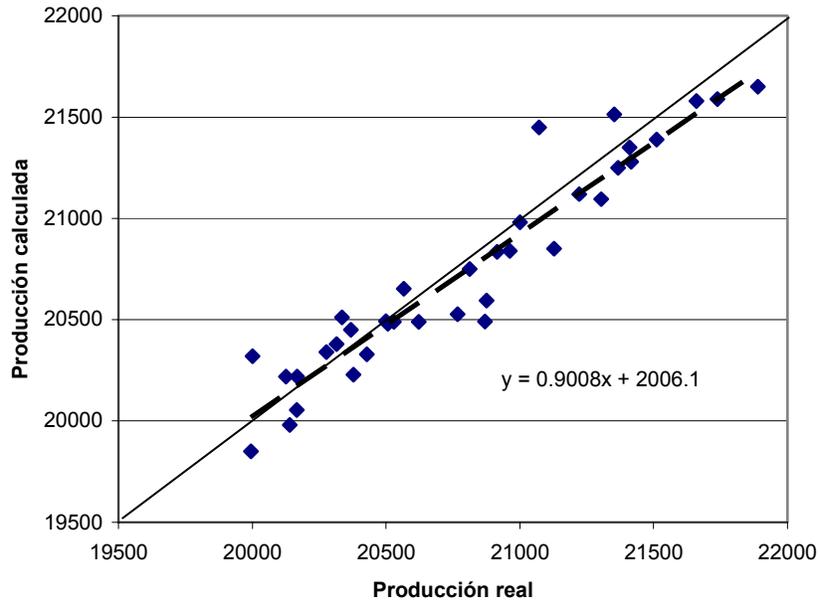


Figura 5.2 Dispersión de datos de producción. Caso Prueba.

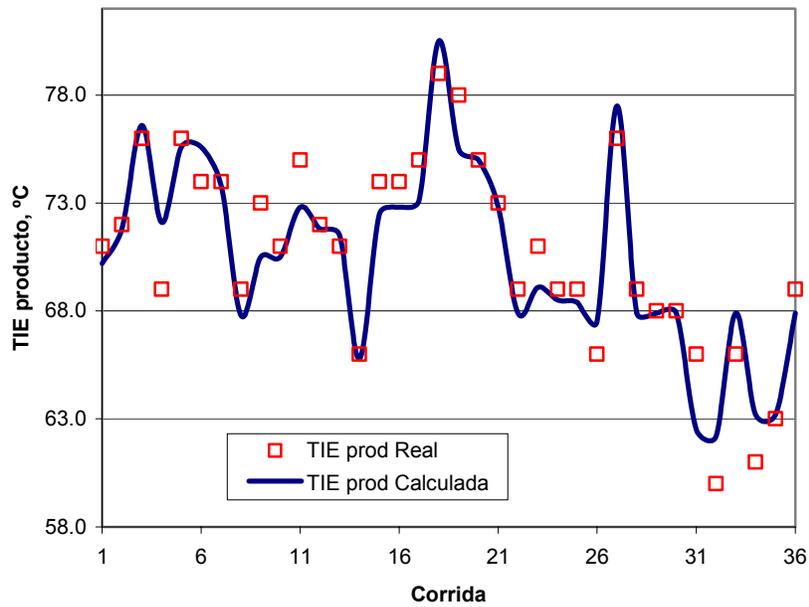


Figura 5.3 Comparación de datos de TIE producto. Caso Prueba.

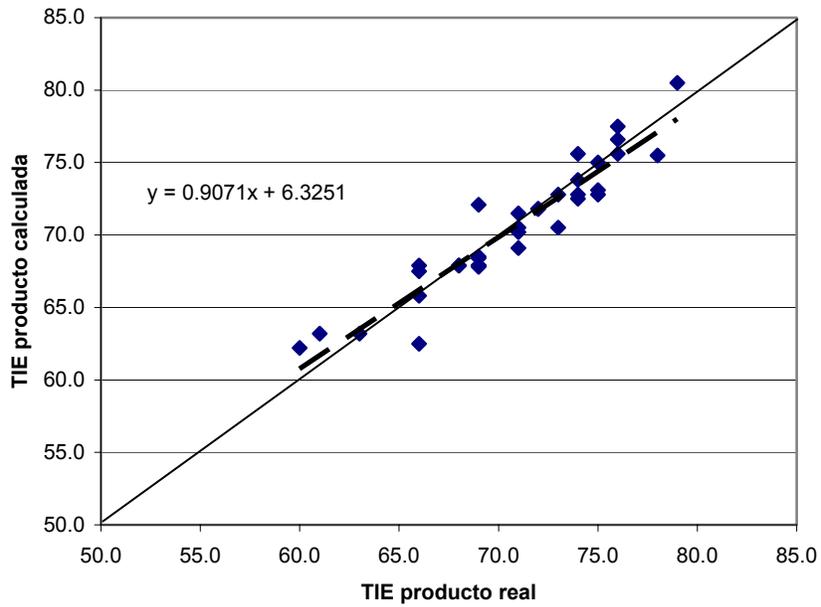


Figura 5.4 Dispersión de datos de TIE producto. Caso Prueba.

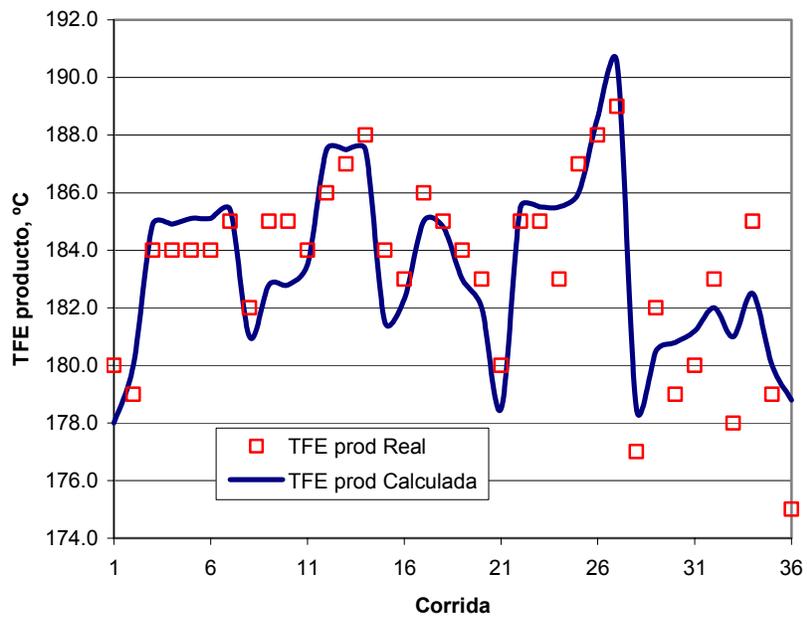


Figura 5.5 Comparación de datos de TFE producto. Caso Prueba.

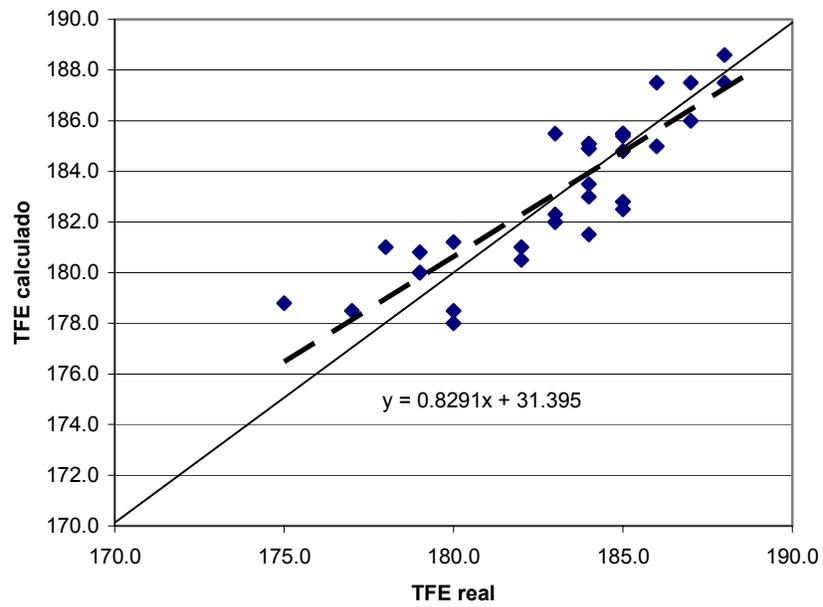


Figura 5.6 Dispersión de datos de TFE producto. Caso Prueba.

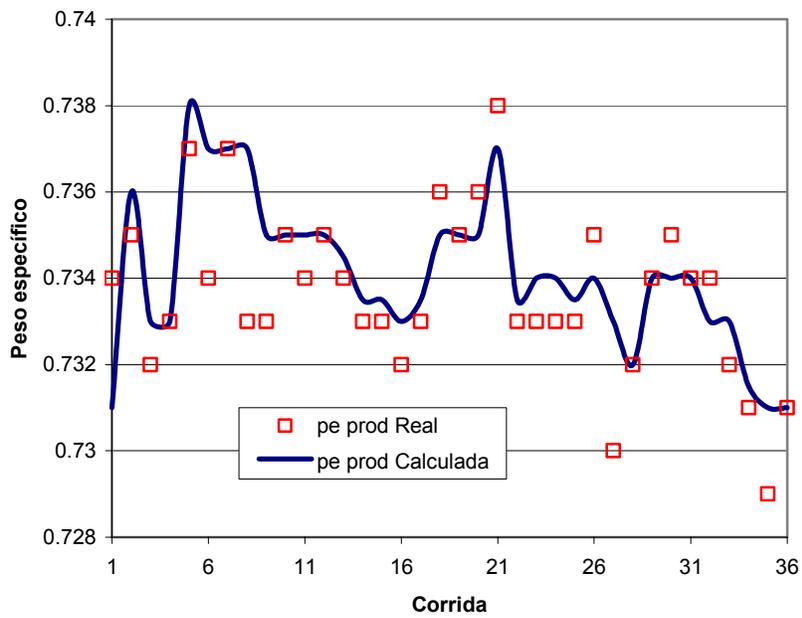


Figura 5.7 Comparación de datos de Peso específico producto. Caso Prueba.

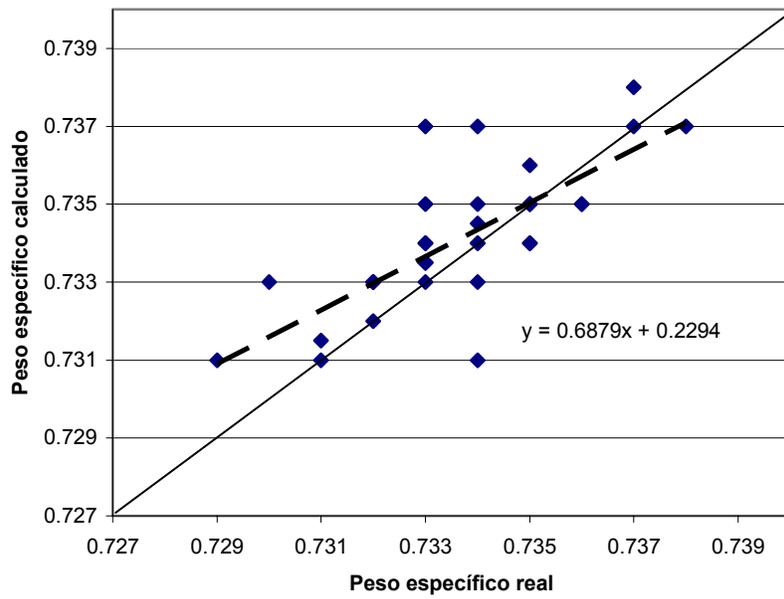


Figura 5.8 Dispersión de datos de Peso específico producto. Caso Prueba.

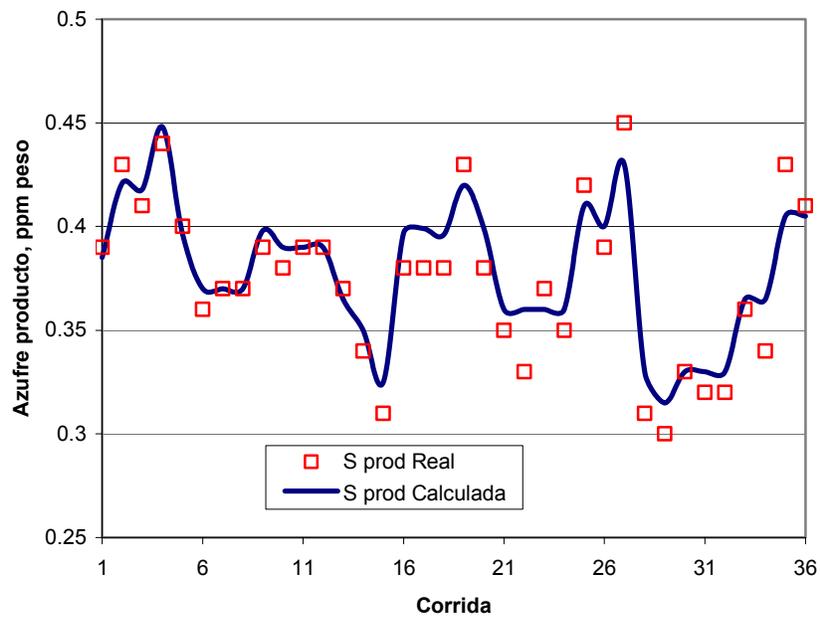


Figura 5.9 Comparación de datos de Azufre en el producto. Caso Prueba.

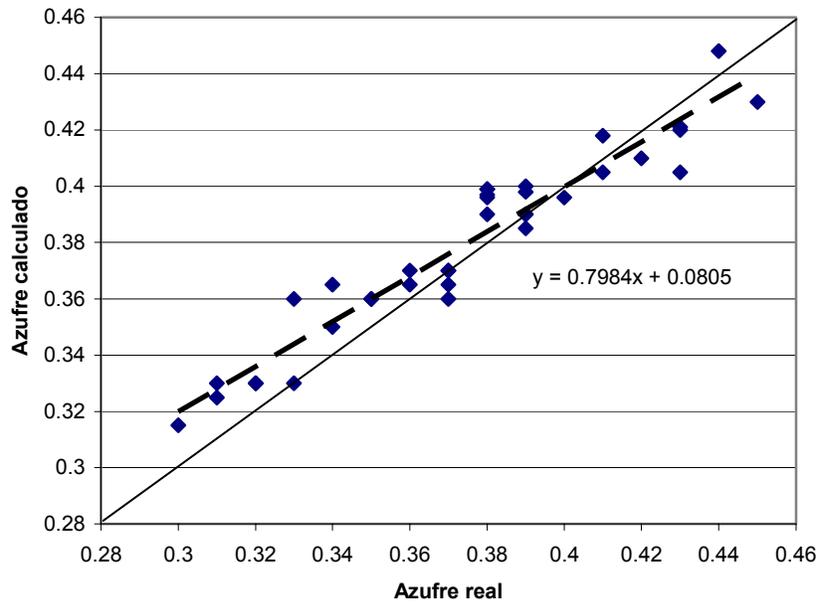


Figura 5.10 Dispersión de datos de Azufre en el producto. Caso Prueba.

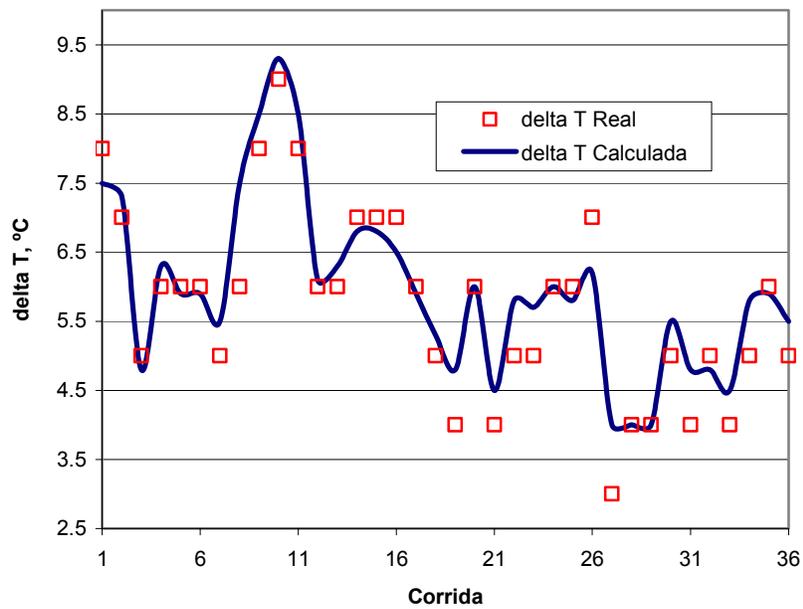


Figura 5.11 Comparación de datos de  $\Delta T$ . Caso Prueba.

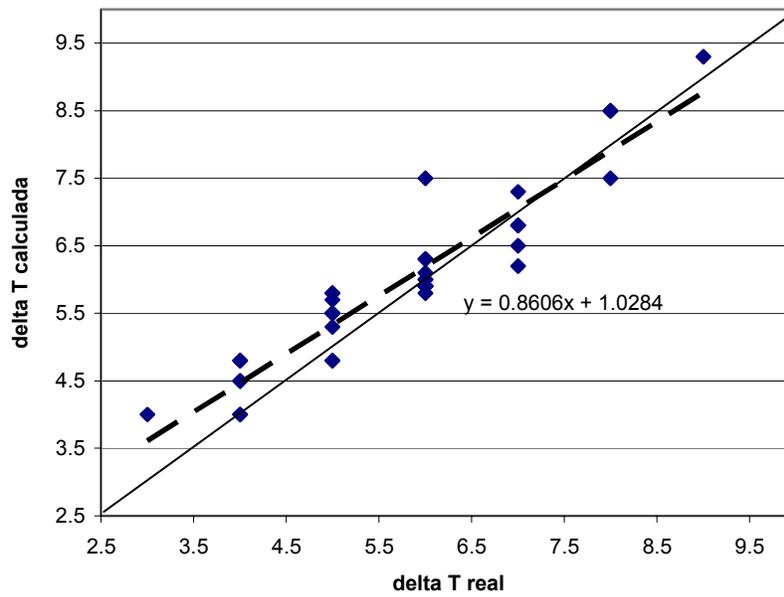


Figura 5.12 Dispersión de datos de  $\Delta T$ . Caso Prueba.

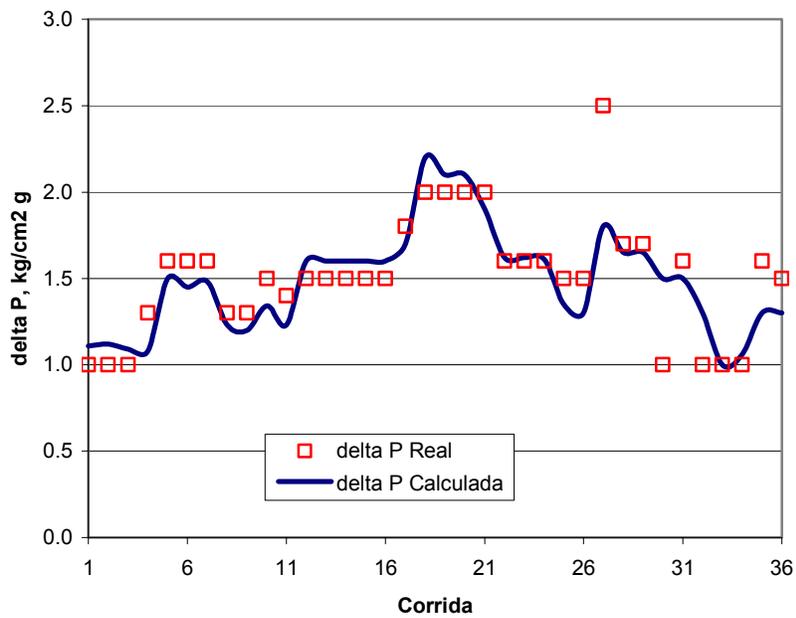


Figura 5.13 Comparación de datos de  $\Delta P$ . Caso Prueba.

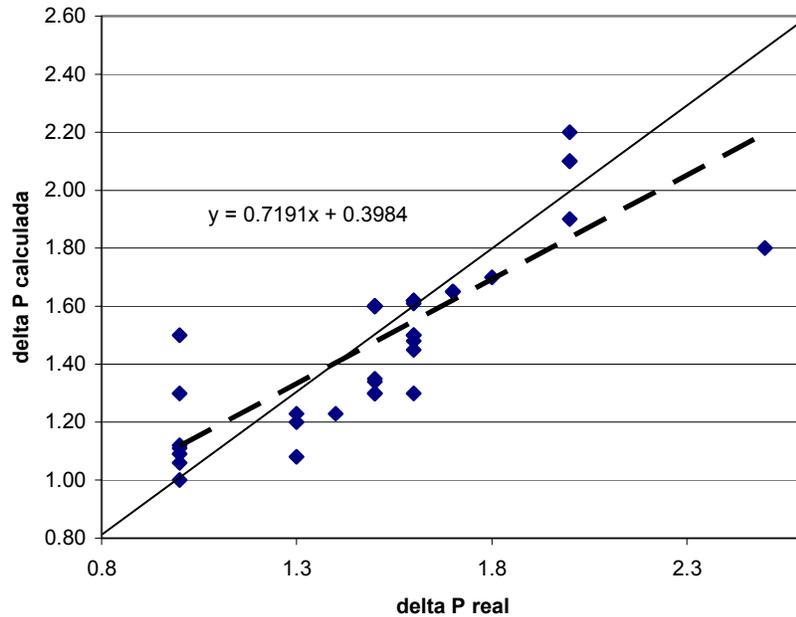


Figura 5.14 Dispersión de datos de  $\Delta P$ . Caso Prueba.

Los errores relativos promedio obtenidos de las predicciones hechas por el modelo de Redes Neuronales se muestran en la siguiente tabla.

Tabla 5.1 Por ciento de error relativo promedio de predicción. Caso Prueba

VARIABLE DE SALIDA	% ERROR
Producción	0.669
Temperatura Inicial de Ebullición (TIE)	1.685
Temperatura Final de Ebullición (TFE)	0.726
Peso específico	0.131
Azufre	2.948
$\Delta T$	7.935
$\Delta P$	9.903

A pesar de que los datos utilizados en la fase de prueba del modelo de RNA para el Proceso de Hidrodesulfuración no habían sido utilizados anteriormente, es decir la RNA nunca los había “visto”; en la tabla 5.1 podemos observar que el modelo predice adecuadamente la mayoría de las variables de salida con errores promedio entre 0.6 y 3.0 %; excepto aquellos para los valores de las deltas de temperatura y presión los cuales muestran un error promedio del 7.9 y 9.9 %, respectivamente. Como ya se había mencionado anteriormente, en el análisis de los datos “crudos” se pudo ver que dichas variables no mostraban prácticamente ninguna correlación, ni aún después del análisis

estadístico, lo cual provoca que tampoco la RNA sea capaz de desarrollar un modelo adecuado para esas variables.

Adicionalmente, en las figuras 5.1 a 5.14 se puede ver que la tendencia de predicción del modelo RNA es buena, sin embargo debido al comportamiento errático de los datos, en algunos casos el modelo no logra predecir correctamente.

Es necesario aclarar que se probaron varios tipos de topologías, en particular una que contenía dos capas intermedias, con la idea de que una topología que contenga más de una capa intermedia puede proporcionar mejores resultados de predicción. Sin embargo, de las predicciones obtenidas se pudo observar, al menos para este caso, que no proporciona ninguna mejora apreciable. Adicionalmente, las topologías más complejas requieren una mayor cantidad de datos para su entrenamiento y su convergencia es más complicada, si es que se alcanza. Por tal motivo, se analizaron topologías con una sola capa oculta.

Asimismo a pesar de que existen varias funciones de transferencia, entre las que encontramos: *tanh*, *gaussiana*, *sech*, *etc*; en este trabajo sólo se investigó la función sigmoideal, ya que es la más ampliamente usada en el desarrollo de modelos RNA aplicados a los procesos químicos (Michalopoulos, et. al., 2001; Chouai, et. al., 2000; Gontarsky, et. al., 2000; Ramasamy, et. al., 1995)..

Adicionalmente al análisis estadístico que normalmente se realiza a los datos, es necesario efectuar lo que en la literatura de las RNA se denomina “Análisis de Componentes Principales” (PCA por sus siglas en inglés) a fin de determinar cuales de las variables de entrada del proceso analizado tienen mayor efecto sobre las variables de salida, y de esta forma eliminar aquellas que no tienen un efecto considerable. Esto reduce la dimensionalidad del problema, lo que a su vez hace que se obtenga una mejor convergencia y que además se requiera menos número de datos, lo cual es importante para procesos donde no se tiene una gran cantidad de estos.

Para este trabajo, no se realizó el PCA debido a las siguientes razones:

-  La cantidad de datos disponible es tal que no hace necesario eliminar variables para mejorar la convergencia, o reducir la dimensionalidad del problema.
  
-  Las variables seleccionadas son las que típicamente están estrechamente relacionadas entre sí, por lo que no se requiere de un análisis adicional para determinar esta interrelación.

## 5.2. COMPARACIÓN CON UN MODELO CINÉTICO.

Tradicionalmente los modelos utilizados en la simulación de sistemas que incluyen reacciones químicas han sido del tipo cinético, y para el caso de la Hidrodesulfuración no es la excepción. Recientemente, De la Paz, Z.C, et. al., 1999; desarrollaron una herramienta denominada *Hidro-Naf*, la cual es un simulador del Proceso de Hidrodesulfuración de Naftas, y que de acuerdo a sus propias palabras se define como sigue:

“El simulador *Hidro-Naf* es un simulador cinético que permite evaluar el comportamiento de la sección de reacción de los proceso de hidrodesulfuración de naftas. Por su naturaleza es adaptable a cualquier planta de este tipo, conociendo su estadística operacional.”

A través de esta herramienta es factible determinar los efectos de alimentar cargas con diferentes características y de distinta calidad, procesadas bajo distintos niveles de severidad, pudiéndose determinar el Balance Global de Materia, el porcentaje de Hidrodesulfuración, el Incremento de Temperatura, la Caída de Presión, el Consumo de Hidrógeno, y el Contenido de Azufre en el Producto.”

Este simulador cuenta adicionalmente con correlaciones para estimar el efecto de la contaminación por agua sobre el catalizador, depósito de carbón, y la caída de presión por ensuciamiento. Cabe aclarar que el modelo RNA desarrollado en esta tesis, no incluye el análisis de la degradación del catalizador.

Se escogió el simulador *Hidro-Naf* ya que fue aplicado a la misma planta que se tomó como base para el desarrollo de este trabajo: la UPH-600 de la Refinería de Cd. Madero, Tamaulipas.

La información estadística de operación utilizada en la aplicación de simulador *Hidro-Naf* a la unidad UPH-600 es similar a la usada en el desarrollo del modelo *NeuroNaf*, tanto en condiciones de operación como en el tipo de catalizador.

Los resultados obtenidos por *Hidro-Naf* se muestran en la tabla 5.2, y son estos mismos datos los que se analizan usando el modelo *NeuroNaf*.

Tabla 5.2 Resultados obtenidos por el simulador cinético *Hidro-Naf*

Variable	Operación	Simulación	Operación	Simulación	Operación	Simulación
<b>CONDICIONES DE OPERACIÓN</b>						
Corrida	1		2		3	
T, °C	291.0	291.0	292.0	292.0	292.0	292.0
<b>ΔT</b>	<b>8.00</b>	<b>6.05</b>	<b>5.00</b>	<b>5.50</b>	<b>5.00</b>	<b>4.69</b>
P, MPa g	5.3	5.3	5.5	5.5	5.5	5.5
<b>ΔP</b>	<b>0.22</b>	<b>0.08</b>	<b>0.23</b>	<b>0.10</b>	<b>0.24</b>	<b>0.10</b>
Rel H <sub>2</sub> /HC	35.4	35.4	35.1	35.1	34.5	34.5
<b>CARGA</b>						
Flujo, BPD	20039	20039	20654	20654	20517	20517
Peso esp.	0.732	0.732	0.730	0.730	0.731	0.731
Azufre, ppm	1248	1248	1012	1012	1081	1081
<b>PRODUCTO</b>						
<b>Azufre, ppm</b>	<b>0.27</b>	<b>0.30</b>	<b>0.25</b>	<b>0.29</b>	<b>0.39</b>	<b>0.40</b>
<b>CONDICIONES DE OPERACIÓN</b>						
Corrida	4		5		6	
T, °C	292.0	292.0	292.0	292.0	292.0	292.0
<b>ΔT</b>	<b>6.00</b>	<b>6.04</b>	<b>7.00</b>	<b>6.70</b>	<b>6.00</b>	<b>6.64</b>
P, MPa g	5.6	5.6	5.6	5.6	5.6	5.6
<b>ΔP</b>	<b>0.25</b>	<b>0.10</b>	<b>0.25</b>	<b>0.10</b>	<b>0.29</b>	<b>0.09</b>
Rel H <sub>2</sub> /HC	36.9	36.9	35.3	35.3	36.3	36.3
<b>CARGA</b>						
Flujo, BPD	20964	20964	21041	21041	21704	21704
Peso esp.	0.734	0.734	0.735	0.735	0.736	0.736
Azufre, ppm	989	989	1372	1372	1412	1412
<b>PRODUCTO</b>						
<b>Azufre, ppm</b>	<b>0.29</b>	<b>0.30</b>	<b>0.27</b>	<b>0.30</b>	<b>0.30</b>	<b>0.31</b>

Las variables comparables entre los dos modelos se distinguen en la tabla anterior por encontrarse resaltados en negritas. Los resultados obtenidos aplicando el modelo *NeuroNaf*, usando las mismas condiciones de operación y el mismo catalizador, se muestran en la tabla 5.3.

La comparación de los resultados obtenidos por los dos simuladores se muestran en la tabla 5.4, donde se puede observar que el modelo *NeuroNaf* a pesar de no ser un modelo riguroso proporciona resultados equivalentes al modelo cinético *Hidro-Naf*; y para la caída de presión y la  $\Delta T$  inclusive reporta mejores resultados.

Por supuesto es claro que el modelo cinético tienen algunas ventajas como incluir expresiones para estimar la vida útil del catalizador, el efecto del agua sobre el mismo, y la deposición de carbón; los cuales no fueron considerados en el presente trabajo debido

a no contar con datos estadísticos de estos parámetros. Sin embargo en el momento que estén disponibles se pueden integrar al modelo *NeuroNaf*.

Tabla 5.3 Resultados obtenidos por el simulador *NeuroNaf*

Variable	Operación	Simulación	Operación	Simulación	Operación	Simulación
<b>CONDICIONES DE OPERACIÓN</b>						
Corrida	1		2		3	
$\Delta T$	8.00	7.65	5.00	5.2	5.00	5.15
$\Delta P$	0.22	0.20	0.23	0.20	0.24	0.20
<b>PRODUCTO</b>						
Azufre, ppm	0.27	0.32	0.25	0.30	0.39	0.41
<b>CONDICIONES DE OPERACIÓN</b>						
Corrida	4		5		6	
$\Delta T$	6.00	5.58	7.00	6.80	6.00	6.30
$\Delta P$	0.25	0.23	0.25	0.22	0.29	0.23
<b>PRODUCTO</b>						
Azufre, ppm	0.29	0.31	0.27	0.29	0.30	0.32

Tabla 5.4 Comparación entre simuladores.

Corrida	$\Delta T$ , °C			% error relativo	
	Operación	HidroNaf	NeuroNaf	HidroNaf	NeuroNaf
1	8.00	6.05	7.65	24.38	4.38
2	5.00	5.50	5.20	-10.00	-4.00
3	5.00	4.69	5.15	6.20	-3.00
4	6.00	6.04	5.58	-0.67	7.00
5	7.00	6.70	6.80	4.29	2.86
6	6.00	6.64	6.30	-10.67	-5.00
% error relativo promedio				2.25	0.37
Corrida	$\Delta P$ , MPa man			% error relativo	
	Operación	HidroNaf	NeuroNaf	HidroNaf	NeuroNaf
1	0.22	0.08	0.20	63.64	9.09
2	0.23	0.10	0.20	56.52	13.04
3	0.24	0.10	0.20	58.33	16.67
4	0.25	0.10	0.23	60.00	8.00
5	0.25	0.10	0.22	60.00	12.00
6	0.29	0.09	0.23	68.97	20.69
% error relativo promedio				61.24	13.25
Corrida	Azufre en el producto, ppm			% error relativo	
	Operación	HidroNaf	NeuroNaf	HidroNaf	NeuroNaf
1	0.27	0.30	0.32	-11.11	-18.52
2	0.25	0.29	0.27	-16.00	-8.00
3	0.39	0.40	0.40	-2.56	-2.56
4	0.29	0.30	0.30	-3.45	-3.45
5	0.27	0.30	0.29	-11.11	-7.41
6	0.30	0.31	0.31	-3.33	-3.33
% error relativo promedio				-7.93	-7.21

Analizando los resultados presentados en la tabla 5.4 se puede pensar que el modelo *NeuroNaf* es superior al modelo cinético, sin embargo es conveniente aclarar algunos puntos al respecto.

El modelo cinético ha sido desarrollado como un modelo general del Proceso de Hidrodesulfuración de Naftas, es decir que es aplicable a cualquier planta de hidrodesulfuración con el sólo requisito de contar con la estadística de operación (De la Paz, et. al., 1999); lo cual hace evidente que requiere efectuar algunas regresiones a fin de ajustar los parámetros del modelo, y esto lo tiene que realizar para cada planta analizada. A fin de proporcionar una adecuada predicción, el modelo cinético requiere contar con datos de una corrida estable, y que representen el comportamiento típico de la planta, lo que en caso de no contar con datos adecuados la capacidad predictiva del modelo puede disminuir considerablemente.

Por otra parte, para el desarrollo del modelo *NeuroNaf* se aseguró desde el inicio del procedimiento que se tuvieran los mejores datos disponibles, y por tal motivo el modelo da mejores resultados; por supuesto también se debe tomar en cuenta que el modelo *NeuroNaf* ha sido desarrollado específicamente para la Planta Hidrodesulfuradora de Naftas UPH-600, mientras que el modelo cinético es de aplicación general.

Estas mismas características hacen que el modelo cinético presente mejor capacidad de predicción en intervalos fuera de los cuales fue creado, es decir que es capaz de extrapolar; mientras que el modelo de RNA está limitado al respecto (Ramasamy, 1995), ya que sólo es aplicable para la planta de la cual se obtuvieron los datos y en los intervalos de las variables con que fue entrenado; sin embargo es capaz de ampliar su intervalo de aplicación integrándole los nuevos datos de interés.

Adicionalmente, el modelo cinético es capaz de evaluar el comportamiento de la planta para diferentes cargas; y por otro lado para aplicar el modelo *NeuroNaf* de esta forma, es necesario contar con todos los datos de las variables del proceso para las diferentes cargas. Lo mismo sucede para el caso en que se requiera analizar el impacto de usar varios catalizadores.



## **CONCLUSIONES**

1. Las RNA son capaces de modelar procesos para los cuales no se tiene una idea precisa de como se llevan a cabo, y que además cuentan con una buena cantidad de datos experimentales aun cuando esta información esté incompleta o viciada, como en el caso de este trabajo para el Proceso de Hidrodesulfuración de Naftas.
2. A pesar de que no existe una metodología sistemática a través de la cual se pueda construir un modelo de RNA, el procedimiento propuesto en este trabajo ofrece una buena alternativa para este propósito, ya que no contiene reglas heurísticas sino que está basado en criterios estadísticos. Sin embargo, existen áreas donde aún se está realizando investigación a fin de establecer reglas claras, como en el caso de los intervalos iniciales de pesos y el número óptimo de nodos en la capa intermedia.
3. Se estableció la importancia de contar con suficientes datos experimentales, lo cual está íntimamente relacionado con el tamaño de la RNA propuesto, es decir entre más grande la RNA, más pesos involucrados en su estructura y por lo tanto se requiere mayor cantidad de datos. Asimismo, también es importante contar cierta calidad de los datos, ya que aun cuando las RNA son capaces de manejar datos con “ruido”, se deben eliminar juegos de datos que están demasiado alejados de ciertos intervalos, este criterio de eliminación lo proporciona un análisis estadístico de los mismos.
4. Se analizaron tres topologías: Global, Proceso, y Propiedades. Las dos últimas topologías tienen menos nodos, pero en conjunto cubren todas las variables incluidas en la topología Global. Para este proceso en particular, el hecho de dividir la topología Global en dos más pequeñas, mejora considerablemente la capacidad predictiva del la RNA, debido a que se reduce la dimensionalidad de cada una de las topologías pequeñas.
5. El modelo RNA se evaluó de dos formas: 1) Presentando datos del conjunto de prueba que no habían sido mostrados anteriormente; 2) Comparándolo con un modelo cinético. La primera opción permite observar que el modelo predice adecuadamente la mayoría de las variables de salida, con errores promedio entre 0.6 y 3.0%, excepto para los valores de  $\Delta T$  y  $\Delta P$  para las cuales da un error promedio de 7.9 y 9.9% respectivamente. La comparación con el modelo cinético muestra que el modelo RNA predice valores similares de contenido de azufre en la nafta producto, mientras que para las  $\Delta T$  y  $\Delta P$  proporciona resultados superiores.

6. Queda establecida la capacidad de los modelos de RNA para representar el comportamiento de un Reactor de Hidrodesulfuración de Naftas, ya que predice la mayoría de las variables de salida con una precisión similar a la proporcionada por un modelo cinético y en algunas otras se obtienen mejores resultados.

## **GLOSARIO**

RNA	Red Neuronal Artificial.- es un sistema computacional que está constituido por cierto número de elementos simples, altamente interconectados, los cuales procesan información por medio de un estado dinámico de respuesta a los estímulos externos.
Topología	Estructura de la Red Neuronal Artificial, a determinar como parte de la construcción del modelo.
Neurona o nodo	Cada uno de los elementos de procesamiento de la Red Neuronal Artificial.
Capa	Cada uno de los vectores que forman la estructura de la Red Neuronal Artificial.
Cada de entrada	Vector de la topología formado por todas y cada una de las variables de entrada de la Red Neuronal Artificial.
Cada oculta	Vector intermedio de la Red Neuronal Artificial, cuyo número de neuronas será determinado como resultado de la construcción del modelo.
Capa de salida	Vector de la topología formado por todas y cada una de las variables de salida de la Red Neuronal Artificial.
Pesos	Factores de ponderación que correlacionan las variables de entrada con las variables de salida, se determinan como parte del modelo.
Retropropagación	Algoritmo de entrenamiento de la Red Neuronal Artificial.
Entrenamiento	Determinación de la topología que mejor reproduce los datos.
Validación	Determinación de los pesos que mejor reproducen los datos.
Prueba	Presentación de datos no usados para entrenamiento ni validación, para verificar la capacidad predictiva del modelo desarrollado.

Época	Iteración completa durante la convergencia del algoritmo de entrenamiento, validación y prueba.
Función sigmoideal	Función de transferencia de la RNA la cual permite introducir la no-linealidad del modelo
Sesgo	Valor aditivo del algoritmo de retropropagación que evita que la convergencia pase por el origen de la superficie de error.
Momentum	Factor del algoritmo de retro-propagación que permite ajustar el tamaño de paso de la convergencia.
Validación cruzada	Técnica que permite determinar el mejor momento para detener el entrenamiento a fin de evitar que la Red Neuronal Artificial se sobre-entrene.

## **BIBLIOGRAFÍA**

- Abilov, A.; Zeybek, Z. "Use of neural network for modeling of non-linear process integration technology in chemical engineering", *Chem. Eng. Proc.*, **39**, 449, (2000).
- Aguilera, L.M.; Sánchez, M.J.F.; Salazar, S.D.; Maya, Y.R.; González, S.A. "Modelos Estadísticos versus Redes Neuronales Artificiales en el Modelado de Unidades Industriales de Craqueo Catalítico Fluido (FCC)". *Premio IMIQ 2003 'Ing. César O. Baptista'*. XLIII Convención Nacional IMIQ. Villahermosa, Tab. (2003).
- Anders, U.; Korn, O. "Model Selection in Neural Networks". *Neural Networks*, **12**(2), 309, (1999).
- Antsaklis, P.J. "Neural Networks in Control Systems". *IEEE Control Systems Magazine*, **3**, (1990).
- Arce, M.E. "Redes Neuronales para el Control de Procesos". *Correspondencia Particular* (1994).
- Ayyagari, R. "Kolmogorov's Superposition Theorem and its Relationship with Neural Networks". (2003).
- Baghat, P. "An Introduction to Neural Nets". *Chem. Eng. Prog.*, **55**, (1990).
- Bailey, D.; Thompson, D. "How to develop Neural-Networks Applications". *AI Expert*, (1990).
- Bath, N.; McAvoy, T.J. "Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems", *Comp. Chem. Eng.*, **14**(4/5), 573, (1990).
- Bath, N.V.; Minderman, P.A. Jr.; McAvoy, T.; Wang, N.S. "Modeling Chemical Process Systems via Neural Computation". *IEEE Control Systems Magazine*, **10**(24), (1990).
- Bishop, C.M. "Neural Networks for Pattern Recognition". *New York: Oxford University Press, Inc.*, (1998).
- Bodor, N.; Harget, A.; Huang, M. "Neural Network studies. 1. Estimation of the Aqueous Solubility of Organic Compounds", *J. Am. Chem. Soc.* **113**, 9480, (1991).
- Bravo, M.M.; Campos, P.B. "Sistema de cálculo para predicción de rendimientos y propiedades de productos en una unidad reductora de viscosidad", *Revista del IMP*,

**XVI**(3), 76, (1984).

Bravo, M.M. "Sistema de cálculo para la estimación de rendimientos y propiedades de productos de las unidades desintegradoras catalíticas tipo fluidas", *Revista del IMP*, **XVII**(1), 18, (1985).

Bulsari, A.B., ed. "Neural Networks for Chemical Engineers". *Elsevier, Amsterdam, The Netherlands*, (1995).

Carrillo Villa, M.A.; Flores Sánchez P.; Bravo Medina, M.A. "Modelo de Predicción de Rendimientos y propiedades de Productos en Unidades Hidrodesulfuradoras Catalíticas". *Revista de IMP*, **XIX**(1), (1987).

Chitra, S.P. "Use Neural Networks for Problem Solving". *Chem. Eng. Prog.*, 44, (1993).

Chouai, A.; Cabassud, M.; Le Lann, M.V.; Gourdon, C.; Casamatta, G. "Use of neural networks for liquid-liquid extraction column modeling: an experimental study", *Chem. Eng. Proc.*, **39**, 171, (2000).

Chu, S.R.; Shoureshi, R.; Tenorio, M. "Neural Networks for System Identification", *IEEE, American Control Conference, Pittsburgh, Pennsylvania*, (1989).

Cooper, D.J.; Megan, L.; Hinde, R.F. Jr. "Comparing two Neural Networks for Pattern Based Adaptive Process Control", *AIChE Journal*, **38**(1), (1992).

De la Paz, Z.C.; Vega, M.P.; Martínez, O.V.; Rodríguez, R. J.E.; Burgos, V.E.; Cruz, M.E.; Leyva, N.M.A. "HIDRO-NAF Simulador del proceso de Hidrodesulfuración de Naftas", *Revista del IMIQ*, **XL**(5-6), (1999).

Epping, W.J.; Nitters, G. "A Neural Network for Analysis and Improvement of Gas Well Production". *Proc. Of Summer Computer Simulation Conference*, 791, (1990).

Fishwick, P.A. "Neural Networks Models in Simulation: A Comparison with Traditional Modeling Approaches". *Proceedings of the 1989 Winter Simulation Conference*, 702, (1989).

Fullana, M.; Trabelsi, F.; Recasens, F. "Use of neural net computing for statistical and kinetic modeling and simulation of supercritical fluid extractors", *Chem. Eng. Sci.*, **55**, 79, (2000).

Garson, G.D. "Interpreting Neural-Network Connection Weights". *AI Expert*, 47, (1991).

Gontarski, C.A.; Rodrigues, P.R.; Mori, M.; Prenem, L.F. "Simulation of an Industrial Wastewater treatment Plant using Artificial Neural Networks", *Comp. Chem. Eng.*, **24**, 1719, (2000).

Haykin, S., "Neural Networks: A Comprehensive Foundation", *Prentice-Hall*, (1998).

Hecht-Nielsen, R. "Neurocomputing", *Addison-Wesley*, (1990).

Hinton, G.E. "Redes neuronales que aprenden de la experiencia". *Investigación y Ciencia*. (1992).

Hoskins, J.C.; Himmelblau, D.M. "Artificial Neural Network Models of Knowledge Representation in Chemical Engineering". *Comp. Chem. Eng.*, **12**(9/10), 881, (1988).

Hoskins, J.C.; Himmelblau, D.M. "Process Control via Artificial Neural Networks and Reinforcement Learning". *Comp. Chem. Eng.* **16**(4), 241, (1992).

Hussain, M.A.; Kershbaum, L.S. "Implementation of an Inverse-Model-Based Control Strategy using Neural Networks on a Partially Simulated Exothermic Reactor", *Trans Écheme.*, **78**(A), (2000).

Instituto Mexicano del Petróleo. "Manual de Operación de Catalizadores IMP-DSD para Hidrotratamiento", (1994).

Jia, L.; Yu, J. "Novel neural fuzzy network for product quality monitoring", *Hydroc. Proc.*, (2001).

Jones, W.P.; Hoskins, J. "Back-Propagation, a generalized delta learning rule. Heuristic algorithms", *Byte*, (1987).

Kavchak, M.; Budman, H. "Adaptive neural network structures for non-linear process estimation and control", *Comp. Chem. Eng.* **23**, 1209, (1999).

Korn, G.A. "A New Environment for Interactive Neural Network Experiments". *Neural Networks*, **2**, 229, (1989).

Kramer, M.A. "Autoassociative Neural Networks". *Comp. Chem. Eng.*, **16**(4), 313, (1992).

Kramer, M.A. "Nonlinear Principal Component Analysis Using Autoassociative Neural Networks". *AIChE Journal*, **37**(2), (1991).

Kung, S.Y.; Hwang, J.N. "An Algebraic Projection Analysis for Optimal Hidden Units Size

- and Learning Rates in Back-Propagation Learning". *Proceedings of IEEE International Neural Nets Conference, II*, I-363, (1988).
- Kuttisupakorn, P.; Hussain, M.A.; Petcherdask, J. "Studies on the Use of Neural Networks in Nonlinear Control Strategies", *J. Chem. Eng. Japan*, **34**(4), 453, (2001).
- Lennox, B.; Rutheford, P.; Montague, G.A.; Haughin, C. "Case study investigating the applications of neural networks for process modeling and condition monitoring", *Comp. Chem. Eng.*, **22**(11), 1573, (1998).
- Leonard, J.; Kramer, M.A. "Improvement of the Back-Propagation Algorithm for Training Neural Networks". *Comp. Chem. Eng.*, **14**(3), 337, (1990).
- Lipman, R. "An Introduction to Computing with Neural Networks", *IEEE ASSP Magazine*, **3**, 4, (1987).
- Martins, F.G.; Coelho, M.A.N. "Application of feedforward artificial neural networks to improve process control of PID-based control algorithms", *Comp. Chem. Eng.*, **24**, 853, (2000).
- Mavrovouniotis, M.L.; Chang, S. "Hierarchical Neural Networks". *Comp. Chem. Eng.*, **16**(4), 347, (1992).
- McGreavy, C.; Lu, M.L.; Wang, X.Z.; Kam, E.K.T. "Characterisation of the Behavior and Product Distribution in Fluid Catalytic Cracking Using Neural Networks". *Chem. Eng. Sci.*, **49**(24A), 4717, (1994).
- McGregor, J.F. "Statistical Process Monitoring and Quality Control of Multivariable Processes", *Proceedings of ADCHEM'94-IFAC Symposium of Advanced Control of Chemical Processes*, (1994).
- Michalopoulos, J.; Papadokonstadakis, S.; Arampatzis, G.; Lygeros, A. "Modelling of an Industrial Fluid Catalytic Cracking Unit Using Neural Networks". *Trans. I. Chem. E.*, **79**(A), (2001).
- Minai, A.A.; Williams, R.D. "Acceleration of Back-Propagation through Learning Rate and Momentum Adaptation". *International Joint Conference on Neural Networks*, **I**, 676, (1990).
- Molga, E.; Cherbanski, R. "Hybrid first-principle-neural network approach to modeling of the liquid-liquid reacting system", *Chem. Eng. Sci.*, **54**, 2467, (1999).

- Molga, E.J.; van Woezik, B.A.A.; Westerterp, K.R. "Neural networks for modeling of chemical reaction system with complex kinetics: oxidation of 2-octanol with nitric acid", *Comp. Eng. Proc.*, **39**, 323, (2000).
- Morris, A.J.; Montague, G.A.; y Willis, M.J. "Artificial Neural Networks: Studies in Process Modelling and Control". *Trans. I. Chem. E.*, **72(A)**, 3, (1994).
- Morrison, J.D. "A 'neural' network model that supports realtime learning of temporal relationships in complex engineering domains". *Simulation*, **59(3)**, 152, (1992).
- Nacimiento, C.A.O.; Giudici, R.; Scherbakoff, N. "Modeling of Industrial Nylon-6,6 Polymerization Process in a Twin-Screw Extruder Reactor. II. Neural Networks and Irbid Models", *J. Appl. Poly. Sci.*, **72**, 905, (1999).
- Nares, O.H.R.; Moreno, T.A.; Bermúdez, M.O. "Introducción al Hidrotratamiento". *Primera Semana Internacional de Reformación Catalítica e Hidrotratamiento*, (1994).
- Palomares, D.; Limón, M.; Méndez, F.J.; Villavicencio, A. "Redes Neuronales Aplicadas a Identificación de Sistemas". *Avances en Ingeniería química*, 273, (1991).
- Park, D.C.; El-Sharkawi, M.A.; Marks, R.J. II. "An Adaptively Trained Neural Network". *IEEE Transsactions on Neural Networks*, **2(3)**, (1991).
- Petersen, R.; Fredenslund, A.; Rasmussen, P. "Artificial Neural Networks as a Predictive Tool for Vapor.Liquid Equilibrium". *Comp. Chem. Eng.*, **18**, s63, (1994).
- Pollard, J.F.; Broussard, M.R.; Garrison, D.B.; San, K.Y. "Process Identification using Neural Networks". *Comp. Chem. Eng.*, **16(4)**, 253, (1992).
- Psaltis, D.; Sideris, A.; Yamamura, A.A. "A Multilayered Neural Network Controler". *IEEE Control System Magazine*, 17, (1988).
- Psichogios, D.C.; Ungar, L.H. "Direct and Indirect Model Based Control Using Artificial Neural Networks", *Ind. Eng. Chem. Res.*, **30**, 2564, (1991).
- Psichogios, D.C.; Ungar, L.H. "A Hybrid Neural Network-First Principles Approach to Process Modeling", *AIChE Journal.*, **38(10)**, (1992).
- Qiang, N.; Sejnowski, T.J. "Predicting the Secondary Structure of Globular Proteins Using Neural Networks Models". *J. Mol. Biol.*, **202**, 865, (1988).
- Ramasamy, S.; Desphande; P.B., Paxton; G.E. and Hajare, R.P., "Consider Neural

- Networks for Process Identification”, *Hydroc. Proc.*, 59, (june, 1995).
- Ripley, B.D. “Pattern Recognition and Neural Networks”, *Cambridge University Press*, (1996).
- Rodríguez Otal, L.M. “El hidrotratamiento y la Reformación de Naftas ante el reto Ecológico”. *Primera Semana Internacional de Reformación Catalítica e Hidrotratamiento*. (1994).
- Rumelhart, D.E.; Mc Clelland, J.L., “Learning Internal representations by Error Propagation”. *Parallel Distributed Processing, Chap. 8.*, (1986).
- Sabharwal, A.; Bath, N.V.; Wada, T. “Integrate empirical and physical modeling”. *Hydroc. Proc.*, (1997).
- Sablani, S.S. “A neural network approach for non-iterative calculation of heat transfer coefficient in fluid-particle systems”, *Comp. Eng. Proc.*, **40**, 363, (2001).
- Sarle, W.S., ed. “Neural Network FAQ, part 1 of 7: Introduction”. *Periodic posting the Usenet newsgroup comp.ai.neural-nets*, URL: <ftp://ftp.sas.com/pub/neural/FAQ.html>, (1999).
- Savkovic-Stevanovic , J. “Neural Networks for Process Analysis and Optimization: Modeling and Applications”. *Comp. Chem. Eng.*, **18**(11/12), 1149, (1994).
- Savkovic-Stevanovic, J., ”A Neural Network Model for Analysis and Optimization of Processes”. *Comp. Chem. Eng.*, **17**, s411, (1993).
- Savkovic-Stevanovic, J.; Vico-Stevanovic, M. “An Artificial Neural Networks for Variables Estimation and State Identification in Process Fermentation”. *Proced. Of Summer Computer Simulation Conference, Boston, Ma.*, 307, (1993).
- Shriver, B.D. “Artificial Neural Networks”. *Computer*, 8, (1988).
- Sietsma, J.; Dow,R.J.F. “Creating Artificial Neural Networks that Generalize”. *Neural Networks*, **4**, 67, (1991).
- Tan, S.; Mavrovouniotis, M.L. “Reducing Data Dimensionality through Optimizing Neural Networks Inputs”. *AIChE Journal*, 41(6), (1995).
- Tholudur, A.; Ramirez, W.F.; McMillan, J.D. “Interpolated parameter functions for neural

- networks models”, *Comp. Chem. Eng.*, **24**, 2545, (2000).
- Ungar, L.H.; Powell, B.A.; Kamens, S.N., ”Adaptive Networks for Fault Diagnosis and Process Control”. *Comp. Chem. Eng.*, **14**(4/5), 561, (1990).
- Venkatasubramanian, V.; Chan, K. “A Neural Methodology for Process Fault Diagnosis”. *AIChE Journal*, **35**(12), (1989).
- Venkatasubramanian, V.; McAvoy, T.J. “Neural Network Applications in Chemical Engineering”. *Comp. Chem. Eng.*, **16**(4), v, (1992).
- Venkatasubramanian, V.; Vaidyanathan, R.; Yamamoto, Y. “Process Fault Detection and Diagnosis using Neural Networks-I. Steady-State Processes”. *Comp. Chem. Eng.*, **14**(7), 699, (1990).
- Wang, D.; Hsu, C. “SLONN: A Simulation Language for modeling of Neural Networks”. *Simulation*, 69, (1990).
- Weller, S.W. “Kinetics of Heterogeneous Catalyzed Reactions”. *Catal. Rev. Sci. Eng.*, **34**(3), 227, (1992).
- Werbos, P.J. “Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences”, *Ph.D. Thesis, Harvard University*, (1974).
- Weisstein, W.E. “Confidence Interval”. From *MathWorld* — A Wolfram Web Resource. <http://mathworld.wolfram.com/ConfidenceInterval.html>
- White, H. “Neural-Networks Learning and Statistics”, *AI Expert*, **48**, (1989)
- Willis, M.J.; Di Massimo, C.; Montague, G.A.; Tham, M.T.; Morris, A.J. ”Artificial neural networks in process engineering”. *IEE Proceeding-D*, **138**(3), (1991).
- Zha, Z.; Louhi-Kultanen, M.; Palosaari, S. “Neural Network simulation for non-MSMPR Crystallization”, *Chem. Eng. J.*, **81**, 101, (2001).
- Zhao, W.; Chen, D.; Hu, S. “Optimizing operations conditions based an ANN and modified Gas”, *Comp. Chem. Eng.*, **24**, 61, (2000).
- Zorzetto, L.F.M.; Filho, R.M.; Wolf-Maciel, M.R. “Process modelling development through artificial neural networks and hybrid models”, *Comp. Chem. Eng.*, **24**, 1355, (2000).



# **APÉNDICES**



# A. ALGORITMO DE RETROPROPAGACIÓN

(Haykin, 1999)

El error de la salida de la neurona  $j$  en la iteración  $n$  esta definida por:

$$e_j(n) = d_j(n) - y_j(n), \text{ la neurona } j \text{ es un nodo de salida} \quad (1)$$

Si se define el error instantáneo de la neurona  $j$  como  $\frac{1}{2}e_j^2(n)$ , entonces el valor instantáneo del error total se obtiene sumando  $\frac{1}{2}e_j^2(n)$  sobre todas las neuronas de la capa de salida, estos nodos son los únicos “visibles” y para los cuales los valores de error se pueden calcular directamente. Entonces se puede escribir:

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (2)$$

donde el conjunto  $C$  incluye todas las neuronas de la capa de salida de la red. Sea  $N$  el número total de patrones (ejemplos) contenidos en el conjunto de entrenamiento. El error cuadrado promedio se obtiene sumando  $E(n)$  sobre todas las  $n$  y posteriormente normalizando dicha suma respecto del tamaño del conjunto  $N$ ,

$$E_{prom} = \frac{1}{N} \sum_{n=1}^N E(n) \quad (3)$$

El error instantáneo  $E(n)$ , y por lo tanto el error promedio  $E_{prom}$ , son una función de todos los parámetros libres (pesos, y valores de sesgo) de la red. Para un conjunto de entrenamiento dado,  $E_{prom}$  representa la *función de costo* como una medida de la capacidad de aprendizaje de la red. El objetivo del proceso de aprendizaje es ajustar los parámetros libres de la red para minimizar  $E_{prom}$ . A fin de efectuar dicha minimización se considera un método simple de entrenamiento en el cual los pesos se actualizan sobre una base patrón-a-patrón hasta que una época, esto es, una presentación completa del conjunto de entrenamiento ha sido procesada. El ajuste de los pesos se hace de acuerdo a los errores respectivos calculados para cada patrón presentado a la red.

El promedio aritmético de esos cambios de los pesos individuales sobre el conjunto de entrenamiento, es una estimación del cambio verdadero que resultará de modificar los pesos, sobre la base de minimizar la función  $E_{prom}$  sobre el conjunto de entrenamiento completo.

Si se considera que una neurona  $j$  esta siendo alimentada por un conjunto de señales producidas por una capa de neuronas previa, entonces el campo local inducido  $v_j(n)$  producido a la entrada de la función de activación asociada a la neurona  $j$  es:

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n) \quad (4)$$

donde  $m$  es el número total de entradas (excepto el valor del sesgo) aplicado a la neurona  $j$ . El peso  $w_{j0}$  (correspondiente al valor fijo  $y_0=+1$ ) es igual al valor de sesgo  $b_j$  aplicado a la neurona  $j$ . Por lo tanto la señal  $y_j(n)$  que se obtiene a la salida de la neurona  $j$  en la iteración  $n$  es:

$$y_j(n) = \varphi_j(v_j(n)) \quad (5)$$

El algoritmo de retropropagación aplica una corrección  $\Delta w_{ji}(n)$  al peso  $w_{ji}(n)$ , el cual es proporcional a la derivada parcial  $\partial E(n)/\partial w_{ji}(n)$ . De acuerdo a la regla de la cadena del cálculo elemental, este gradiente se puede expresar como:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (6)$$

La derivada parcial  $\partial E(n)/\partial w_{ji}(n)$  representa un *factor de sensibilidad*, el cual determina la dirección de búsqueda del peso  $w_{ji}$  en el espacio de los pesos.

Diferenciando ambos lados de la ecuación (2) con respecto a  $e_j(n)$ , se obtiene:

$$\frac{\partial E(n)}{\partial e_j(n)} = e_j(n) \quad (7)$$

Diferenciado ambos lados de la ecuación (1) respecto a  $y_j(n)$ , tenemos:

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \quad (8)$$

Ahora, diferenciado la ecuación (5) respecto a  $v_j(n)$ , obtenemos:

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi'_j(v_j(n)) \quad (9)$$

donde la comilla significa diferenciación respecto al argumento. Finalmente, diferenciado (4) respecto a  $w_{ji}(n)$  obtenemos:

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n) \quad (10)$$

Sustituyendo las ecuaciones (7) a (10) en (6) tenemos:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n) \varphi'_j(v_j(n)) y_i(n) \quad (11)$$

La corrección  $\Delta w_{ji}(n)$  aplicada al peso  $w_{ji}(n)$  esta definida por la *Regla Delta* de la siguiente forma:

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad (12)$$

donde  $\eta$  es el parámetro denominado *velocidad de aprendizaje* del algoritmo de retropropagación. El uso del signo menos en la ecuación (12) toma en cuenta el descenso del gradiente en el espacio de errores (buscando en una dirección en la cual un cambio en los pesos reduce el valor de  $E(n)$ ). Sustituyendo la ecuación (11) en la (12), tenemos.

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (13)$$

donde el gradiente local  $\delta_j(n)$  está definido por:

$$\begin{aligned} \delta_j(n) &= -\frac{\partial E(n)}{\partial v_j(n)} \\ &= -\frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \end{aligned}$$

$$= e_j(n)\varphi'_j(v_j(n)) \quad (14)$$

El gradiente local establece los cambios requeridos de los pesos. De acuerdo a la ecuación (14), el gradiente local  $\delta_j(n)$  de la neurona  $j$  es igual al producto de su error correspondiente  $e_j(n)$  y la derivada  $\varphi'_j(v_j(n))$  de la función de activación asociada.

De las ecuaciones (13) y (14) se puede notar que el factor clave involucrado en el cálculo del ajuste de los pesos  $\Delta w_{ji}(n)$  es el error  $e_j(n)$  de la salida de la neurona  $j$ . En este contexto se pueden identificar dos casos, dependiendo de donde esté colocada la neurona  $j$  dentro de la red. En el caso 1, la neurona es un nodo de salida. Este caso es simple de manejar ya que a cada nodo de salida de la red se le proporciona el valor de salida deseado, y de esta forma el cálculo del error es directo. En el caso 2, la neurona es un nodo oculto. Aun cuando las neuronas ocultas no son directamente accesibles, comparten la responsabilidad de cualquier error obtenido en la salida de la red. Aquí la pregunta es saber como penalizar o premiar las neuronas ocultas por compartir esta responsabilidad. Este problema se resuelve retropropagando los errores a través de la red.

Caso 1. La neurona  $j$  es un nodo de salida.

Cuando la neurona  $j$  esta localizada en la capa de salida de la red, el valor de respuesta deseado se proporciona a dicha neurona. Entonces se puede usar la ecuación (1) para calcular el error  $e_j(n)$  asociado con esta neurona. Una vez que se ha determinado el error, el gradiente local  $\delta_j(n)$  se puede calcular de forma directa usando la ecuación (14).

Caso 2. La neurona  $j$  es un nodo oculto.

Cuando la neurona  $j$  esta localizada en una capa oculta de la red, no hay valor deseado disponible para ese nodo. Entonces, el error para la neurona oculta debe ser determinada recursivamente en términos de los errores de todas las neuronas a las cuales esa neurona oculta está directamente conectada, y es aquí donde el desarrollo del algoritmo de retropropagación se vuelve complicado.

De acuerdo a la ecuación (14) es posible definir el gradiente local  $\delta_j(n)$  para una neurona oculta  $j$  como:

$$\delta_j(n) = -\frac{\partial E(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)}$$

$$(15)$$

$$= \frac{\partial E(n)}{\partial y_j(n)} \varphi'_j(v_j(n)), \text{ la neurona } j \text{ es un nodo oculto}$$

donde es la segunda línea se ha usado la ecuación (9). Para calcular la derivada parcial se procede de la siguiente forma:

$$E(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n), \text{ la neurona } k \text{ es un nodo de salida} \quad (16)$$

la cual es la ecuación (2) usando el índice k en lugar del índice j. esto se hace de esta forma a fin de evitar confusión con el uso del índice j que se refiere a una neurona oculta en el caso 2. Diferenciando la ecuación (16) respecto a la señal  $y_j(n)$ , tenemos:

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)} \quad (17)$$

Enseguida se usa la regla de la cadena para calcular la derivada parcial del segundo término, y rescribiendo la ecuación (17) en su forma equivalente, tenemos:

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k(n) \frac{\partial e_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial y_j(n)} \quad (18)$$

Sin embargo se debe hacer notar que:

$$e_k(n) = d_k(n) - y_k(n)$$

$$= d_k(n) - \varphi'_k(v_k(n)), \text{ la neurona } k \text{ es un nodo de salida} \quad (19)$$

Por lo tanto,

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\varphi'_k(v_k(n)) \quad (20)$$

También se debe notar que para la neurona k, el campo inducido es:

$$v_k(n) = \sum_{j=0}^m w_{kj}(n) y_j(n) \quad (21)$$

donde  $m$  es el número total de valores de entradas (excepto el valor de sesgo) aplicado a la neurona  $k$ . Aquí nuevamente, el peso  $w_{k0}(n)$  es igual al valor de sesgo  $b_k(n)$  aplicado a la neurona  $k$ , y su valor se fija a un valor de  $+1$ . Diferenciado la ecuación (21) respecto a  $y_j(n)$  obtenemos:

$$\frac{\partial v_k(n)}{y_j(n)} = w_{kj}(n) \quad (22)$$

Sustituyendo las ecuaciones (20) y (22) en la ecuación (18) obtenemos la derivada parcial requerida:

$$\begin{aligned} \frac{\partial E(n)}{\partial y_j(n)} &= -\sum_k e_k(n) \varphi'_k(v_k(n)) w_{kj}(n) \\ &= \sum \delta_k(n) w_{kj}(n) \end{aligned} \quad (23)$$

donde en la segunda línea se ha usado la definición del gradiente local  $\delta_k(n)$  dado en la ecuación (14) donde se ha sustituido el índice  $k$  por el  $j$ .

Finalmente, se usa la ecuación (23) en la ecuación (15), para obtener la fórmula de retropropagación para el gradiente local  $\delta_j(n)$ .

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_k \delta_k(n) w_{kj}(n), \quad \text{la neurona } j \text{ es un nodo oculto} \quad (24)$$

El factor  $\varphi'_j(v_j(n))$  involucrado en el cálculo del gradiente local  $\delta_j(n)$  en la ecuación (24) solo depende de la función de activación asociada con la neurona oculta  $j$ . El factor restante involucrado en este cálculo, llamémosle la sumatoria sobre  $k$ , depende de dos conjuntos de términos. El primer conjunto de términos  $\delta_k(n)$ , requiere del conocimiento de los errores  $e_k(n)$ , para todas las neuronas que yacen en la capa siguiente de la neurona  $j$ , y que están directamente conectadas a dicha neurona. El segundo conjunto de términos  $w_{kj}(n)$ , consiste de los pesos asociados con esas conexiones.

Enseguida se resumen las relaciones que han sido derivadas para el algoritmo de retropropagación. Primero se muestran las correcciones aplicadas a los pesos que conectan a la neurona  $i$  con la neurona  $j$ , las cuales están definidas por la *Regla Delta*:

$$\begin{pmatrix} \text{Corrección} \\ \Delta w_{ji}(n) \end{pmatrix} = \begin{pmatrix} \text{Velocidad} \\ \text{de aprendizaje} \\ \eta \end{pmatrix} \cdot \begin{pmatrix} \text{Gradiente} \\ \text{Local} \\ \delta_j(n) \end{pmatrix} \cdot \begin{pmatrix} \text{Entrada} \\ \text{neurona } j \\ y_i(n) \end{pmatrix} \quad (25)$$

Segundo, el gradiente local  $\delta_j(n)$  depende de si el nodo considerado es un nodo de salida u oculto:

1. Si la neurona  $j$  es un nodo de salida,  $\delta_j(n)$  es igual al producto de la derivada  $\varphi'_j(v_j(n))$  y el error  $e_j(n)$ , ambos asociados con la neurona  $j$ . (Ver ecuación 14)
2. Si la neurona  $j$  es un nodo oculto,  $\delta_j(n)$  es igual al producto de la derivada  $\varphi'_j(v_j(n))$  asociada y la suma ponderada de las  $\delta$ 's calculadas para las neuronas en la siguiente capa oculta o de salida que están conectadas a la neurona  $j$ . (Ver ecuación 24).



## **B. PROGRAMA NEURONAF**

Una vez determinada la topología óptima, así como los parámetros de la Red Neuronal Artificial que mejor representa el comportamiento de un Reactor de Hidrodesulfuración de Naftas, se desarrolla un programa denominado “**NeuroNaf**”.

Este programa fue desarrollado en Visual Basic® Versión 5.0, ya que se consideró que la lógica del código no es muy complicada, ni demanda demasiados recursos al sistema operativo, además de que se requería tener una interfase gráfica que facilitara el uso de dicho programa.

El programa consta básicamente de cuatro ventanas que aparecen en forma secuencial, conforme avanza la ejecución. En primer lugar aparece la ventana de inicio donde se especifica el nombre del programa, las plataformas en las que corre, la versión, el nombre del desarrollador, y una pequeña advertencia respecto a los derechos de autor.

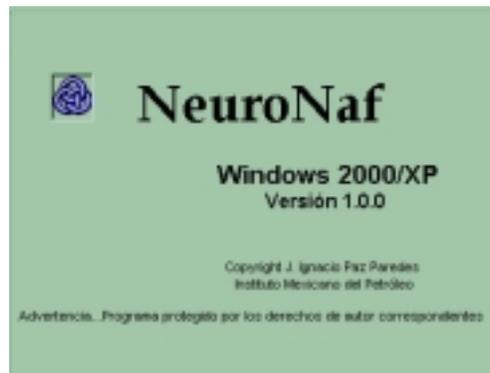


Figura B.1 Ventana de arranque del Programa NeuroNaf

Esta ventana sólo es visible por unos instantes, y da paso al cuadro de dialogo que permite seleccionar cual de las topologías se utilizará:

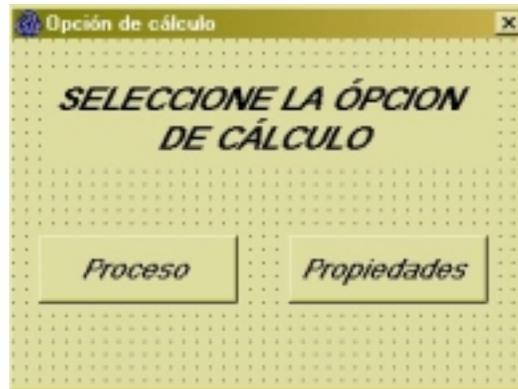


Figura B.2 Ventana de selección de opciones del programa Neuronaf

En caso de seleccionar la opción “Proceso”, aparecerá la ventana correspondiente a la topología “Proceso”. (Figura B.3)

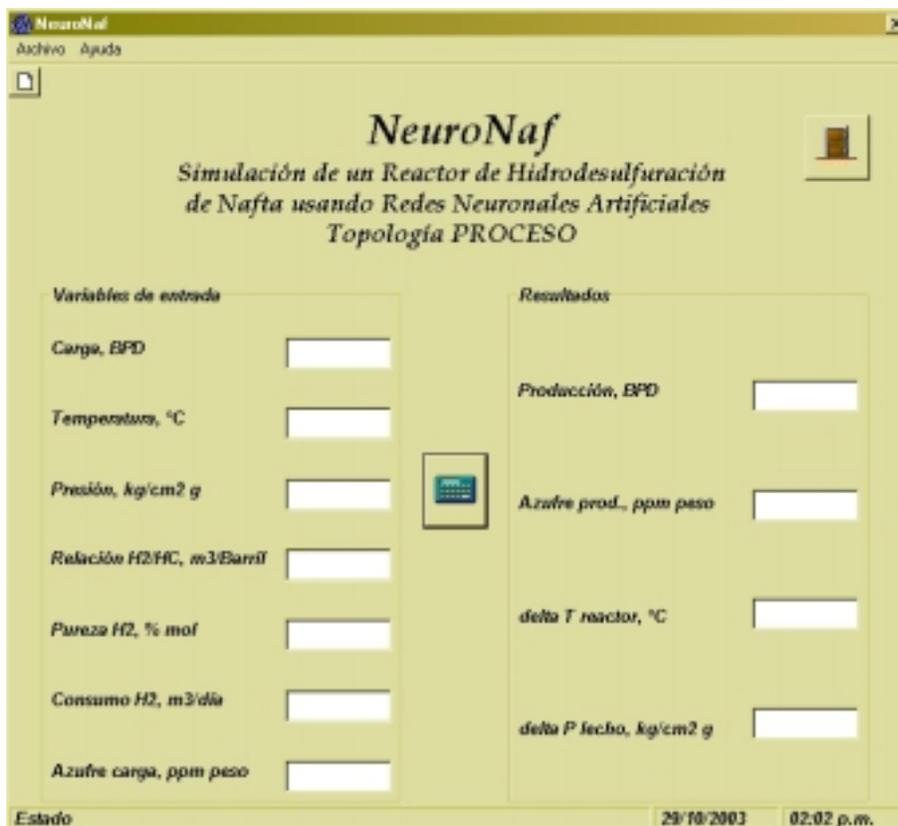


Figura B.3 Ventana de la topología Proceso del Programa NeuroNaf

En la ventana de la figura B.3 se realizan las operaciones requeridas para la simulación de las variables de proceso de un Reactor de Hidrodesulfuración de Naftas. Consta de

dos secciones que contienen: la primera todos los datos de entrada requeridos para efectuar la simulación, y la segunda, los valores de salida del proceso.

En caso de que se seleccione la opción “Propiedades”, aparece la ventana mostrada en la figura B.4.



Figura B.4 Ventana de la topología “Propiedades” del Programa NeuroNaf

En esta ventana se efectúan todas las operaciones para proporcionar los valores de las variables de salida correspondientes a las propiedades del producto.

El programa cuenta con algunos procedimientos de seguridad; el primero es que asegura que todas las variables de entrada tengan valores válidos mayores de cero, y que sean numéricos, para lo cual en caso de omitir alguno de ellos tomará el promedio de los valores usados en el entrenamiento de la red de la variable correspondiente. El segundo procedimiento de seguridad verifica que cada una de las variables de entrada se encuentre entre los intervalos válidos, sobre los cuales fue entrenada la Red Neuronal.

Dentro de ambas ventanas de las topologías están colocados dos botones de ejecución.

Uno sirve para iniciar los cálculos , mientras que el segundo se usa para salir del programa .

El programa cuenta con un menú de herramientas, el cual sólo contiene el botón , el cual limpia todos los campos de entrada y salida, para introducir un nuevo caso de simulación.

También se cuenta con un menú principal el cual está formado por dos opciones: Archivo, y Ayuda.

La opción Archivo,



esta formada a su vez por dos opciones; “Nuevo”, tiene la misma funcionalidad que el botón , mientras que “Salir”, termina el programa.

La opción Ayuda,



sólo contiene la opción “Acerca de NeuroNaf...”, la cual se encarga de desplegar la ventana de información del programa, como se muestra en la figura B.5.

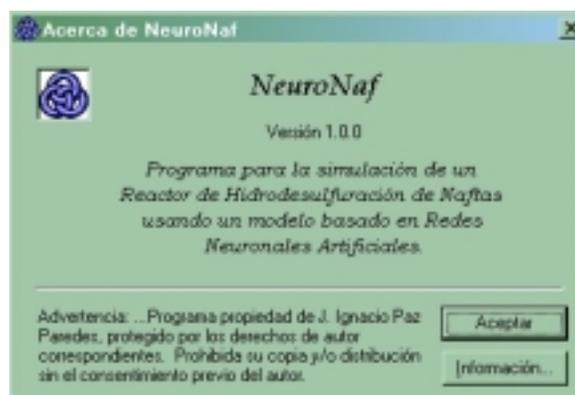


Figura B.5 Ventana de información del Programa NeuroNaf.

El programa está diseñado para que al oprimir la tecla “Enter” se inicien los cálculos, mientras que si se oprime la tecla “Esc” el programa termina.

## **C. LISTADO DEL PROGRAMA**

```
Sub Main()  
  
    frmSplash.Show  
  
    frmSplash.Refresh  
  
    Set fMainForm = New frmMain  
    Load fMainForm  
  
    Unload frmSplash  
  
    fMainForm.Show  
  
End Sub  
  
Private Sub Form_Load()  
    lblVersion.Caption = "Version " & App.Major & "." & App.Minor & "." & App.Revision  
    lblProductName.Caption = App.Title  
End Sub  
  
Public Function ErrorCheck()  
  
    If Val(fMainForm.txtEntrada(0).Text) <= 0 Then  
  
        intPress = MsgBox("El dato de carga no es valido", vbCritical, "Error en los datos  
de entrada")  
        fMainForm.txtEntrada(0).SetFocus  
        ErrorCheck = 1  
  
    ElseIf Val(fMainForm.txtEntrada(1).Text) <= 0 Then  
  
        intPress = MsgBox("El dato de Temperatura no es valido", vbCritical, "Error en los  
datos de entrada")  
        fMainForm.txtEntrada(1).SetFocus  
        ErrorCheck = 1  
  
    ElseIf Val(fMainForm.txtEntrada(2).Text) <= 0 Then  
  
        intPress = MsgBox("El dato de Presion no es valido", vbCritical, "Error en los datos  
de entrada")  
        fMainForm.txtEntrada(2).SetFocus  
        ErrorCheck = 1  
  
    ElseIf Val(fMainForm.txtEntrada(3).Text) <= 0 Then  
  
        intPress = MsgBox("El dato de la Relacion H2/HC no es valido", vbCritical, "Error en  
los datos de entrada")  
        fMainForm.txtEntrada(3).SetFocus  
        ErrorCheck = 1  
  
    ElseIf Val(fMainForm.txtEntrada(4).Text) <= 0 Then  
  
        intPress = MsgBox("El dato de la Pureza de H2 no es valido", vbCritical, "Error en  
los datos de entrada")  
        fMainForm.txtEntrada(4).SetFocus  
        ErrorCheck = 1  
  
    ElseIf Val(fMainForm.txtEntrada(5).Text) <= 0 Then  
  
        intPress = MsgBox("El dato del Consumo de H2 no es valido", vbCritical, "Error en  
los datos de entrada")  
        fMainForm.txtEntrada(5).SetFocus  
        ErrorCheck = 1
```

```
ElseIf Val(fMainForm.txtEntrada(6).Text) <= 0 Then

    intPress = MsgBox("El dato de TIE carga no es valido", vbCritical, "Error en los
datos de entrada")
    fMainForm.txtEntrada(6).SetFocus
    ErrorCheck = 1

ElseIf Val(fMainForm.txtEntrada(7).Text) <= 0 Then

    intPress = MsgBox("El dato de TFE carga no es valido", vbCritical, "Error en los
datos de entrada")
    fMainForm.txtEntrada(7).SetFocus
    ErrorCheck = 1

ElseIf Val(fMainForm.txtEntrada(8).Text) <= 0 Then

    intPress = MsgBox("El dato de Peso especifico de la carga no es valido", vbCritical,
"Error en los datos de entrada")
    fMainForm.txtEntrada(8).SetFocus
    ErrorCheck = 1

ElseIf Val(fMainForm.txtEntrada(9).Text) <= 0 Then

    intPress = MsgBox("El dato de Azufre carga no es valido", vbCritical, "Error en los
datos de entrada")
    fMainForm.txtEntrada(9).SetFocus
    ErrorCheck = 1

ElseIf ((22938 < Val(fMainForm.txtEntrada(0).Text)) Or _
        (Val(fMainForm.txtEntrada(0).Text) < 18838)) Then

    intPress = MsgBox("El dato de Carga esta fuera del intervalo: 18838-22938 BPD. Los
resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 0
    fMainForm.txtEntrada(0).SetFocus

ElseIf ((279 > Val(fMainForm.txtEntrada(1).Text)) Or _
        (Val(fMainForm.txtEntrada(1).Text) > 299)) Then

    intPress = MsgBox("El dato de Temperatura esta fuera del intervalo: 279-299 C. Los
resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(1).SetFocus

ElseIf ((44 > Val(fMainForm.txtEntrada(2).Text)) Or _
        (Val(fMainForm.txtEntrada(2).Text) > 60)) Then

    intPress = MsgBox("El dato de Presion esta fuera del intervalo: 44-60 kg/cm2 g. Los
resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(2).SetFocus

ElseIf ((26.63 > Val(fMainForm.txtEntrada(3).Text)) Or _
        (Val(fMainForm.txtEntrada(3).Text) > 43.85)) Then

    intPress = MsgBox("El dato de Relacion H2/HC esta fuera del intervalo: 26.43-43.85
m3/Bbl. Los resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(3).SetFocus

ElseIf ((74.4 > Val(fMainForm.txtEntrada(4).Text)) Or _
        (Val(fMainForm.txtEntrada(4).Text) > 96.4)) Then

    intPress = MsgBox("El dato de Pureza de H2 esta fuera del intervalo: 74.4-96.4 %
mol. Los resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(4).SetFocus

ElseIf ((50562 > Val(fMainForm.txtEntrada(5).Text)) Or _
        (Val(fMainForm.txtEntrada(5).Text) > 124566)) Then

    intPress = MsgBox("El dato de Consumo de H2 esta fuera del intervalo: 50562-124566
m3/dia. Los resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(5).SetFocus
```

```
ElseIf ((54 > Val(fMainForm.txtEntrada(6).Text)) Or _
        (Val(fMainForm.txtEntrada(6).Text) > 74)) Then

    intPress = MsgBox("El dato de TFE carga esta fuera del intervalo: 54-74 C. Los
resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(6).SetFocus

ElseIf ((170 > Val(fMainForm.txtEntrada(7).Text)) Or _
        (Val(fMainForm.txtEntrada(7).Text) > 193)) Then

    intPress = MsgBox("El dato de TFE carga esta fuera del intervalo: 170-193 C. Los
resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(7).SetFocus

ElseIf ((0.725 > Val(fMainForm.txtEntrada(8).Text)) Or _
        (Val(fMainForm.txtEntrada(8).Text) > 0.738)) Then

    intPress = MsgBox("El dato de Peso especifico de la carga esta fuera del intervalo:
0.725-0.738. Los resultados pueden no ser confiables", vbCritical, "Dato fuera de
intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(8).SetFocus

ElseIf ((360 > Val(fMainForm.txtEntrada(9).Text)) Or _
        (Val(fMainForm.txtEntrada(9).Text) > 1532)) Then

    intPress = MsgBox("El dato de Azufre de carga esta fuera del intervalo: 360-1532.
Los resultados pueden no ser confiables", vbCritical, "Dato fuera de intervalo")
    ErrorCheck = 1
    fMainForm.txtEntrada(9).SetFocus

End If

End Function

Private Sub cmdCalcula_Click()
'Verifica la validez de los datos de entrada a través de la funcion
'Errorcheck

If ErrorCheck = 1 Then

    Exit Sub

End If

'Si los datos son mayores de cero y sus valores entran dentro del intervalo válido
'de cada variable, los calculos continuan, en caso contrario el programa se
'interrumpe y el cursor se coloca en el campo que se debe modificar.

Asigna

Calcula

End Sub

Private Sub cmdSalida_Click()
'Descarga el formulario principal
Unload Me
End Sub

Private Sub Form_Load()
Me.Left = GetSetting(App.Title, "Settings", "MainLeft", 1000)
Me.Top = GetSetting(App.Title, "Settings", "MainTop", 1000)
Me.Width = GetSetting(App.Title, "Settings", "MainWidth", 6500)
Me.Height = GetSetting(App.Title, "Settings", "MainHeight", 6500)
End Sub

Private Sub Form_Unload(Cancel As Integer)
Dim i As Integer

'close all sub forms
```

```
For i = Forms.Count - 1 To 1 Step -1
    Unload Forms(i)
Next
If Me.WindowState <> vbMinimized Then
    SaveSetting App.Title, "Settings", "MainLeft", Me.Left
    SaveSetting App.Title, "Settings", "MainTop", Me.Top
    SaveSetting App.Title, "Settings", "MainWidth", Me.Width
    SaveSetting App.Title, "Settings", "MainHeight", Me.Height
End If
End Sub

Private Sub Frame1_DragDrop(Source As Control, x As Single, Y As Single)

End Sub

Private Sub mnuHelpAbout_Click()
    frmAbout.Show vbModal, Me
End Sub

Private Sub tbToolBar_ButtonClick(ByVal Button As ComctlLib.Button)

    Select Case Button.Key

        Case "New"
            mnuFileNew_Click
        Case "Open"
            mnuFileOpen_Click
        Case "Save"
            mnuFileSave_Click
        Case "Print"
            mnuFilePrint_Click

    End Select

End Sub

Private Sub mnuHelpContents_Click()

    Dim nRet As Integer

    'Si no hay archivo de Ayuda para este proyecto, muestra un mensaje al usuario
    'puede establecer el archivo de Ayuda para su aplicación en el cuadro de
    'diálogo Propiedades del proyecto
    If Len(App.HelpFile) = 0 Then
        MsgBox "Imposible mostrar los contenidos de la Ayuda. No hay una Ayuda asociada con
este proyecto.", vbInformation, Me.Caption
    Else
        On Error Resume Next
        'nRet = OSWinHelp(Me.hWnd, App.HelpFile, 3, 0)
        If Err Then
            MsgBox Err.Description
        End If
    End If
End Sub

Private Sub mnuHelpSearch_Click()

    Dim nRet As Integer

    'Si no hay archivo de Ayuda para este proyecto, muestra un mensaje al usuario
    'puede establecer el archivo de Ayuda para su aplicación en el cuadro de
    'diálogo Propiedades del proyecto
    If Len(App.HelpFile) = 0 Then
        MsgBox "Imposible mostrar los contenidos de la Ayuda. No hay una Ayuda asociada con
este proyecto.", vbInformation, Me.Caption
    Else
        On Error Resume Next
```

```
        'nRet = OSWinHelp(Me.hWnd, App.HelpFile, 261, 0)
    If Err Then
        MsgBox Err.Description
    End If
End If
End Sub

Private Sub mnuFileOpen_Click()
    Dim sFile As String

    With dlgCommonDialog
        'Para hacer
        'Establece los indicadores y atributos del
        'control Common Dialog
        .Filter = "Todos los archivos (*.*)|*.*"
        .ShowOpen
        If Len(.filename) = 0 Then
            Exit Sub
        End If
        sFile = .filename
    End With
    'Para hacer
    'Procesa el archivo abierto
End Sub

Private Sub mnuFileClose_Click()
    'Para hacer
    MsgBox "Aquí se sitúa el código para cerrar"
End Sub

Private Sub mnuFileSave_Click()
    'Para hacer
    MsgBox "Aquí se sitúa el código para guardar"
End Sub

Private Sub mnuFileSaveAs_Click()
    'Para hacer
    'Configura el control Common Dialog
    'antes de llamar a ShowSave
    dlgCommonDialog.ShowSave
End Sub

Private Sub mnuFileSaveAll_Click()
    'Para hacer
    MsgBox "Aquí se sitúa el código para guardar todo"
End Sub

Private Sub mnuFileProperties_Click()
    'Para hacer
    MsgBox "Aquí se sitúa el código de las propiedades"
End Sub

Private Sub mnuFilePageSetup_Click()
    dlgCommonDialog.ShowPrinter
End Sub

Private Sub mnuFilePrintPreview_Click()
    'Para hacer
    MsgBox "Aquí se sitúa el código de vista previa"
End Sub

Private Sub mnuFilePrint_Click()
    'Para hacer
    MsgBox "Aquí se sitúa el código de impresión"
End Sub
```

```
Private Sub mnuFileSend_Click()  
    'Para hacer  
    MsgBox "Aquí se sitúa el código de enviar"  
End Sub  
  
Private Sub mnuFileMRU_Click(Index As Integer)  
    'Para hacer  
    MsgBox "Aquí se sitúa el código de archivos recientes"  
End Sub  
  
Private Sub mnuFileExit_Click()  
    'Descarga el formulario  
    Unload Me  
End Sub  
  
Private Sub mnuFileNew_Click()  
    'Para hacer  
    fMainForm.txtEntrada(0).Text = ""  
    fMainForm.txtEntrada(1).Text = ""  
    fMainForm.txtEntrada(2).Text = ""  
    fMainForm.txtEntrada(3).Text = ""  
    fMainForm.txtEntrada(4).Text = ""  
    fMainForm.txtEntrada(5).Text = ""  
    fMainForm.txtEntrada(6).Text = ""  
    fMainForm.txtEntrada(7).Text = ""  
    fMainForm.txtEntrada(8).Text = ""  
    fMainForm.txtEntrada(9).Text = ""  
  
    fMainForm.txtResultados(0).Text = ""  
    fMainForm.txtResultados(1).Text = ""  
    fMainForm.txtResultados(2).Text = ""  
    fMainForm.txtResultados(3).Text = ""  
    fMainForm.txtResultados(4).Text = ""  
    fMainForm.txtResultados(5).Text = ""  
    fMainForm.txtResultados(6).Text = ""  
  
    fMainForm.txtEntrada(0).SetFocus  
  
    'MsgBox "New File Code goes here!"  
End Sub  
  
'Public Sub Verifica()  
  
    ' If ErrorCheck() = 1 Then  
  
        Exit Sub  
  
    ' End If  
  
'End Sub  
  
Public fMainForm As frmMain  
  
Public Sub Calcula()  
'Esta subrutina calcula las variables de salida de un Reactor de Hidrodesulfuración  
'de Naftas, para ello toma los valores del formato frmMain que previamente ya fueron  
'asignadas a las variables de entrada, como carga, T,P, etc.  
  
'Como primer paso los valores de entrada se deben escalar para que puedan ser usados  
'por el programa. Para este fin se llama a la subrutina "Escalar".  
  
    Escalar  
  
'Una vez que se obtienen los datos escalados se procede a llamar a las subrutinas  
  
    'Principal, y  
    'Propiedades.  
  
    Principal  
  
    Propiedades  
  
    Descalar
```

```
Resultados

End Sub

Public prin(1 To 7) As Double
Public propi(1 To 3) As Double

Public Sub Escalar()
'Esta subrutina escala las variables de entrada del proceso de Hidrodesulfuración de
'Naftas, para que sea utilizado por el programa.
'
'           x_nor = (x - x_min)/(x_max - x_min)
'Escalamiento de la carga
prin(1) = (carga - 18838#) / (22938# - 18838#)
'Escalamiento de la temperatura de operacion
prin(2) = (T - 279#) / (299# - 279#)
'Escalamiento de la presion de operacion
prin(3) = (P - 44#) / (60# - 44#)
'Escalamiento de la relacion H2/HC
prin(4) = (relacion - 26.63) / (43.85 - 26.63)
'Escalamiento de la pureza de H2
prin(5) = (pureza - 74.4) / (96.4 - 74.4)
'Escalamiento del consumo de H2
prin(6) = (consumo - 50562#) / (124566# - 50562#)
'Escalamiento de TIE de la carga
propi(1) = (tie_carga - 54#) / (74# - 54#)
'Escalamiento de TFE de la carga
propi(2) = (tfe_carga - 170#) / (193# - 170#)
'Escalamiento del peso especifico de la carga
propi(3) = (pe_carga - 0.725) / (0.738 - 0.725)
'Escalamiento del Azufre de la carga
prin(7) = (azufre_carga - 360#) / (1532# - 360#)

End Sub

Public carga As Long
Public T As Long
Public P As Long
Public relacion As Long
Public pureza As Long
Public consumo As Long
Public tie_carga As Long
Public tfe_carga As Long
Public pe_carga As Long
Public azufre_carga As Long
Public Sub Asigna()

carga = Val(fMainForm.txtEntrada(0).Text)

T = Val(fMainForm.txtEntrada(1).Text)

P = Val(fMainForm.txtEntrada(2).Text)

relacion = Val(fMainForm.txtEntrada(3).Text)
```

```
pureza = Val(fMainForm.txtEntrada(4).Text)

consumo = Val(fMainForm.txtEntrada(5).Text)

tie_carga = Val(fMainForm.txtEntrada(6).Text)

tfe_carga = Val(fMainForm.txtEntrada(7).Text)

pe_carga = Val(fMainForm.txtEntrada(8).Text)

azufre_carga = Val(fMainForm.txtEntrada(9).Text)

End Sub

Public hidden1_index As Integer
Public output_index As Integer
Public numneurons As Integer
Public inputs As Integer
Public hidden1 As Integer
Public outputs As Integer
Public prin_sal(1 To 4) As Double
Option Explicit
Public Sub Principal()
'
'Esta subrutina calcula las variables de salida:
' (1) Producción.
' (2) Azufre en el producto.
' (3) delta T del proceso, y
' (4) delta P del lecho catalítico.
'
inputs = 7
hidden1 = 5
outputs = 4

Static neuron(1 To 16) As Double

hidden1_index = 7
output_index = 11
numneurons = 16

'Se definen los arreglos de los pesos

Static W_In_H1(1 To 5, 1 To 8) As Double
Static W_H1_Out(1 To 4, 1 To 6) As Double

'Pesos de la capa de entrada a la capa intermedia u oculta.

W_In_H1(1, 1) = 1.5059
W_In_H1(1, 2) = 1.39202
W_In_H1(1, 3) = 0.814144
W_In_H1(1, 4) = -1.08094
W_In_H1(1, 5) = 0.382455
W_In_H1(1, 6) = -0.495952
W_In_H1(1, 7) = 0.660286
W_In_H1(1, 8) = 0.477905

W_In_H1(2, 1) = -1.26218
W_In_H1(2, 2) = 1.99939
W_In_H1(2, 3) = -6.01269
W_In_H1(2, 4) = -1.57681
W_In_H1(2, 5) = 9.58442
W_In_H1(2, 6) = -7.52902
W_In_H1(2, 7) = -1.04505
W_In_H1(2, 8) = -2.01765

W_In_H1(3, 1) = -1.15039
W_In_H1(3, 2) = -0.813772
W_In_H1(3, 3) = -1.64175
W_In_H1(3, 4) = 3.96504
W_In_H1(3, 5) = 1.79479
W_In_H1(3, 6) = 2.05351
W_In_H1(3, 7) = -1.99471
W_In_H1(3, 8) = 3.62329

W_In_H1(4, 1) = -1.92015
W_In_H1(4, 2) = 1.73387
```

```
W_In_H1(4, 3) = -11.0361
W_In_H1(4, 4) = -10.0171
W_In_H1(4, 5) = 2.79984
W_In_H1(4, 6) = -3.91067
W_In_H1(4, 7) = 2.29023
W_In_H1(4, 8) = 3.68283
```

```
W_In_H1(5, 1) = -3.44805
W_In_H1(5, 2) = 3.2998
W_In_H1(5, 3) = 5.18195
W_In_H1(5, 4) = -1.85304
W_In_H1(5, 5) = -0.24463
W_In_H1(5, 6) = -2.02917
W_In_H1(5, 7) = 2.62674
W_In_H1(5, 8) = -0.257114
```

'Pesos de la capa intermedia u oculta a la capa de salida

```
W_H1_Out(1, 1) = 3.85194
W_H1_Out(1, 2) = -0.261515
W_H1_Out(1, 3) = -0.613194
W_H1_Out(1, 4) = -0.524965
W_H1_Out(1, 5) = -1.59257
W_H1_Out(1, 6) = -0.424198
```

```
W_H1_Out(2, 1) = 0.647935
W_H1_Out(2, 2) = 0.427976
W_H1_Out(2, 3) = -0.0954302
W_H1_Out(2, 4) = -0.529808
W_H1_Out(2, 5) = -1.59257
W_H1_Out(2, 6) = -0.424198
```

```
W_H1_Out(3, 1) = -1.73332
W_H1_Out(3, 2) = 3.58732
W_H1_Out(3, 3) = -4.20251
W_H1_Out(3, 4) = -2.10412
W_H1_Out(3, 5) = -0.875547
W_H1_Out(3, 6) = 3.49447
```

```
W_H1_Out(4, 1) = 4.1864
W_H1_Out(4, 2) = -2.1257
W_H1_Out(4, 3) = 2.38989
W_H1_Out(4, 4) = -0.10407
W_H1_Out(4, 5) = 2.96715
W_H1_Out(4, 6) = -5.46649
```

'Definición de los factores de escalamiento usando internamente por la Red

```
Static R_min(1 To 11) As Double
Static R_max(1 To 11) As Double
Static S_min(1 To 11) As Double
Static S_max(1 To 11) As Double
```

'Asignación de los vectores de los factores de escalamiento

```
R_min(1) = 0#
R_min(2) = 0#
R_min(3) = 0#
R_min(4) = 0#
R_min(7) = 0#
R_min(6) = 0#
R_min(7) = 0#
R_min(8) = 0.1
R_min(9) = 0.1
R_min(10) = 0.1
R_min(11) = 0.1
```

```
R_max(1) = 1#
R_max(2) = 1#
R_max(3) = 1#
R_max(4) = 1#
R_max(5) = 1.1
R_max(6) = 1#
R_max(7) = 1#
R_max(8) = 0.9
R_max(9) = 0.9
```

```
R_max(10) = 0.9
R_max(11) = 0.9
```

```
S_min(1) = -1#
S_min(2) = -1#
S_min(3) = -1#
S_min(4) = -1#
S_min(5) = -1#
S_min(6) = -1#
S_min(7) = -1#
S_min(8) = 0#
S_min(9) = 0#
S_min(10) = 0#
S_min(11) = 0#
```

```
S_max(1) = 1#
S_max(2) = 1#
S_max(3) = 1#
S_max(4) = 1#
S_max(5) = 1#
S_max(6) = 1#
S_max(7) = 1#
S_max(8) = 1#
S_max(9) = 1#
S_max(10) = 1#
S_max(11) = 1#
```

```
'Se calcula la respuesta de la Red Neuronal Artificial.
```

```
Dim F As Double
Dim i, j As Integer
```

```
'Se inicializa a cero la suma de las neuronas
```

```
For i = 1 To numneurons
```

```
    neuron(i) = 0#
```

```
Next i
```

```
'Se efectúa el escalamiento interno de la Red Neuronal
```

```
For i = 1 To inputs
```

```
    F = (S_max(i) - S_min(i)) / (R_max(i) - R_min(i)) 'Factor de escalamiento
```

```
    neuron(i) = F * prin(i) + S_min(i) - F * R_min(i)
```

```
Next i
```

```
'Se calcula la capa interna u oculta
```

```
For i = 1 To hidden1
```

```
    For j = 1 To inputs
```

```
        neuron(hidden1_index + i) = neuron(hidden1_index + i) + _
        neuron(j) * W_In_H1(i, j)
```

```
    Next j
```

```
    neuron(hidden1_index + i) = neuron(hidden1_index + i) + _
    W_In_H1(i, inputs) 'Sesgo (bias)
```

```
    neuron(hidden1_index + i) = sigmoide(neuron(hidden1_index + i))
```

```
Next i
```

```
'Se calculan las salidas de la Red Neuronal
```

```
For i = 1 To outputs
```

```
    For j = 1 To hidden1
```

```
        neuron(output_index + i) = neuron(output_index + i) + _
        neuron(hidden1_index + j) * W_H1_Out(i, j)
```

```
Next j

neuron(output_index + i) = neuron(output_index + i) + _
W_H1_Out(i, hidden1) 'Sesgo (bias)

neuron(output_index + i) = sigmoide(neuron(output_index + i))

Next i

'Se copian las salidas al arreglo prin_sal

For i = 1 To outputs

F = (R_max(i + inputs) - R_min(i + inputs)) / _
(S_max(i + inputs) - S_min(i + inputs))

prin_sal(i) = F * neuron(output_index + 1) + R_min(i + inputs) - _
F * S_min(i + inputs)

Next i

End Sub

Public Function sigmoide(X) As Double
'
'Función que calcula la función de transferencia
'
' Dim x As Double
sigmoide = 1# / (1# + Exp(-X))

End Function

Option Explicit
Public propi_sal(1 To 3) As Double

Public Sub Propiedades()
'
'Esta subrutina corre la red neuronal que se encarga de calcular las propiedades
'de la gasolina producto del proceso de Hidrodesulfuración.
'
' TIE de la gasolina
' TFE de la gasolina
' Peso específico de la gasolina

inputs = 3
hidden1 = 3
outputs = 3

Static neuron(1 To 16) As Double

hidden1_index = 3
output_index = 6
numneurons = 9

'Se definen los arreglos de los pesos

Static W_In_H1(1 To 3, 1 To 4) As Double
Static W_H1_Out(1 To 3, 1 To 4) As Double

'Pesos de la capa de entrada a la capa intermedia u oculta.

W_In_H1(1, 1) = -4.43867
W_In_H1(1, 2) = -2.98272
W_In_H1(1, 3) = -4.20848
W_In_H1(1, 4) = 0.610347

W_In_H1(2, 1) = 0.756819
W_In_H1(2, 2) = 0.938212
W_In_H1(2, 3) = -0.834173
W_In_H1(2, 4) = -0.856755

W_In_H1(3, 1) = -0.379684
W_In_H1(3, 2) = 5.21926
```

```
W_In_H1(3, 3) = 1.15631
W_In_H1(3, 4) = -0.0991506

'Pesos de la capa intermedia u oculta a la capa de salida

W_H1_Out(1, 1) = -1.95456
W_H1_Out(1, 2) = 1.38145
W_H1_Out(1, 3) = -1.54145
W_H1_Out(1, 4) = 1.50676

W_H1_Out(2, 1) = 0.0681781
W_H1_Out(2, 2) = 2.116
W_H1_Out(2, 3) = 2.16627
W_H1_Out(2, 4) = -2.14771

W_H1_Out(3, 1) = -1.19823
W_H1_Out(3, 2) = -4.49574
W_H1_Out(3, 3) = 0.810694
W_H1_Out(3, 4) = 1.32744

'Definición de los factores de escalamiento usando internamente por la Red

Static R_min(1 To 6) As Double
Static R_max(1 To 6) As Double
Static S_min(1 To 6) As Double
Static S_max(1 To 6) As Double

'Asignación de los vectores de los factores de escalamiento

R_min(1) = 0#
R_min(2) = 0#
R_min(3) = 0#
R_min(4) = 0.1
R_min(5) = 0.1
R_min(6) = 0.1

R_max(1) = 1#
R_max(2) = 1#
R_max(3) = 1#
R_max(4) = 0.9
R_max(5) = 0.9348
R_max(6) = 0.9

S_min(1) = -1#
S_min(2) = -1#
S_min(3) = -1#
S_min(4) = 0#
S_min(5) = 0#
S_min(6) = 0#

S_max(1) = 1#
S_max(2) = 1#
S_max(3) = 1#
S_max(4) = 1#
S_max(5) = 1#
S_max(6) = 1#

'Se calcula la respuesta de la Red Neuronal Artificial.

Dim F As Double
Dim i, j As Integer

'Se inicializa a cero la suma de las neuronas

For i = 1 To numneurons
    neuron(i) = 0#
Next i

'Se efectúa el escalamiento interno de la Red Neuronal

For i = 1 To inputs
    F = (S_max(i) - S_min(i)) / (R_max(i) - R_min(i)) 'Factor de escalamiento
```

```
neuron(i) = F * propi(i) + S_min(i) - F * R_min(i)
'neuron(i) = propi(i)
Next i
'Se calcula la capa interna u oculta
For i = 1 To hidden1
  For j = 1 To inputs
    neuron(hidden1_index + i) = neuron(hidden1_index + i) + _
    neuron(j) * W_In_H1(i, j)
  Next j
  neuron(hidden1_index + i) = neuron(hidden1_index + i) + _
  W_In_H1(i, inputs) 'Sesgo (bias)
  neuron(hidden1_index + i) = sigmoide(neuron(hidden1_index + i))
Next i
'Se calculan las salidas de la Red Neuronal
For i = 1 To outputs
  For j = 1 To hidden1
    neuron(output_index + i) = neuron(output_index + i) + _
    neuron(hidden1_index + j) * W_H1_Out(i, j)
  Next j
  neuron(output_index + i) = neuron(output_index + i) + _
  W_H1_Out(i, hidden1) 'Sesgo (bias)
  neuron(output_index + i) = sigmoide(neuron(output_index + i))
Next i
'Se copian las salidas al arreglo prin_sal
For i = 1 To outputs
  F = (R_max(i + inputs) - R_min(i + inputs)) / _
  (S_max(i + inputs) - S_min(i + inputs))
  propi_sal(i) = F * neuron(output_index + 1) + R_min(i + inputs) - _
  F * S_min(i + inputs)
  'propi_sal(i) = neuron(output_index + 1)
Next i
End Sub

Public produccion As Double
Public tie_producto As Double
Public tfe_producto As Double
Public pe_producto As Double
Public azufre_producto As Double
Public delta_T As Double
Public delta_P As Double

Public Sub Descalar()
'
'Subrutina que se encarga de desescalar los valores de las variables de salida de
'la Red Neuronal que simula la operación y desempeño de un Reactor de Hidrodesul-
'furación de Naftas.
'
'Las variables son:
'Producción
'TIE del producto.
'TFE del producto.
```

```
'Peso especifico del producto.
'Azufre del producto.
'delta T del proceso.
'delta P del Lecho catalitico.

'Producción.

    produccion = (6845# * prin_sal(1) + 15979.5) * 1.22

'TIE producto

    tie_producto = (21.25 * propi_sal(1) + 58.875) * 1.1

'TFE producto

    tfe_producto = 28.75 * propi_sal(2) + 169.125

'Peso especifico producto

    pe_producto = 0.01375 * propi_sal(3) + 0.726625

'Azufre del producto

    azufre_producto = (0.25 * prin_sal(2) + 0.254) * 1.4

'delta T del proceso

    delta_T = (10# * prin_sal(3)) * 4.8

'Delta P del lecho catalítico

    delta_P = 4.625 * prin_sal(4) + 0.4375

End Sub

Public Sub Resultados()

    fMainForm.txtResultados(0).Text = produccion
    fMainForm.txtResultados(1).Text = tie_producto
    fMainForm.txtResultados(2).Text = tfe_producto
    fMainForm.txtResultados(3).Text = pe_producto
    fMainForm.txtResultados(4).Text = azufre_producto
    fMainForm.txtResultados(5).Text = delta_T
    fMainForm.txtResultados(6).Text = delta_P

End Sub
```