

**APLICACIÓN DE LA REGRESIÓN LINEAL MÚLTIPLE Y LAS REDES
NEURONALES ARTIFICIALES EN EL ANÁLISIS DEL PROCESO DE
HIDROTRATAMIENTO DE ACPM**

JUAN DAVID NIÑO LIZCANO

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISCOQUÍMICAS
ESCUELA DE INGENIERÍA QUÍMICA
BUCARAMANGA
2016**

**APLICACIÓN DE LA REGRESIÓN LINEAL MÚLTIPLE Y LAS REDES
NEURONALES ARTIFICIALES EN EL ANÁLISIS DEL PROCESO DE
HIDROTRATAMIENTO DE ACPM**

JUAN DAVID NIÑO LIZCANO

**Trabajo de grado presentado como requisito para optar por el título de
Ingeniero Químico**

Director:

GIOVANNI MORALES MEDINA, Ing. Qco., Dr.

Codirector:

OMAR ANDRÉS BENAVIDES PRADA, Ing. Qco.

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISCOQUÍMICAS
ESCUELA DE INGENIERÍA QUÍMICA
BUCARAMANGA
2016**

DEDICATORIA

A mis padres y a mi familia por su apoyo incondicional, por confiar en mis capacidades desde un principio. Este es un logro de todos, sin ustedes jamás hubiera alcanzado esta meta.

AGRADECIMIENTOS

Mis más sinceros agradecimientos:

*A mis padres por todos los esfuerzos realizados
para verme hecho un profesional.*

A mi abuelita Silvia por darme su apoyo en los momentos más difíciles.

*Al ingeniero, profesor, amigo y ahora colega Omar Andrés Benavides
por haber hecho parte me formación integral como profesional
y guía en este proceso académico terminado.*

*Al profesor Giovanni Morales por haber depositado su confianza en mí,
por su valioso apoyo, orientación y dedicación a este proyecto.*

*Por último y sin ser menos, a cada una de las personas que hicieron
parte de este trayecto, a todos ustedes amigos y compañeros de estudio.*

TABLA DE CONTENIDO

INTRODUCCIÓN	12
1. MARCO TEÓRICO	17
1.1. PROCESO DE HIDROTRATAMIENTO	17
1.1.1. Variables de proceso.....	19
1.2. REGRESIÓN LINEAL MÚLTIPLE (RLM)	19
1.3. REDES NEURONALES ARTIFICIALES (RNA)	20
1.3.1. Elementos que componen una red neuronal.....	21
2. METODOLOGÍA	23
3. ANÁLISIS DE RESULTADOS	27
3.1. Selección y análisis de datos	27
3.2. REGRESIONES LINEALES MÚLTIPLES	29
3.3. REDES NEURONALES ARTIFICIALES	33
3.3.1. RNA para el ACPM hidrotratado y su contenido de azufre	33
3.3.2. RNA para las condiciones de operación	36
3.3.3. RNA para la predicción de variables individuales.	40
4. CONCLUSIONES Y RECOMENDACIONES	42
REFERENCIAS	43
BIBLIOGRAFÍA	46
ANEXOS	51

LISTA DE FIGURAS

Figura 1. Reacciones de hidrodesulfuración presentes en el proceso de HDT.....	18
Figura 2. Estructura de una RNA Feedforward multicapa.....	21
Figura 3. Diagrama de bloques descripción metodología.	23
Figura 4. Gráfica de probabilidad normal de los residuos para el Caso II.....	31
Figura 5. Gráfica de probabilidad normal de los residuos para el Caso III.....	32
Figura 6. Variación del MSE con N durante la validación en la predicción del HC2 y S2 con redes FA-I.	35
Figura 7. Variación del R^2 con N durante la validación en la predicción del HC2 y S2 con redes FA-I.....	35
Figura 8. Modelo de RNA planteado por Arce <i>et al.</i>	36
Figura 9. Variables de entrada y salida, modelos de RNA propuestos por Lukec <i>et al.</i>	36
Figura 10. Variación de R^2 vs N en las redes FA-I para la predicción de H, dT y P.	38
Figura 11. Variación de R^2 vs N en las redes FA-II para la predicción de H, dT y P.	38
Figura 12. Variación de MSE vs N en las redes FA-I para la predicción de H, dT y P.	39
Figura 13. Variación de MSE vs N en las redes FA-II para la predicción de H, dT y P.	39
Figura 14. Funciones de activación habituales.	51
Figura 15. Diagrama de dispersión para el flujo de ACPM (HC1).....	52
Figura 16. Diagrama de dispersión para el punto final de ebullición (PFB).....	52
Figura 17. Diagrama de dispersión para el punto inicial de ebullición (PIB).	52
Figura 18. Diagrama de dispersión para grados API.	53
Figura 19. Diagrama de dispersión para el contenido de azufre en HC1 (S1).	53
Figura 20. Diagrama de dispersión para el flujo de hidrógeno (H).	53

Figura 21. Diagrama de dispersión para la relación ACPM/hidrógeno (RR).	54
Figura 22. Diagrama de dispersión para la temperatura de entrada (T1).	54
Figura 23. Diagrama de dispersión para el cambio de temperatura en la zona de reacción (dT).....	54
Figura 24. Diagrama de dispersión para la presión de operación en la zona de reacción (P).....	55
Figura 25. Diagrama de dispersión para el flujo de ACPM hidrotratado (HC2).	55
Figura 26. Diagrama de dispersión para el contenido de azufre en HC2 (S2).	55
Figura 27. Esquema RNA para la predicción del flujo de ACPM hidrotratado y su contenido de azufre.	60
Figura 28. Esquema RNA para la predicción del flujo de hidrógeno, cambio de temperatura y presión en la zona de reacción.	60
Figura 29. Esquema RNA para la predicción del flujo de hidrógeno y cambio de temperatura en la zona de reacción.....	61
Figura 30. Gráfica del R^2 vs en número de neuronas en la capa oculta para la predicción del H y dT. Redes neuronales con función de activación FA I.....	62
Figura 31. Gráfica del R^2 vs en número de neuronas en la capa oculta para la predicción el H y dT. Redes neuronales con función de activación FA II.....	62
Figura 32. Gráfica del MSE vs en número de neuronas en la capa oculta para la predicción del H y dT. Redes neuronales con función de activación FA I.....	63
Figura 33. Gráfica del MSE vs en número de neuronas en la capa oculta para la predicción del H y dT. Redes neuronales con función de activación FA II.....	63
Figura 34. Esquema RNA para la predicción del flujo del contenido de azufre en HC2.....	64
Figura 35. Esquema RNA para la predicción del flujo de ACPM hidrotratado.	64
Figura 36. Esquema RNA para la predicción de la relación ACPM/hidrógeno.....	65
Figura 37. Esquema RNA para la predicción del flujo hidrógeno.	65
Figura 38. Esquema RNA para la predicción del cambio de temperatura en la zona de reacción.....	66

LISTA DE TABLAS

Tabla 1. Resumen de estadísticos para cada variable en la ventana operativa. ...	27
Tabla 2. Resumen de estadísticos para cada variable en la ventana operativa sin atípicos.	28
Tabla 3. Matriz de correlación entre las variables.	29
Tabla 4. Valores de R^2 y $R^2_{ajustado}$ para cada caso de predicción.	30
Tabla 5. Valores de R^2 y MSE en la validación externa de las RLM.....	32
Tabla 6. Valores de R^2 y $R^2_{ajustado}$ para las RLM en función de HC1.	32
Tabla 7. Valores de R^2 y MSE en la validación externa de las RLM en función de HC1.....	33
Tabla 8. Estadísticos para las redes FA-I en la predicción del HC2 y S2.....	34
Tabla 9. Estadísticos para las redes FA-II en la predicción del HC2 y S2.....	34
Tabla 10. Desempeño de las RNA FA-I en la predicción de H, dT y P.....	37
Tabla 11. Desempeño de las RNA FA-II en la predicción de H, dT y P.....	38
Tabla 12. Validación de los modelos de red neuronal con funciones de activación FA I.	41
Tabla 13. Validación de los modelos de red neuronal con funciones de activación FA II.	41
Tabla 14. Resumen de las mejores RNA obtenidas.	41
Tabla 15. Resumen del ajuste de RLM para el contenido de azufre en el producto hidrotratado (S2). ($R^2= 0,2423$; $R^2_{ajustado}=0,2419$).....	57
Tabla 16. Resumen del ajuste de RLM para el flujo de ACPM hidrotratado (HC2).	57
Tabla 17. Resumen del ajuste de RLM para la relación ACPM/hidrógeno (RR). ..	58
Tabla 18. Resumen del ajuste de RLM para el flujo de hidrógeno (H).	58
Tabla 19. Resumen del ajuste de RLM para el cambio de temperatura en la zona de reacción (dT). ($R^2= 0,2582$; $R^2_{ajustado}=0,2578$).....	59

Tabla 20. Resumen del ajuste de RLM para el flujo de ACPM hidrotratado (HC2) en función de HC1. ($R^2= 0,9352$; $R^2_{ajustado}=0,9352$).....	59
Tabla 21. Resumen del ajuste de RLM para la relación ACPM/hidrógeno (RR) en función de HC1. ($R^2= 0,8934$; $R^2_{ajustado}=0,8924$).....	59
Tabla 22. Resultados modelos de redes neuronales con las funciones de activación FA I para la predicción de H y dT.....	61
Tabla 23. Resultados modelos de redes neuronales con las funciones de activación FA II para la predicción de H y dT.....	62

LISTA DE ANEXOS

Anexo A. Funciones típicas de transferencia o activación.	51
Anexo B. Gráficas de dispersión de las variables del proceso de HDT.	52
Anexo C. Método de normalización Min-Max.....	56
Anexo D. Detalles calibración modelos de RLM.	57
Anexo E. Detalles calibración modelos de RLM en función de HC1.....	59
Anexo F. Esquemas de RNA para la predicción de variables en conjunto.	60
Anexo G. Resultados RNA para las condiciones de operación H y dT.....	61
Anexo H. Esquemas RNA para la predicción de variables individuales.....	64

RESUMEN

TÍTULO: APLICACIÓN DE LA REGRESIÓN LINEAL MÚLTIPLE Y LAS REDES NEURONALES ARTIFICIALES EN EL ANÁLISIS DEL PROCESO DE HIDROTRATAMIENTO DE ACPM.*

AUTOR: JUAN DAVID NIÑO LIZCANO.**

PLABRAS CLAVE: REGRESIÓN LINEAL MÚLTIPLE, REDES NEURONALES ARTIFICIALES, DIÉSEL, HIDROTRATAMIENTO, HIDRODESULFURACIÓN.

DESCRIPCIÓN:

El desarrollo de modelos de predicción con regresión lineal múltiple (RLM) basados en datos históricos tiene múltiples ventajas. Entre ellas la posibilidad de: establecer la influencia de las variables en los procesos, determinar los estadísticos que presentan los datos, análisis de datos atípicos e influyentes y predicción de nuevos valores. Por otra parte, Las redes neuronales artificiales (RNA) son sistemas de procesamiento, que emulan de manera esquemática la estructura neuronal del cerebro, logrando reproducir sus capacidades en la solución de problemas. Estas redes son útiles en aquellos casos en los cuales el problema no posee una solución computacional precisa o se requiere de algoritmos complejos y extensos

En el presente trabajo se desarrollaron modelos RLM y RNA para la predicción de las variables de salida y de operación del proceso de hidrotreatmento de ACPM, los procedimientos matemáticos fueron aplicados en la versión libre de prueba del *software* estadístico Minitab 17 y en el *toolbox* respectivo de MATLAB. Para el desarrollo de cada uno de los modelos fueron usados datos históricos del proceso de hidrotreatmento correspondientes a la unidad U-4700 de la refinería de Ecopetrol en Barrancabermeja. Los resultados de este trabajo demuestran la no linealidad de las variables en el proceso, siendo los modelos de RNA los indicados para la predicción de las variables de salida y de operación.

* Trabajo de grado

** Facultad de Ingenierías Físicoquímicas. Escuela de Ingeniería Química. Director: Giovanni Morales Medina, Dr. Ing. Químico

ABSTRACT

TITLE: APLICATION OF THE MULTIPLE LINEAR REGRESSION AND ARTIFICIAL NEURAL NETWORKS IN THE ANALYSIS OF ACPM HYDROTREATING PROCESS.*

AUTHOR: JUAN DAVID NIÑO LIZCANO.**

KEY WORDS: MULTIPLE LINEAR REGRESSION; ARTIFICIAL NEURAL NETWORKS; DIESEL; HYDROTREATMENT; HYDRODESULFURIZATION.

DESCRIPTION OR CONTENT:

The development of predictive models with multiple linear regression (MLR) based on historical data has multiple advantages. We can count among them the possibility of establishing the influence of the variables in the processes, determining the statistics data presented, analyzing the outlier and influential data, and predicting new values for the properties. Moreover, Artificial Neural Networks (ANN) are processing systems that emulate schematically the neuronal structure of the brain, reproducing its skills in solving problems. These networks are useful in those cases where the problem does not have a precise computational solution, or complex and extensive algorithms are required.

In the present study, MLR and ANN models were developed for the prediction of the outcome and operational variables of the hydrotreatment process of Diesel Fuel (derived from petroleum), the mathematical procedures were performed using the free trial version of the statistics software *Minitab 17* and the corresponding toolbox of *MATLAB*. For the development of the models, historical data of the hydrotreatment process corresponding the U-4700 of the Barrancabermeja's Ecopetrol refinery were used. The results of this work demonstrate the nonlinearity of the variables in the process, being the ANN models the more suitable for the prediction of the outcome and operational variables.

* Degree Work

** Faculty of Phisico-Chemical Engineering. School of Chemical Engineerign. Director: Giovanni Morales Medina, Dr. Ing.Chemical.

INTRODUCCIÓN

El uso de combustibles con altos contenidos de azufre conlleva a la generación de gases contaminantes para el medio ambiente. Con las regulaciones ambientales actuales, es necesario implementar procesos donde sea viable remover la mayor cantidad de azufre posible, de tal forma que el producto final se ajuste a las leyes de cada país. En este contexto, el hidrotratamiento es un proceso utilizado en la industria del petróleo con el fin de remover principalmente compuestos azufrados como mercaptanos, sulfuros y disulfuros, entre otros [1] [2]. Dicho tratamiento utiliza el hidrógeno en presencia de un catalizador para remover el azufre presente en las cargas de crudos pesados y ACPM [2].

En el hidrotratamiento, al igual que en cualquier otro proceso, es de gran importancia disponer de modelos matemáticos que permitan realizar un seguimiento continuo de sus variables, con el fin de implementar acciones de control que conlleven a la consecución de los estándares de calidad requeridos. Los paquetes de simulación de procesos, tales como Aspen Plus, Aspen Hysys y Pro II corresponden a una de las opciones disponibles a nivel comercial para el seguimiento del proceso [3].

Alaei & Nariman [4] mediante el *software* Aspen HYSYS simularon una unidad de hidrotratamiento de gases, en la refinería de South Pars en Irán, con el objetivo de obtener la combinación de variables que permitieran predecir los contenidos de azufre en el producto hidrotratado. El uso de simuladores comerciales tiene la desventaja del conocimiento previo de los parámetros específicos para cada proceso, como es el caso de los parámetros cinéticos de reacción, los cuales deben ser obtenidos mediante pruebas industriales o de plantas piloto. Por otra parte, los modelos matemáticos pueden ser desarrollados con base en datos a escala piloto o industrial. Menderos-Nieto [5] y Rodríguez- Huerta [6] soportaron sus modelos matemáticos para el proceso de hidrotratamiento a escala piloto para

predecir la información de salida del reactor. Asimismo, Álvarez *et al.* [3] basados en datos de planta piloto propusieron una cinética para el hidrotreatmento de mezclas de diésel, aceite liviano de ciclo y jet. De otro lado, Arce & Paz [1] entrenaron redes neuronales artificiales en la regresión de variables de proceso de hidrotreatmento en la refinería mexicana PEMEX.

El uso de redes neuronales en comparación con los paquetes de simulación comerciales, tiene la ventaja de requerir sólo datos históricos del proceso y definir una buena arquitectura de red. Rodríguez & Ancheyta [7] y Mederos & Ancheyta [8] plantearon modelos de simulación para el hidrotreatmento de gasóleos de vacío, donde tuvieron en cuenta la transferencia de masa de hidrogeno y de sulfuro en la fase gaseosa. En estos trabajos se aprecia el poder predictivo de las redes neuronales y su aplicación en los procesos industriales.

En este orden de ideas, el presente documento analiza el desempeño de las regresiones basadas en datos históricos de proceso y en las redes neuronales artificiales para el proceso de hidrotreatmento de ACPM en la refinería de Barrancabermeja; también, se comparan algunos modelos obtenidos utilizando regresión lineal múltiple. Las regresiones se dirigieron a la predicción del contenido de azufre en el ACPM hidrotreatado, el flujo de ACPM hidrotreatado, la relación ACPM/hidrogeno, el flujo de hidrogeno, el cambio de temperatura en la zona de reacción y la temperatura de entrada a la misma. Los distintos modelos predictivos fueron desarrollados utilizando el *toolbox* de Matlab R2015a y la versión de prueba del *software* estadístico Minitab 17.

1. MARCO TEÓRICO

1.1. PROCESO DE HIDROTRATAMIENTO

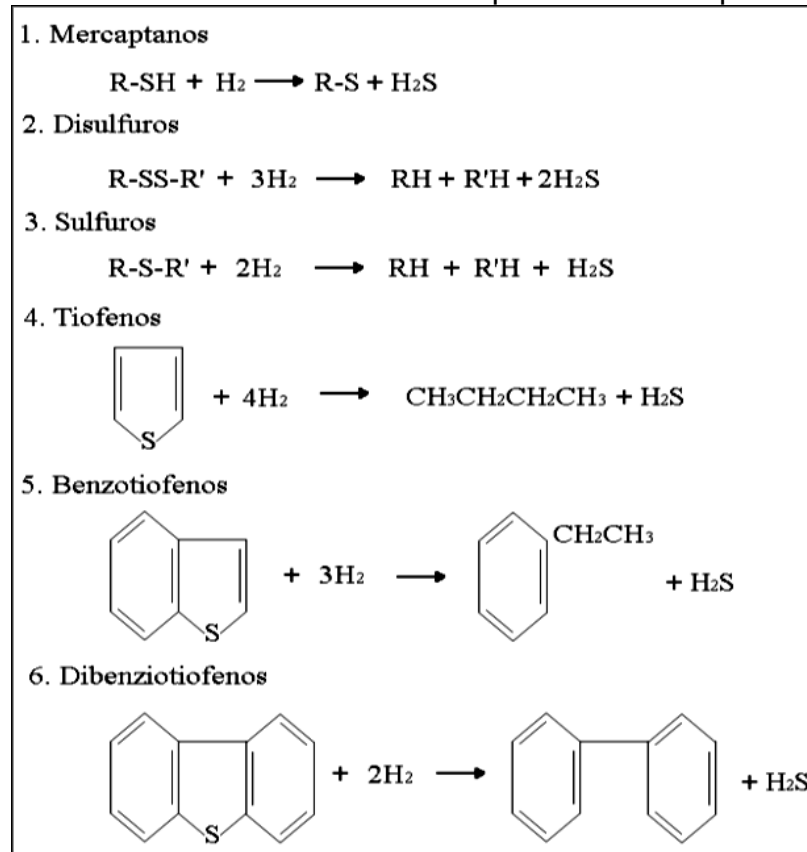
El proceso de hidrotratamiento (HDT) consiste en la hidrogenólisis de compuestos sulfurados, nitrogenados y oxigenados a partir de moléculas de hidrocarburos con heteroátomos como los mercaptanos y las porfirinas entre otros; también, en el HDT se hidrogenan olefinas y compuestos aromáticos [6]. Este proceso se efectúa para mejorar la calidad del petróleo crudo o de sus fracciones. Industrialmente, el hidrotratamiento se lleva a cabo en reactores tubulares de lecho fijo, donde las fracciones del petróleo reaccionan con el hidrogeno en presencia de un catalizador, en condiciones severas de presión y temperatura. Generalmente los catalizadores son utilizados con formulaciones de metales de transición tales como molibdeno, wolframio, cobalto, níquel y/o hierro, ya que presentan una actividad elevada frente a otros compuestos [3] [6].

El proceso de hidrotratamiento de ACPM en la refinería de Barrancabermeja de ECOPETROL se lleva a cabo en la unidad U-4700, la cual está constituida por las secciones de carga, compresión y reciclo, reacción, depuración de H₂ con DEA, despojo de nafta y la sección de depuración de H₂. La sección de reacción consta de dos reactores de lechos empacados con catalizadores cobalto-molibdeno soportados sobre una alúmina. Dicha sección está diseñada para operar bajo condiciones de temperatura entre 725 y 800 °F y presión entre 840 y 950 psig, con carga promedio de 56000 barriles por día de ACPM con contenido de azufre entre 3000 y 6000 ppm. El contenido de azufre en el producto hidrotratado presenta valores inferiores a 50 ppm.

En el hidrotratamiento se llevan a cabo reacciones de hidrodesaromatización (HDA), hidrodeshidrogenación (HDN), hidrodeshidrometalización (HDM), hidrodeshidroxigenación (HDO) e hidrodeshidrosulfuración (HDS) dependiendo de la composición del flujo de entrada. Para el hidrotratamiento del diésel, la reacción

principal corresponde a la hidrodesulfuración, a partir de la cual los compuestos con presencia de azufre (Figura 1), reaccionan con el hidrogeno de forma exotérmica e irreversible para generar un diésel de bajo contenido de azufre [6][9].

Figura 1. Reacciones de hidrodesulfuración presentes en el proceso de HDT.



Fuente: Rodríguez Huerta (2007)

El contenido de azufre es un parámetro de control en la calidad del diésel comercial, debido a las emisiones de óxidos de azufre (SO_x) generados durante la combustión, los cuales se transforman en ácidos al entrar en contacto con el vapor de agua concentrado en la troposfera [1]. Por esta razón, la ley colombiana en la resolución 182087/07 y en la más reciente 90963/14, en relación con los criterios de calidad de los biocombustibles para su uso en motores diésel como componente de la mezcla con el combustible diésel de origen fósil en procesos de combustión, establece que el contenido de azufre permitido en el diésel comercial debe estar por debajo de 50 ppm [10].

1.1.1. Variables de proceso.

Dentro de las variables que tienen influencia sobre el proceso de hidrotratamiento se encuentran temperatura, presión parcial de hidrogeno, relación gas/hidrocarburo y la velocidad espacial del líquido. En cuanto a las variables de temperatura, presión parcial de hidrogeno y la relación de gas/hidrocarburo, a medida que sus valores incrementan, favorecen notablemente las reacciones de HDS, debido a la disolución del sulfuro de hidrogeno, el cual es un inhibidor de la reacción, permitiendo de esta forma obtener mejores productos hidrotratados [11-13]. De otro lado el incremento en la temperatura no deber ser excesivo, ya que puede generar daños en el catalizador. Por otra parte, la velocidad espacial del líquido que está definida como la relación entre el flujo volumétrico del alimento y el volumen de catalizador, es inversamente proporcional al tiempo de residencia; lo anterior implica que valores bajos en la velocidad espacial repercuten en tiempos de residencia elevados, favoreciendo las reacciones de HDS [9][13][14].

1.2. REGRESIÓN LINEAL MÚLTIPLE (RLM)

Un modelo de RLM, es un modelo matemático simplificado que puede explicar las interacciones existentes entre un grupo determinado de variables independientes o factores con una variable dependiente [15]. En la ecuación (1), y es la variable dependiente o variable a predecir, β_i son el conjunto de coeficientes de regresión de las variables predictoras x_i y ε es el error presente entre la predicción y el valor real.

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \varepsilon = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_{k-1} x_{k-1} + \beta_k x_k + \varepsilon \quad (1)$$

La RLM parte de la suposición de que el error (ε) es independiente, posee una distribución normal, tiene una media cero y una varianza en común para cualquier valor de x_i [16]. La ecuación (1) sin el término del error (ε) representa la ecuación de predicción de un modelo RLM. En la RLM se minimiza la suma de los

cuadrados de las diferencias entre la variable de respuesta y la predicción obtenida. El resultado de la minimización reporta los valores de los coeficientes de regresión según [17]:

$$\beta = (X'X)^{-1}X'y \quad (2)$$

Donde X y y son la matriz de datos iniciales, X' es la matriz transpuesta de X y β es el vector de coeficientes de regresión. La significancia del modelo (*i.e.* los supuestos detrás de la RLM) pueden ser verificados por medio de un análisis de varianza (ANOVA) y el análisis de los residuos [16].

El coeficiente de determinación R^2 establece el porcentaje de predicción de la varianza de la variable dependiente utilizando los predictores seleccionados; este coeficiente se expresa según:

$$R^2 = 1 - \frac{SSR}{SST} \quad (3)$$

Donde SSR es la suma de los cuadrados de los errores y SST es la varianza total de los datos. El coeficiente de determinación no está en función de la cantidad de variables presentes en el modelo; por esta razón, puede presentar un valor incorrecto debido a la adición de más variables predictoras. La corrección del coeficiente de determinación considerando el número de variables independientes conlleva al coeficiente ajustado [17][18].

$$R^2_{ajustado} = \bar{R}^2 = 1 - \frac{SSR/(n-k-1)}{SST/(n-1)} \quad (4)$$

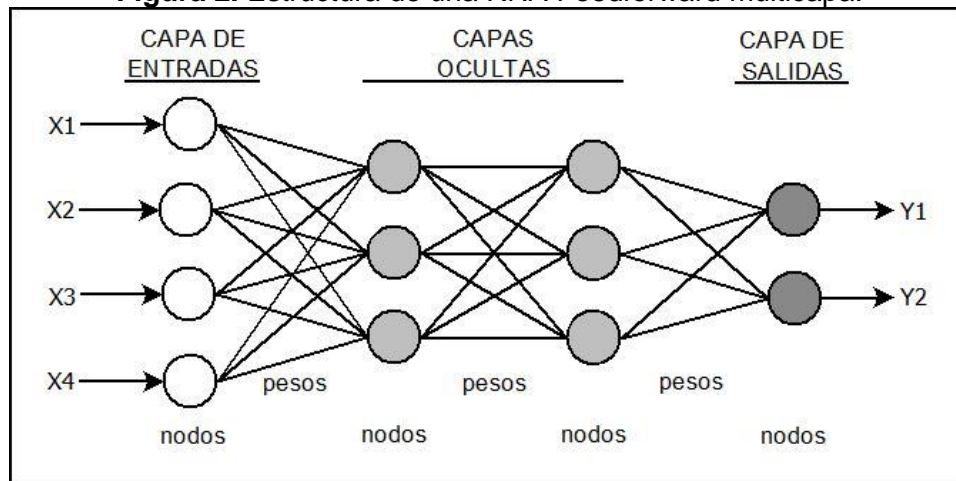
1.3. REDES NEURONALES ARTIFICIALES (RNA)

Las RNA son sistemas de procesamiento, que emulan de manera esquemática la estructura neuronal del cerebro, logrando reproducir sus capacidades en la

solución de problemas. Así como el cerebro humano aplica el conocimiento adquirido a través de la experiencia y la práctica, una RNA toma problemas resueltos y con ellos forma un sistema con el cual toma decisiones y realiza clasificaciones para dar solución a un determinado problema. Estas redes son útiles en aquellos casos en los cuales el problema no posee una solución computacional precisa o se requiere de algoritmos complejos y extensos [19] [20].

Como se puede observar en la Figura 2, las RNA están constituidas por un conjunto de neuronas o nodos interconectados a través de señales con pesos, los cuales corresponden a la intensidad de los enlaces sinápticos entre neuronas; cada nodo posee una función de activación que le permite cambiar de nivel de activación dependiendo si la totalidad de la información recibida supera el umbral interpuesto de la función de activación [21].

Figura 2. Estructura de una RNA Feedforward multicapa.



1.3.1. Elementos que componen una red neuronal

Las RNA se componen de neuronas de entrada y salida, las cuales se encargadas de recibir la información proveniente del exterior y de entregar la información arrojada por el sistema, respectivamente. En el intermedio de las capas de entrada y salida se encuentran las neuronas de procesamiento (capas ocultas), las cuales se encargan de transmitir la información a elementos internos para su

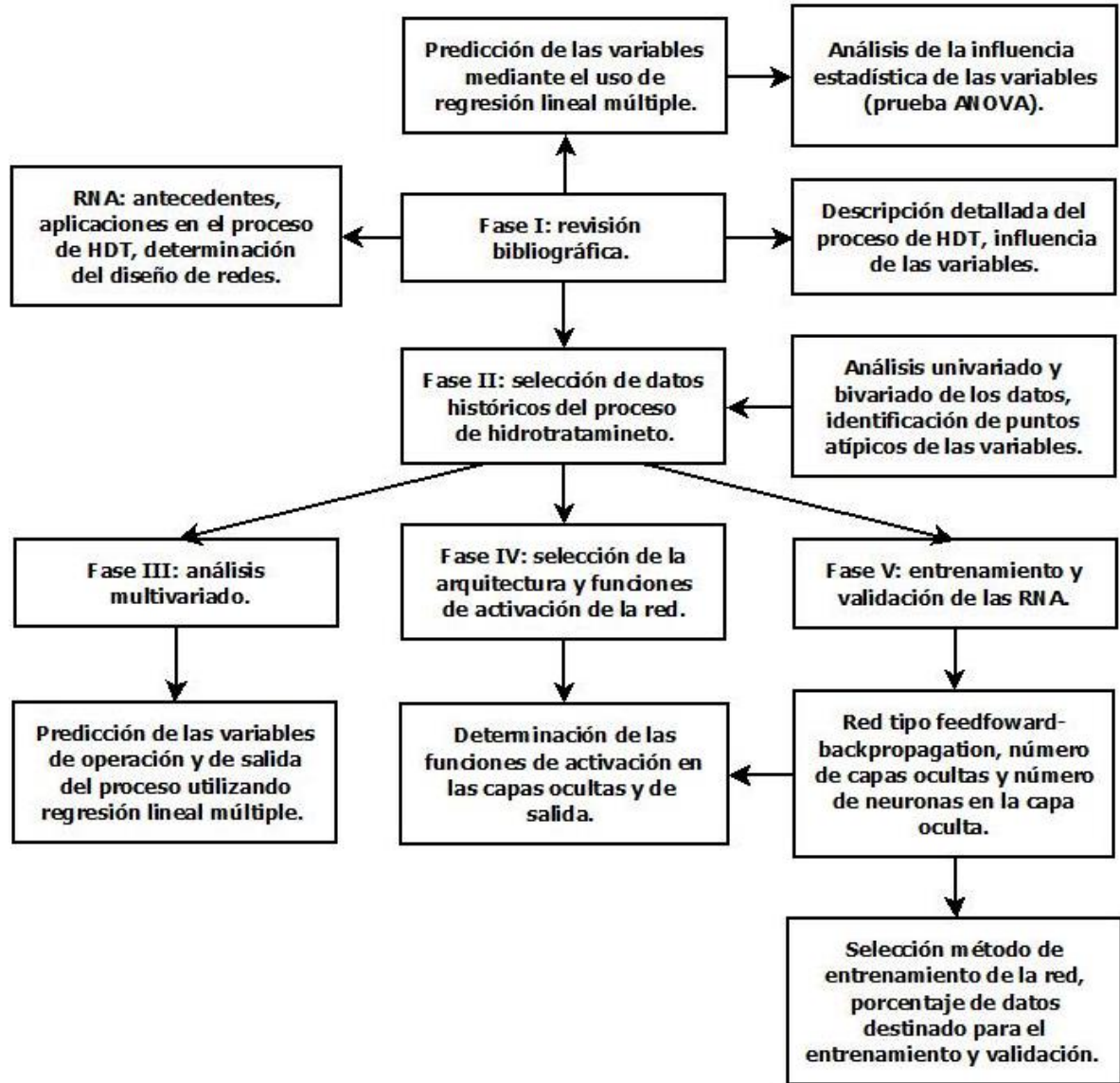
procesamiento [19]. Entre los nodos o neuronas existe una serie de conexiones asociadas a una magnitud llamada peso. Determinada neurona recibe una cantidad de información de entrada, cada señal de entrada se multiplica por el peso asociado a cada una de las conexiones. Posteriormente se suma las entradas ponderadas con los pesos y se aplica la función de activación para generar la señal de salida. Estas funciones de activación deben ser continuas y derivables; las funciones más utilizadas se pueden detallar en el Anexo A [19-21].

El proceso de aprendizaje o entrenamiento consiste en hallar los pesos que codifican los conocimientos. Uno de los algoritmos de aprendizaje más utilizados es el denominado *backpropagation*, el cual es una generalización del procedimiento de mínimos cuadrados, donde se actualizan los pesos con base en el error cuadrático promedio. En este tipo de aprendizaje se requiere conocer o disponer de un conjunto de instrucciones de entrenamiento que describa a cada una de las entradas a la red y sus respectivos valores de salida [20][21].

2. METODOLOGÍA

La siguiente figura ilustra la metodología seguida durante el desarrollo del presente trabajo:

Figura 3. Diagrama de bloques descripción metodología.



FASE I: se desarrolló una consulta bibliográfica acerca de los análisis estadísticos univariado, bivariado y métodos de predicción con RLM. Además, se buscaron

antecedentes para la simulación del proceso de hidrotratamiento tanto para los métodos de predicción con RLM y RNA como para modelos matemáticos y paquetes comerciales. Por otra parte, se definieron las variables que tienen influencia sobre las reacciones de HDS en el proceso de hidrotratamiento y las bases para establecer las arquitecturas de RNA desarrolladas en el presente trabajo.

Fase II: los datos históricos de planta utilizados corresponden a registros minuto a minuto del proceso de hidrotratamiento de la unidad U-4700 de la refinería de Barrancabermeja. Estos datos conforman una matriz de 23400x12 datos, cuyas columnas corresponden a: el flujo de ACPM con alto contenido de azufre y algunas de sus propiedades como el punto inicial de ebullición, punto final de ebullición, grado API, y el contenido de azufre; también se consideraron ciertas condiciones de operación en la zona de reacción como presión de operación, temperatura de entrada y cambio de temperatura en la zona de reacción, la relación hidrógeno/hidrocarburo y el flujo de hidrógeno alimentado, además del flujo de diésel hidrotratado y su contenido de azufre.

Se aplicaron diferentes análisis estadísticos univariados y bivariados a los datos de proceso; estos condujeron a la obtención de los estadísticos valores mínimos, máximos y promedios, desviaciones estándar y coeficientes de correlación. A su vez, se realizaron las gráficas de dispersión con las cuales se observó el comportamiento de las variables y sus interacciones entre sí. Por otra parte, mediante el análisis estadístico de caja y bigotes (*box-plot*) se tomaron como valores atípicos aquellos que se encontraban fuera del rango. Las condiciones de operación atípicas fueron analizadas para determinar la eliminación de los puntos respectivos de la base de datos [15] [24].

Fase III: se realizaron las siguientes actividades con el fin de obtener modelos de regresión lineal múltiple para la predicción del flujo de diésel hidrotratado, su contenido de azufre y las condiciones de operación en la zona de reacción.

- 1. Pretratamiento de datos:** todas las variables se escalaron a un rango entre 0,1 y 0,9 mediante la normalización min-max [25], para garantizar órdenes de magnitud similares.
- 2. Planteamiento de los modelos de regresión múltiple:** para la calibración de los modelos RLM se utilizaron el 70% de los datos. Inicialmente, se consideraron la totalidad de las variables para cada caso de predicción; cada modelo fue analizado utilizando la tabla ANOVA, estableciendo un intervalo de confianza del 95% para las pruebas de valor-p. Mediante los criterios de R^2 y $R^2_{ajustado}$ se determinó el desempeño de los modelos.
- 3. Validación modelos de regresión lineal múltiple:** la validación se realizó con un conjunto externo conformado por el 30% de los datos históricos restantes.

Fase IV: El entrenamiento y la validación de las RNA fue efectuado utilizando la herramienta de *Toolbox nntool* de Matlab R2015a (*The MathWorks, inc.*), a través de las siguientes actividades.

- 1. Diseño y arquitectura de las redes neuronales:** las RNA fueron tipo *feed-forward Backpropagation* con funciones de activación *logsig* y arquitecturas de red con una capa oculta de la forma (X:N:Y), donde X, N y Y hacen referencia al número de neuronas en la capa de entrada, capa oculta y capa de salida, respectivamente. Estas arquitecturas para las RNA fueron propuestas por Arce & Paz [1] y Lukec & Sertic [26] en sus análisis relacionados con el proceso de hidrotatamiento. El desempeño de la arquitectura de la red fue analizado cambiando diferentes funciones de activación en la capa oculta y en la capa de salida, utilizando las funciones *logsig-logsig* y *logsig-purelin*. Asimismo, el número de neuronas fue alterado hasta encontrar el mejor modelo de red, siguiendo la restricción para el número de neuronas en la capa oculta (*i.e.* el número de neuronas en las capas ocultas no debe superar al doble del número de variables de entrada) [27][28].

2. **Entrenamiento de la red neuronal:** para la función de entrenamiento y adaptación, se utilizó la función de Matlab denominada *trainlm*, considerando el 70% de los datos. El algoritmo de esta función actualiza los pesos y las ganancias de acuerdo a la optimización de Levenberg-Marquardt; este algoritmo de entrenamiento es el más rápido para redes tipo *Backpropagation* [4].
3. **Validación de las redes neuronales:** las RNA entrenadas fueron validadas utilizando el 30% de los datos históricos restantes. Los parámetros R^2 y el error cuadrático promedio (*MSE*) de entrenamiento y de validación fueron determinados para establecer la mejor RNA para el proceso de hidrotratamiento.

3. ANÁLISIS DE RESULTADOS

3.1. Selección y análisis de datos

Los datos históricos del proceso HDT, unidad U-4700, fueron seleccionados en una ventana operativa desde el 07 de diciembre a la hora 00:00:00 hasta el 23 de diciembre a la hora 02:25:00, cada minuto. Los estadísticos para cada variable en la ventana operativa son presentados en la Tabla 1. Asimismo, las gráficas de dispersión se presentan en el Anexo B. A partir del análisis univariado, es posible inferir la presencia de valores que se desviaban del conjunto de datos típicos para cada variable. Para la presión, los posibles datos atípicos pueden deberse a variaciones en los flujos de ACPM e hidrógeno alimentados a la zona de reacción, los datos atípicos no fueron eliminados debido a la posibilidad de operación estable de la unidad HDT con diferentes valores de flujo de ACPM e hidrogeno.

Tabla 1. Resumen de estadísticos para cada variable en la ventana operativa.

SIMBOLO	DESCRIPCIÓN DE LAS VARIABLE	UNIDADES	VALOR MÍNIMO	VALOR PROMEDIO	VALOR MÁXIMO	DESVIACIÓN ESTÁNDAR
HC1	Flujo de ACPM con alto contenido de azufre	$\frac{\text{bls}}{\text{min}}$	31853.660	52627.826	64941.879	5517.487
PIB	Temperatura inicial de ebullición	°F	719.100	737.095	759.700	8.275
PIF	Temperatura final de ebullición	°F	333.500	390.025	414.900	19.966
API	Densidad relativa para líquidos del petróleo	°API	28.400	29.720	30.899	0.383
S1	Concentración de azufre en HC1	ppm	3985.000	4786.103	5603.000	349.327
H	Hidrógeno alimentado en la entrada a la zona de reacción	$\frac{\text{mol}}{\text{min}}$	105.532	126.336	142.636	5.864
RR	Relación ACPM / hidrógeno	$\frac{\text{bls}}{\text{mol}}$	233.424	418.359	560.212	54.896
T1	Temperatura de entrada a la zona de reacción	°F	577.521	636.000	669.330	3.521
dT	Cambio de temperatura en la zona de reacción	°F	5.569	41.941	100.777	1.921
P	Presión de operación en la zona de reacción	psia	675.511	684.918	688.478	0.677
HC2	Flujo de ACPM hidrotratado	$\frac{\text{bls}}{\text{min}}$	26942.305	47578.168	60366.645	5473.037
S2	Concentración de azufre en HC2	ppm	5.000	26.923	50.000	10.623

Por otra parte, para la temperatura y el cambio de temperatura (Anexo B Figuras 22 y 23), la presencia de datos atípicos pudo ser ocasionada por problemas en los sensores de medición. Por esta razón, mediante el método de cajas y bigotes (*box-plot*) los puntos que se encontraban por fuera de los valores máximos extremos fueron eliminados. Después de esta reducción, se obtuvo una matriz de 23179 filas y 12 columnas, cuyos estadísticos se encuentran resumidos en la Tabla 2, en donde se pueden apreciar los valores máximos y mínimos, el valor promedio y la desviación estándar para cada una de las variables consideradas.

Tabla 2. Resumen de estadísticos para cada variable en la ventana operativa sin atípicos.

SIMBOLO	DESCRIPCIÓN DE LAS VARIABLE	UNIDADES	VALOR MÍNIMO	VALOR PROMEDIO	VALOR MÁXIMO	DESVIACIÓN ESTÁNDAR
HC1	Flujo de ACPM con alto contenido de azufre	$\frac{\text{bls}}{\text{min}}$	31853.660	52627.826	64941.879	5517.487
PIB	Temperatura inicial de ebullición	°F	719.100	737.095	759.,700	8.275
PIF	Temperatura final de ebullición	°F	333.500	390.025	414.900	19.966
API	Densidad relativa para líquidos del petróleo	°API	28.400	29.720	30.899	0.383
S1	Concentración de azufre en HC1	ppm	3985.000	4786.103	5603.000	349.327
H	Hidrógeno alimentado en la entrada a la zona de reacción	$\frac{\text{mol}}{\text{min}}$	105.532	126.336	142.636	5.864
RR	Relación ACPM / hidrógeno	$\frac{\text{bls}}{\text{mol}}$	233.424	418.359	560.212	54.896
T1	Temperatura de entrada a la zona de reacción	°F	623.386	636.005	653.533	3.424
dT	Cambio de temperatura en la zona de reacción	°F	36.112	41.990	47.910	1.596
P	Presión de operación en la zona de reacción	psia	675.511	684.918	688.478	0.677
HC2	Flujo de ACPM hidrotratado	$\frac{\text{bls}}{\text{min}}$	26942.305	47578.168	60366.645	5473.037
S2	Concentración de azufre en HC2	ppm	5.000	26.923	50.000	10.623

Por otra parte, aunque diferentes reportes incluyen la variable velocidad espacial del líquido [12][13], ésta no fue tomada en cuenta en el presente trabajo debido a que el volumen de catalizador en la zona de reacción permaneció constante durante la ventana operativa; la velocidad espacial del líquido en la unidad HDT (U-4700) está representada por el flujo de ACPM de entrada. Asimismo, la variable ACPM/hidrógeno (RR) no fue tomada en cuenta entre las variables de entrada o

predictoras, ya que está altamente correlacionada con el flujo de ACPM (HC1), como se puede observar en la matriz de correlación (Tabla 3).

Tabla 3. Matriz de correlación entre las variables.

	HC1	PFB	PIB	API	S1	H	RR	T1	dT	P	HC2	S2
HC1	1											
PFB	0,267	1										
PIB	0,072	0,082	1									
API	-0,015	-0,382	-0,303	1								
S1	-0,195	0,079	-0,166	-0,225	1							
H	-0,440	-0,01	0,293	-0,34	0,095	1						
RR	0,947	0,204	-0,046	0,103	-0,191	-0,703	1					
T1	0,523	0,483	0,108	-0,334	0,053	-0,062	0,439	1				
dT	0,374	-0,017	0,117	-0,034	-0,114	-0,032	0,305	-0,043	1			
P	-0,032	-0,013	-0,065	0,030	0,008	0,029	-0,037	-0,043	0,074	1		
HC2	0,967	0,257	0,110	-0,065	-0,182	-0,432	0,919	0,541	0,358	-0,032	1	
S2	0,217	0,187	-0,293	0,068	0,264	-0,178	0,22	0,176	-0,001	-0,022	0,218	1

3.2. REGRESIONES LINEALES MÚLTIPLES

Para el desarrollo de las regresiones lineales, se normalizaron los datos en un intervalo de 0,1 y 0,9 mediante el método min-max [25] (descrito en el Anexo C), y posteriormente, diversos datos fueron seleccionados en forma aleatoria para conformar los conjuntos de calibración y de validación. Se usaron un total de 16225 datos de proceso para conformar la matriz de calibración (70% de la matriz total).

Los casos de regresión establecidos corresponden:

- I. Predicción del contenido de azufre en el ACPM hidrotratado (S2).
- II. Predicción del flujo de ACPM hidrotratado (HC2).
- III. Predicción de la relación hidrocarburo / hidrógeno (RR).
- IV. Predicción del flujo de hidrógeno (H).
- V. Predicción del cambio de temperatura en la zona de reacción (dT).

Los modelos de RLM fueron ajustados utilizando la versión de prueba del *software* estadístico *Minitab* 17, mediante el método de regresión por eliminación hacia atrás (*Backward*), con un intervalo de confianza del 95%. Los resultados iniciales obtenidos se encuentran resumidos en la Tabla 4 y en el Anexo D, donde se presentan las regresiones y los valores de R^2 y $R^2_{ajustado}$ respectivos.

Como se aprecia en la Tabla 4, el caso II y el caso III presentan un coeficiente de determinación elevado, lo cual indica que el azufre en el ACPM hidrotratado y el flujo de ACPM hidrotratado pueden ser predichos por las variables independientes establecidas; las regresiones se presentan en las ecuaciones 5 y 6. Cabe resaltar que el programa estadístico *Minitab* 17, entrega como resultado modelos de regresión en los cuales ya se han eliminado las variables que no tienen significancia (valor- $p \geq 0.05$).

Tabla 4. Valores de R^2 y $R^2_{ajustado}$ para cada caso de predicción.

Caso de estudio	R^2	$R^2_{ajustado}$
I	0.2423	0.2419
II	0.9426	0.9426
III	0.9186	0.9185
IV	0.3995	0.3992
V	0.2584	0.2580

$$\mathbf{HC2} = 0,0660 + 0,8950\mathbf{HC1} - 0,0315\mathbf{PFB} + 0,0272\mathbf{PIB} - 0,0622\mathbf{API} - 0,0039\mathbf{S1} - 0,0605\mathbf{H} + 0,0945\mathbf{T1} + 0,0141\mathbf{dT} + 0,0127\mathbf{P} \quad (5)$$

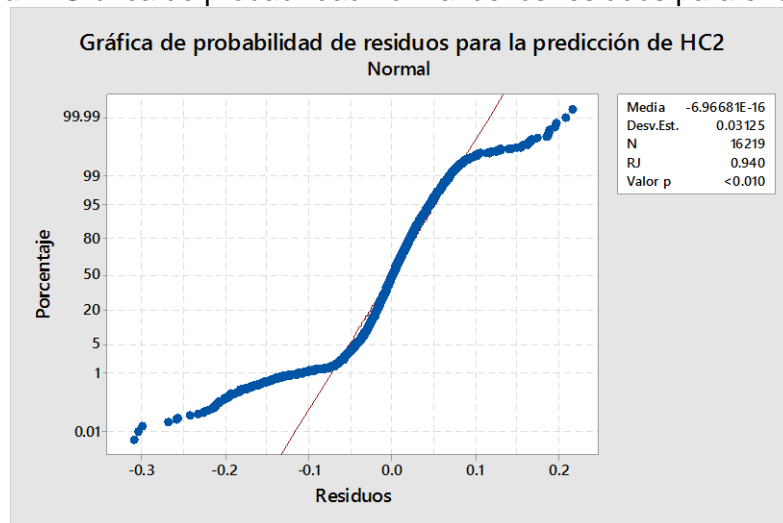
$$\mathbf{RR} = 0,0335 + 1,0250\mathbf{HC1} - 0,0605\mathbf{PIB} + 0,0825\mathbf{API} + 0,0080\mathbf{S1} - 0,0875\mathbf{T1} - 0,0802\mathbf{dT} - 0,0312\mathbf{P} - 0,0194\mathbf{S2} \quad (6)$$

Al encontrar valores significativos de R^2 y $R^2_{ajustado}$ en los modelos de predicción para los casos II y III, fue necesario desarrollar una prueba de normalidad para confirmar los supuestos en las RLM obtenidas [16]. Por esta razón, mediante la

versión de prueba del *software* Minitab 17, se realizó la prueba de normalidad de Ryan-Joiner [29], la cual establece que el valor-p debe ser mayor a 0,01 para asegurar la normalidad de los residuos.

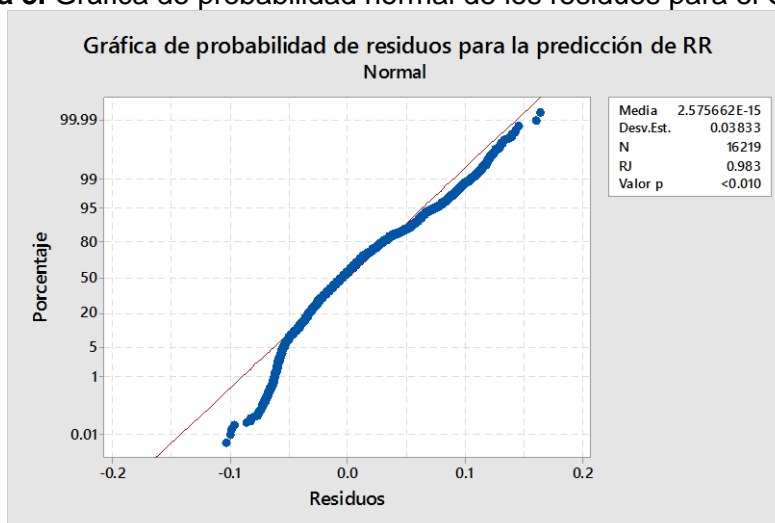
Como se observa en las Figuras 4 y 5, los residuos no poseen una distribución normal y la prueba de normalidad de Ryan-Joiner indica un valor-p menor a 0,01; lo cual invalida la suposición de normalidad de los residuos. Aunque los modelos de regresión no cumplen con el supuesto de normalidad, al realizar la validación externa con 30% de los datos restantes, se encontró que poseen la capacidad de predecir las variables independientes dentro del rango del conjunto de datos de calibración (Tabla 5). No se recomienda el uso de estos modelos para extrapolaciones debido a que no cumplen el supuesto de normalidad.

Figura 4. Gráfica de probabilidad normal de los residuos para el Caso II.



Fuente: obtenido con la versión de prueba de Minitab 17, Autor.

Figura 5. Gráfica de probabilidad normal de los residuos para el Caso III.



Fuente: obtenido con la versión de prueba de Minitab 17, Autor.

Tabla 5. Valores de R^2 y MSE en la validación externa de las RLM

Caso de estudio	R^2	MSE
II	0.9477	0.0526
III	0.9208	0.0014

Por otra parte, como se observa en las ecuaciones 5 y 6 la variable del flujo de ACPM (HC1) tiene mayor peso en los coeficientes de regresión en comparación con las otras. Por lo tanto, se replantearon las RLM para los casos II y III. En las Tabla 6 y 7, y el Anexo E se encuentran los estadísticos R^2 , $R^2_{ajustado}$ y MSE para la calibración y validación externa, y los ajustes de las RLM. Se puede observar que los valores de $R^2_{ajustado}$ disminuyeron y los valores de MSE aumentaron ligeramente, por lo cual se recomienda la aplicación de las ecuaciones 7 y 8 para la estimación rápida del azufre en el ACPM hidrotratado y el flujo de ACPM hidrotratado, respectivamente.

Tabla 6. Valores de R^2 y $R^2_{ajustado}$ para las RLM en función de HC1.

Caso de estudio	R^2	$R^2_{ajustado}$
II	0.9352	0.9352
III	0.8934	0.8934

Tabla 7. Valores de R2 y MSE en la validación externa de las RLM en función de HC1.

Caso de estudio	R²	MSE
II	0.9410	0.0520
III	0.8935	0.0020

$$HC2 = 0,0217 + 0.9499 * HC1 \quad (7)$$

$$RR = -0.0227 + 0.9558 * HC1 \quad (8)$$

3.3. REDES NEURONALES ARTIFICIALES

El flujo de ACPM hidrotratado con su respectivo contenido de azufre y a las condiciones de operación fueron predichas utilizando RNA. La siguiente nomenclatura fue establecida para las diferentes funciones de activación: FA-I para las redes con la función *logsig-purelin* y FA-II para las redes con la función *logsig-logsig*.

3.3.1. RNA para el ACPM hidrotratado y su contenido de azufre

Las topologías asumidas corresponden a la forma 9:N:2, donde N, el número de neuronas en la capa oculta varía desde 2 a 18; 9 la cantidad de variables de entrada y 2 las variables de salida (Anexo F, Figura 27) Los estadísticos de desempeño de las redes FA-I y FA-II con las diferentes topologías para la predicción del flujo de ACPM hidrotratado y su contenido de azufre pueden ser analizados en las Tablas 8 y 9. Según estas tablas, la red neuronal FA-I de topología 9:18:2 presentó los mejores estadísticos tanto para el entrenamiento como para la validación. También las redes FA-II presenta su mejor desempeño en la topología 9:16:2 (Tabla 9).

Tabla 8. Estadísticos para las redes FA-I en la predicción del HC2 y S2.

N	R^2	MSE de entrenamiento	MSE de validación
2	0.317	1.24E-02	1.21E-02
4	0.774	7.53E-03	7.41E-03
6	0.868	4.17E-03	4.10E-03
8	0.940	1.63E-03	1.57E-03
10	0.957	1.02E-03	9.71E-04
12	0.969	6.50E-04	5.78E-04
14	0.968	6.71E-04	6.20E-04
16	0.967	7.32E-04	6.97E-04
18	0.975	5.17E-04	4.75E-04

Tabla 9. Estadísticos para las redes FA-II en la predicción del HC2 y S2.

N	R^2	MSE de entrenamiento	MSE de validación
2	0.185	1.84E-02	1.86E-02
4	0.546	1.61E-02	1.64E-02
6	0.708	1.26E-02	1.27E-02
8	0.735	1.23E-02	1.24E-02
10	0.724	1.25E-02	1.26E-02
12	0.733	1.24E-02	1.25E-02
14	0.729	1.24E-02	1.25E-02
16	0.738	1.23E-02	1.24E-02
18	0.762	1.19E-02	1.20E-02

Por otra parte, aunque la red FA-I con topología 9:18:2 presentó los mejores estadísticos, MSE y R^2 , las Figuras 6 y 7 que reportan el desempeño durante la validación, indican que para valores superiores de $N=12$ la variación en los estadísticos es mínima. Lo anterior indica que el número de neuronas óptimas para la capa oculta debe corresponder a 12 para obtener valores de predicción de HC2 y S2 que puedan ser de utilidad a nivel cuantitativo.

Figura 6. Variación del MSE con N durante la validación en la predicción del HC2 y S2 con redes FA-I.

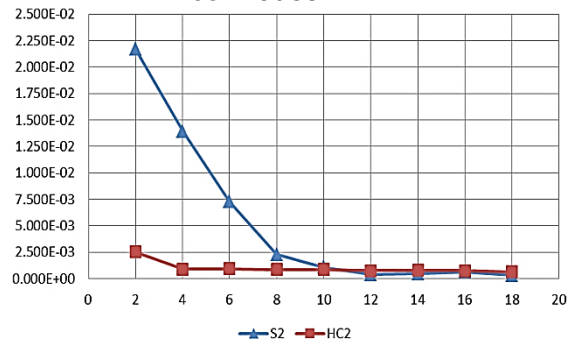
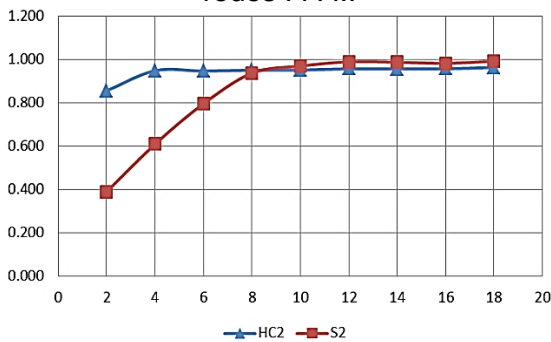
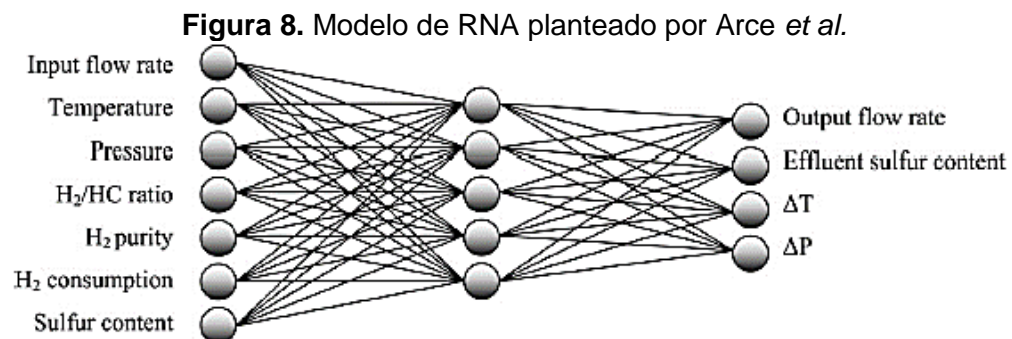


Figura 7. Variación del R² con N durante la validación en la predicción del HC2 y S2 con redes FA-I..



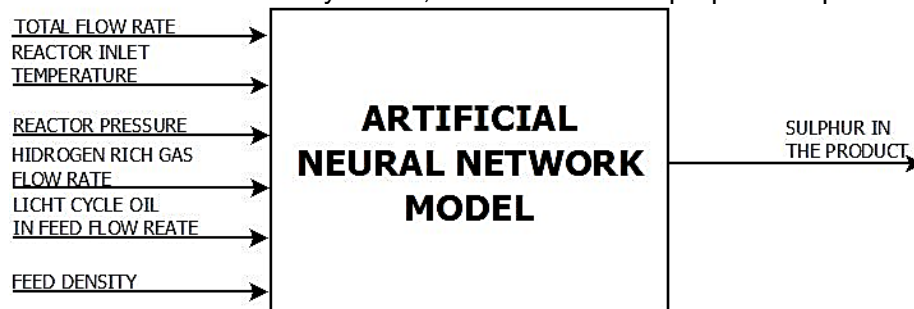
Al comparar los resultados obtenidos con los trabajos desarrollados por Arce & Paz [1] y Lukec & Sertic [26] se observa que las tendencias de predicción son similares para el contenido de azufre en el producto hidrotratado (R² mayores a 0,9); cabe resaltar que los modelos propuestos por Arce *et al.* se desarrollaron para el hidrotratamiento de nafta de la refinería de PEMEX en México, con un conjunto de datos de 197 tomados cada día y los de Lukec *et al.* para el hidrotratamiento de gasóleo ligero (GLO) y gasóleo de vacío (GVO) de la refinería de Rijeka en Croacia, con un conjunto de datos de 220 (GLO) y 125 (GVO) tomados cada hora. En el presente trabajo las RNA difieren de las planteados por Arce *et al.* y Lukec *et al.* en cuanto al número de neuronas en la capa de entrada, oculta y de salida; estos autores obtuvieron modelos de red con topologías 7:5:4 y 6:6:1, en las Figuras 8 y 9 se observan las respectivas variables utilizadas. Las RNA desarrolladas en el presente trabajo tienen una capacidad predictora con un

R^2 igual a 0,989 para el contenido de azufre en el producto hidrotratado, en comparación con las RNA propuestas por Arce *et al.* y Lukec *et al.* que obtuvieron respectivamente valores de R^2 iguales a 0,918; 0,9514 (GLO) y 0,9368 (GVO). La diferencia en las topologías y capacidad de predicción, puede deberse a la diferencia de variables de entrada y salida y al conjunto de datos utilizados.



Fuente: Arce *et al.* (2009).

Figura 9. Variables de entrada y salida, modelos de RNA propuestos por Lukec *et al.*



Fuente: Lukec *et al.* (2007).

3.3.2. RNA para las condiciones de operación

La topología considerada para la predicción de las condiciones de operación fue de la forma 7:N:3, donde 7 corresponde a la cantidad de variables de entrada, N el número de neuronas en la capa oculta y 3 las variables de salida H, dT y P (Anexo F, Figura 28). Según la relación de la capa oculta – capa de entrada [27][28], N debe ser menor o igual a 14. El desempeño de las diferentes RNA puede ser

analizado en las Tablas 10 y 11; según estas tablas el desempeño de las redes es limitado para $N \leq 14$ con el mejor desempeño obtenido en la red FA-I y 14 neuronas en la capa interna. El aumento en N muestra que la mejora en los estadísticos de desempeño de las RNA es exiguo. Las Figuras 10–13 ilustran las tendencias obtenidas para los estadísticos de la etapa de predicción con la variación en el número de neuronas internas para FA-I y FA-II. De acuerdo a estas figuras, las redes FA-I presentan su mejor desempeño en la topología 7:22:3 para la cual la etapa de predicción de la presión es deficiente reportando un valor máximo de 0,027 para el R^2 de predicción (Figura 10); en lo referente al flujo de hidrogeno y el cambio de temperatura, R^2 reporta valores de predicción de 0,926 y 0,728, respectivamente. De otro lado, las redes FA-II muestran un desempeño inapropiado para la etapa de predicción de las variables respectivas (Figura 11). Según las Figuras 12 y 13, las redes FA-I y FA-II reportan valores de *MSE* comparables. Resumiendo lo anterior, la predicción del flujo de hidrogeno, el cambio de temperatura en la zona de reacción y la presión utilizando una sola red, ya sea FA-I o FA-II, reporta aplicaciones que pueden ser de utilidad desde una valoración cualitativa, pero sólo para la predicción del flujo de hidrógeno.

Tabla 10. Desempeño de las RNA FA-I en la predicción de H, dT y P.

N	R²	MSE de entrenamiento	MSE de validación
2	0.753	6.79E-03	6.89E-03
4	0.848	4.40E-03	4.40E-03
6	0.858	4.13E-03	4.12E-03
8	0.872	3.17E-03	3.15E-03
10	0.900	2.97E-03	2.98E-03
12	0.911	2.66E-03	2.64E-03
14	0.919	2.42E-03	2.43E-03
16	0.923	2.32E-03	2.35E-03
18	0.928	2.18E-03	2.15E-03
20	0.930	2.10E-03	2.07E-03
22	0.934	2.00E-03	2.03E-03

Tabla 11. Desempeño de las RNA FA-II en la predicción de H, dT y P.

N	R ²	MSE de entrenamiento	MSE de validación
2	0.741	7.15E-03	7.28E-03
4	0.780	6.33E-03	6.41E-03
6	0.790	6.15E-03	6.24E-03
8	0.827	5.50E-03	5.58E-03
10	0.805	5.82E-03	5.89E-03
12	0.812	5.78E-03	5.83E-03
14	0.823	5.44E-03	5.52E-03
16	0.829	5.38E-03	5.45E-03
18	0.833	5.38E-03	5.42E-03
20	0.832	5.40E-03	5.47E-03
22	0.837	5.29E-03	5.37E-03

Figura 10. Variación de R² vs N en las redes FA-I para la predicción de H, dT y P-

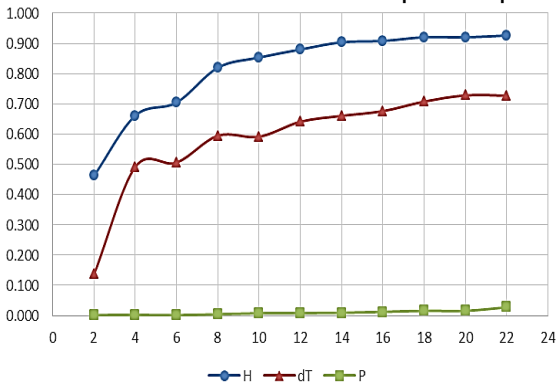


Figura 11. Variación de R² vs N en las redes FA-II para la predicción de H, dT y P.

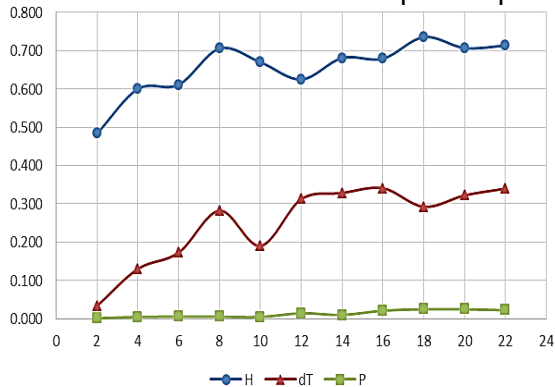


Figura 12. Variación de MSE vs N en las redes FA-I para la predicción de H, dT y P.

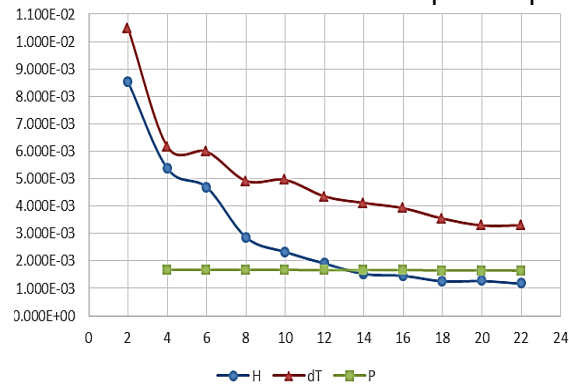
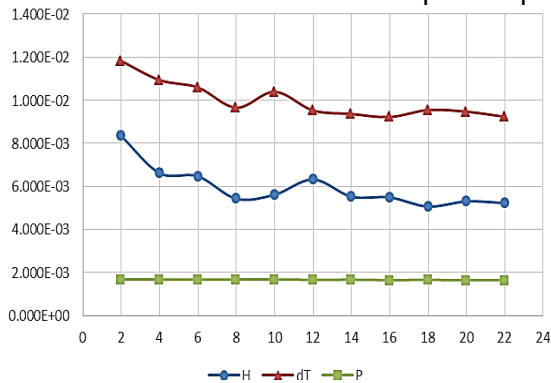


Figura 13. Variación de MSE vs N en las redes FA-II para la predicción de H, dT y P.



Por las razones expuestas anteriormente, se decidió probar otro modelo de red neuronal, donde la presión fuera una variable de entrada y no una variable de salida. De esta forma, se buscó entrenar una RNA para estimar el flujo de hidrogeno y el cambio de temperatura en la zona de reacción. Este modelo de red se desarrolló con la topología 8:N:2, donde 8 era el número de variables de entrada, N el número de neuronas en la capa oculta y 2 el número de salidas (Anexo F, Figura 29).

Según los resultados (Anexo G), no se encontró un coeficiente de determinación aceptable para las RNA dentro del límite establecido por la regla de la capa de entrada – capa oculta [27][28] con $N \leq 16$. Una mayor cantidad de neuronas en la capa oculta fue utilizada para la topología 8:22:2 con función de activación FA-I; el

desempeño de las RNA entrenadas reporta mejores coeficientes de determinación y valores de *MSE* para el flujo de hidrogeno y el cambio de temperatura que el modelo planteado inicialmente (R^2 de 0,930 y 0,759; y valores *MSE* de 1,11E-03 y 2,92E-03, respectivamente).

3.3.3. RNA para la predicción de variables individuales.

Las variables de salida en los casos de las secciones anteriores fueron predichas utilizando una topología 9:N:1, en donde 9 corresponde a las variables de entrada, N las neuronas de la capa oculta y 1 a la neurona de salida para la propiedad en turno (Anexo H); según la regla de la capa de entrada-capas ocultas [27][28], N debe tomar valores menores o iguales a 18. Los resultados mostrados en las Tablas 12 y 13, reportan que la capacidad predictiva de las redes tipo FA-I supera notablemente a las de tipo FA-II. Asimismo, la predicción del cambio de temperatura en la zona de reacción mejora respecto a lo obtenido en la sección anterior. También mejoraron las predicciones para cada una de las variables por separado, en comparación con las RNA planteadas con más de una variable de salida. Es importante mencionar que la predicción de dT requiere de una mayor cantidad de neuronas en la capa oculta que para las otras variables, con el fin de obtener un coeficiente de correlación y un *MSE* aceptables.

Por otra parte, no se plantearon RNA para predecir la presión ya que como se observó anteriormente, no fue posible obtener una predicción aceptable; al analizar esta variables (Anexo B, Figura 24, Tabla 2), se encontraron que la presión presentaba cambios bruscos en la tendencia habitual de los datos medidos, lo cual dificulta el entrenamiento de una RNA. En la Tabla 14 se exponen los mejores modelos de predicción obtenidos en el presente trabajo, donde cabe resaltar, que los modelos con configuración *logsig-purelin* en la capa oculta y de salida, fueron los que arrojaron los mejores resultados tanto en el entrenamiento como en la validación.

Tabla 12. Validación de los modelos de red neuronal con funciones de activación FA I.

Variable a predecir	Topología	R ²	MSE de entrenamiento	MSE de validación
S2	9:14:1	0.997	1.75E-04	1.21E-04
HC2	9:16:1	0.974	4.62E-04	4.57E-04
RR	9:14:1	0.993	1.23E-04	1.27E-04
H	9:16:1	0.942	8.85E-04	9.22E-04
dT	9:24:1	0.814	2.55E-03	2.25E-03

Tabla 13. Validación de los modelos de red neuronal con funciones de activación FA II.

Variable a predecir	Topología	R ²	MSE de entrenamiento	MSE de validación
S2	9:10:1	0.718	1.91E-02	1.90E-02
HC2	9:14:1	0.804	4.52E-03	4.84E-03
RR	9:14:1	0.838	4.87E-03	5.15E-03
H	9:12:1	0.758	4.97E-03	4.79E-03
dT	9:20:1	0.427	8.11E-03	8.64E-03

Tabla 14. Resumen de las mejores RNA obtenidas.

Variable(s) a predecir	Topología	R ²	MSE de entrenamiento	MSE de validación
S2 – HC2	9:12:2	0.969	6.50E-04	5.78E-04
H – dT - P	7:22:3	0.934	2.00E-03	2.03E-03
H - dT	8:22:2	0.928	1.99E-03	2.02E-03
S2	9:14:1	0.997	1.75E-04	1.21E-04
HC2	9:16:1	0.974	4.62E-04	4.57E-04
RR	9:14:1	0.993	1.23E-04	1.27E-04
H	9:16:1	0.942	8.85E-04	9.22E-04
dT	9:22:1	0.814	2.55E-03	2.25E-03

Es importante mencionar que no se encontraron trabajos reportados en la literatura que referencien el entrenamiento de RNA para la predicción de las variables de operación definiendo las condiciones de salida en el hidrotamiento del crudo y sus fracciones.

4. CONCLUSIONES Y RECOMENDACIONES

Según los resultados del presente trabajo, los modelos de regresión lineal planteados muestran un desempeño insuficiente para su aplicación en la predicción del contenido de azufre en el producto hidrotratado, el flujo de hidrogeno, y el cambio de temperatura en la zona de reacción. Por otra parte, el modelo para la relación ACPM/hidrógeno y el modelo para el flujo de ACPM hidrotratado reportan coeficientes de determinación ajustados de 0,9426 y 0,9185, con lo cual pueden ser aplicados en predicciones cualitativas.

Para la predicción con las RNA, la topología 09:12:02 con funciones de activación *logsig-purelin* reportó los menores errores cuadrados promedio (respectivamente, de 7,57E-04 y 3,98E-04) para la predicción del flujo del ACPM hidrotratado y su contenido de azufre; asimismo, los coeficientes de determinación alcanzados fueron de 0,957 y 0,989. El desempeño reportado por esta RNA puede soportar su aplicación en predicciones cuantitativas para las respectivas variables del proceso de HDT. Respecto a la predicción del flujo de hidrogeno y el cambio de temperatura en la zona de reacción, se encontró que la mejor topología fue 08:22:02 utilizando funciones de activación *logsig-purelin*, con coeficientes de determinación respectivamente de 0,930 y 0,759; y errores cuadrados promedio de 1,11E-03 y 2,92E-03.

Para mejorar los desempeños de las RNA se recomienda adicionar otras variables de proceso como el factor de desactivación del catalizador. Asimismo, realizar validaciones adicionales con datos históricos para soportar la aplicación cuantitativa de los modelos propuestos.

REFERENCIAS

- [1] ARCE-MEDINA, E., & PAZ-PAREDES, J. Artificial neural network modeling techniques applied to the hydrodesulfurization process. *Mathematical and Computer Modelling An International Journal*, 2009, (pp. 207-214).
- [2] MIGUEL, V. PASSINI, L. Hidrotratamiento de nafta y diésel en refinería de Bahía Blanca. *Petrotecnia* 2010. Abril. (Pp. 72-79).
- [3] ÁLVAREZ, Magda C.; HOYOS, Luis J.; ZAMBRANO Liliana G.; Diesel hydrotreating process modelling. Instituto Colombiano del Petróleo (ICP), Ecopetrol S.A., aceptado 27/11/2012.
- [4] Alaei, J; Nariman E Simulation of hydrodesulfurization unit for natural gas condensate with high sulfur content. *Applied Petrochemical Research* (107) 2015.
- [5] MENDEROS NIETO, Fabián S., Desarrollo de un modelo dinámico de una planta piloto de hidrotratamiento. Tesis de maestría en ciencias con especialidad en ingeniería química. Instituto politécnico nacional, México D.F., 2004.
- [6] RODRÍGUEZ H., Miguel A. Modelación de un reactor de HDT de crudo pesado y su aplicación en integración energética. Tesis de doctorado en Ingeniería Energética. Universidad Nacional Autónoma de México, Facultad de Ingeniería, 2007.
- [7] RODRÍGUEZ M, ANCHEYTA J. Modeling of hydrodesulfurization (HDS), Hydrodenitrogenation (HDN), and the Hydrogenation of Aromatics (HDA) in a Vacuum Gas Oil Hydrotreater. *Ener. & Fuels*. 2004;18(3):789-94.
- [8] MEDEROS F, ANCHEYTA J. Mathematical modeling and simulation of hydrotreating reactors: Cocurrent versus countercurrent operations. *Appl. Catal. A-Gen*. 2007; 332:8-21.
- [9] VEGA, Andrés F., DE RIO AMADOR, Eduardo, BARBOSA, Aida L. Hidrodesulfuración de crudos de petróleo: base para el mejoramiento de combustibles. una revisión. *Avances en Ciencias e Ingeniería*, 2014, Julio-septiembre, 37-60.

- [10] Ministerio de Minas y Energía. Resolución 90963 de 2014. [Citado el 05 de abril de 2016]. Disponible en <
<https://www.minminas.gov.co/documents/10180//23517//26101-90963.pdf> >
- [11] GONZÁLEZ-CORTÉS, S.L., RUGMINI, S., XIAO, T., GREEN, M.L.H., RODULFO-BAECHLER, S.M. & IMBERT, F.E. Deep hydrotreating of different feedstocks over a highly active Al₂O₃-supported NiMoW sulfide catalyst. *Applied Catalysis A: General*, 475 (0), 2014, 270-281.
- [12] KNUDSEN, K.G., COOPER, B.H. & TOPSØE, H. Catalyst and process technologies for ultra-low sulfur diesel. *Applied Catalysis A: General*, 189 (2), 1999, 205-215.
- [13] JONES, D. & PUJADÓ, P. *Handbook of petroleum processing*: Springer. 2006.
- [14] ANDARI, M.K., ABU-SEEDO, F., STANISLAUS, A. & QABAZARD, H.M. Kinetics of individual sulfur compounds in deep hydrodesulfurization of Kuwait diesel oil. *Fuel*, 75 (14), 1996, 1664-1670
- [15] HAIR JJ, BLACK W, BABIN B, ANDERSON R. *Multivariate Data Analysis*. 7th editio. USA: Prentice Hall; 2009.
- [16] MENDENHALL W. *Introducción a la probabilidad y la estadística*. 1987. p. 551-652.
- [17] DRAPER, N.R.; SMITH, H. *Applied Regression Analysis*. Editorial Pueblo y Educación, 1980.
- [18] WALPOLE, Ronald E., MYERS, Raymond H. *Probabilidad y estadística para ingenieros*. México: Prentice-hall Hispanoamérica s.a., 1999.
- [19] MARTÍN DEL BRÍO, B. Y SANZ, M. *A Redes Neuronales y Sistemas Difusos (2da Ed.)*. México D.F, México. Alfaomega, 2002.
- [20] PONCE, C. Pedro. *Inteligencia artificial con aplicaciones a la ingeniería (1ra Ed.)*. México. Alfaomega, 2010.
- [21] ISASI P., GALVAN I. *Redes neuronales artificiales enfoque práctico*. Pearson Educación, S. A. Madrid. 2004.
- [22] ZEPEDA, José A., AGUILAR, Aldo E., HERNÁNDEZ, Álvaro, SALGADO, María C. Metodología para determinar la topología, patrones de parámetros eléctricos y entrenamiento de una Red Neuronal Artificial para un control inteligente. *Revista Tecnología Digital* Vol. 5 No. 1, 2015. Disponible en: www.revistatecnologiadigital.com

- [23] SOTOLONGO, G.; GUZMÁN, MV. Aplicaciones de las redes neuronales. El caso de la bibliometría. Ciencias de la Información. (preprint) 2001; 32(1):27-34.
- [24] MONGOMERY, Douglas C. Y RUNGER, George C. Probabilidad y estadística aplicadas a la ingeniería, México, Mc Graw-Hill, 1996.
- [25] Berrocal, J. L. A., Paniagua, C. G. F., & Rodríguez, Á. F. Z. Mejoras en la recuperación de información en la Web mediante el tratamiento de la información de los enlaces. Scire: Representación y organización del conocimiento, 12(1), 2006, 197-209.
- [26] IVANA LUKEC, KATICA SERTIĆ-BIONDA, DARKO LUKEC, Prediction of sulphur content in the industrial hydrotreatment process. Publicado el 30 de noviembre de 2007. Revista Elsevier 2007 B.V. All rights reserved.
- [27] SWINGLER, K., Applying Neural Networks, A Practical Guide. Press Limited Oval Road London NW1, 1996.
- [28] BERRY, M.J.A., LINOFF, G., Data Mining Techniques. John Wiley & Sons. Nueva York, 1997.
- [29] DEVORE, Jay L., Probabilidad y Estadística para Ingeniería y Ciencias. Séptima edición, 2008.

BIBLIOGRAFÍA

ALAEI J., NARIMAN E. Simulation of hydrodesulfurization unit for natural gas condensate with high sulfur content. Applied Petrochemical Research (107) 2015. 1-10

ÁLVAREZ, Magda C.; HOYOS, Luis J.; ZAMBRANO Lilliana G.; Diesel hydrotreating process modelling. Instituto Colombiano del Petróleo (ICP), Ecopetrol S.A., Revistas.uis.edu.co 2012. (25). 7-14

ANDARI, M.K., ABU-SEEDO, F., STANISLAUS, A. & QABAZARD, H.M., Kinetics of individual sulfur compounds in deep hydrodesulfurization of Kuwait diesel oil. Fuel, 75 (14), 1996, 1664-1670.

ARCE-MEDINA, E., & PAZ-PAREDES, J., Artificial neural network modeling techniques applied to the hydrodesulfurization process. Mathematical and Computer Modelling An International Journal, 2009 , (pp. 207-214).

BASHEERA, I., HAJMEERB M. Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods. Volumen 43, Issue 1, 1 diciembre 2000, P. 3–31

BELLOS G.D., KALLINIKOS L.E., GOUNARIS C.E., PAPAYANNKOS N.G. Modelling of the performance of industrial HDS reactors using a hybrid neural network approach. Chemical Engineering and Processing: Process Intensification. Volumen 44, Issue 5, May 2005, P. 505–515

BERRY, M.J.A., LINOFF, G., Data Mining Techniques. John Wiley & Sons. Nueva York, 1997.

DRAPER, N.R.; SMITH, H., Applied Regression Analysis. Editorial Pueblo y Educación, 1980.

NUÑEZ A., Eduardo, STEYERBERGC, Ewout W., NUÑEZ, Julio. Estrategias para la elaboración de modelos estadísticos de regresión. Publicado el 29 de abril de 2011. 2011 Sociedad Española de Cardiología. Published by Elsevier España, S.L. All rights reserved.

FERNANDEZ R., Fransisco. Sistema para la predicción de posición y seguimiento de un conjunto de naufragos basado en redes neuronales. Tesis de maestría. Universidad Complutense de Madrid, Madrid, España. 2009-2010.

ISASI P., GALVAN I. Redes neuronales artificiales enfoque práctico. Pearson Educación, S. A. Madrid. 2004.

GONZÁLEZ-CORTÉS, S.L., RUGMINI, S., XIAO, T., GREEN, M.L.H., RODULFO-BAECHLER, S.M. & IMBERT, F.E., Deep hydrotreating of different feedstocks over a highly active Al₂O₃-supported NiMoW sulfide catalyst. Applied Catalysis A: General, 475 (0), 2014, 270-281.

HAIR JJ, BLACK W, BABIN B, ANDERSON R. Multivariate Data Analysis. 7th editio. USA: Prentice Hall; 2009.

IVANA LUKEC, KATICA SERTIĆ-BIONDA, DARKO LUKEC, Prediction of sulphur content in the industrial hydrotreatment process. Publicado el 30 de noviembre de 2007. Revista Elsevier 2007 B.V. All rights reserved.

JONES, D. & PUJADÓ, P., Handbook of petroleum processing: Springer, 2006.

KNUDSEN, K.G., COOPER, B.H. & TOPSØE, H., Catalyst and process technologies for ultra-low sulfurdiesel. Applied Catalysis A: General, 189 (2), 1999, 205-215.

LENNOXA B., MONTAGUEB G., FRITHC A., GENTC C., BEVANC V. Industrial application of neural networks — an investigation. Journal of Process Control. Volumen 11, Issue 5, octubre 2001, P. 497–507

Neural Network Toolbox User's Guide. Matlab R2015a (*The MathWorks, inc.*).

MARTÍN DEL BRÍO, B. Y SANZ, M. A., Redes Neuronales y Sistemas Difusos (2da Ed.). México D.F, México. Alfaomega, 2002.

MENDENHALL W. Introducción a la probabilidad y la estadística. 1987. p. 551-652.

MEDEROS F, ANCHEYTA J. Mathematical modeling and simulation of hydrotreating reactors: Cocurrent versus countercurrent operations. Appl. Catal. A-Gen. 2007; 332:8-21.

MENDEROS NIETO, Fabian S., Desarrollo de un modelo dinámico de una planta piloto de hidrotreamiento. Tesis de maestría en ciencias con especialidad en ingeniería química. Instituto politécnico nacional, México D.F., 2004.

MICHALOPOULOS J., PAPADOKONSTADAKIS S., ARAMPATZIS G., LYGEROS A. Modelling of an Industrial Fluid Catalytic Cracking Unit Using Neural Networks. Chemical Engineering Research and Design. Volumen 79, Issue 2, marzo 2001, P. 137-142.

PASSINI, L., Miguel V. Hidrotreamiento de nafta y diésel en refinería de Bahía Blanca. Petrotecnia 2010. Abril. (Pp. 72-79).

MONGOMERY, Douglas C., RUNGER, George C., Probabilidad y estadística aplicadas a la ingeniería, México, Mc Graw-Hill, 1996.

NÚÑEZ E., STEYERBERG E., NÚÑEZ J. Estrategias para la elaboración de modelos estadísticos de regresión. Sociedad Española de Cardiología. Published by Elsevier España, S.L. All rights reserved. 2011; 64(6):501–507

PONCE, C. Pedro, Inteligencia artificial con aplicaciones a la ingeniería (1ra Ed.). México. Alfaomega, 2010.

RODRÍGUEZ H., MIGUEL A. Modelación de un reactor de HDT de crudo pesado y su aplicación en integración energética. Tesis de doctorado en Ingeniería

Energética. Universidad Nacional Autónoma de México, Facultad de Ingeniería, 2007.

SAFAMIRZAEI M., MODARRESS H. Hydrogen solubility in heavy n-alkanes; modeling and prediction by artificial neural network. Fluid Phase Equilibria. Volumen 310, Issues 1–2, 25 noviembre 2011, P. 150–155

SHOKRI S., SADEGHI M. S., MARVAST M., NARASIMHAN S. Improvement of the prediction performance of a soft sensor model based on support vector regression for production of ultra-low sulfur diésel. Petroleum Science. March 2015, Volumen 12, Issue 1, pp 177–188

SOLÍS G., MARROQUIN P., Aplicación de Técnicas de Análisis Estadístico Multivariado en un Proceso de Control Dimensional. IEEE LATIN AMERICA TRANSACTIONS, VOL. 5, NO. 2, MAYO 2007

SOTOLONGO, G.; GUZMÁN, MV. Aplicaciones de las redes neuronales. El caso de la bibliometría. Ciencias de la Información. (preprint) 2001; 32(1):27-34.

STANISLAUS, A., MARAFI, A. & RANA, M.S., Recent advances in the science and technology of ultra-low sulfur diesel (ULSD) production. Catalysis Today, 153 (1–2), 2010, 1-68.

SWINGLER, K., Applying Neural Networks, A Practical Guide. Press Limited Oval Road London NW1, 1996.

VEGA, Andrés F., DE RIO AMADOR, Eduardo, BARBOSA, Aida L., Hidrodesulfuración de crudos de petróleo: base para el mejoramiento de combustibles. una revisión. Avances en Ciencias e Ingeniería, Julio-septiembre, 2014, 37-60.

W. SAUERBREIA, C. MEIER-HIRMERB, A. BENNERC, P. ROYSTOND. Multivariable regression model building by using fractional polynomials: Description of SAS, STATA and R programs. Publicado Julio 13 de 2005. © 2005 Elsevier B.V. All rights reserved.

WALPOLE, RONALD E., MYERS, RAYMOND H., Probabilidad y estadística para ingenieros. México: Pretice-hall Hispanoamérica s.a., 1999.

WHITEHURST, D.D., ISODA, T. & MOCHIDA, I., Present State of the Art and Future Challenges in the Hydrodesulfurization of Polyaromatic Sulfur Compounds. In W.O.H.B.G.D.D. Eley & K. Helmut (Eds.), Advances in Catalysis (Vol. Volumen 42, 1998, pp. 345-471): Academic Press.

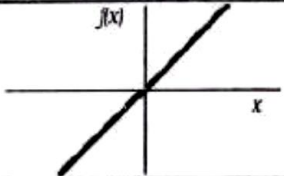
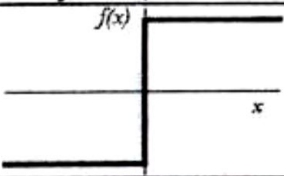
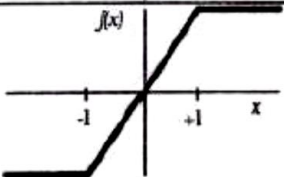
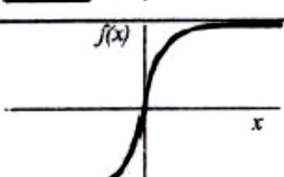
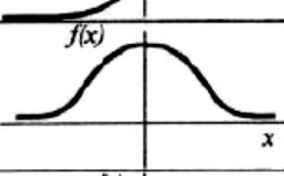
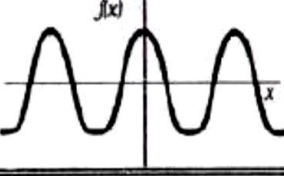
ZAHEDIA G., ELKAMELA A., LOHIC A., JAHANMIRID A., RAHIMPORD M.R. Hybrid artificial neural network—First principle model formulation for the unsteady state simulation and analysis of a packed bed reactor for CO₂ hydrogenation to metanol. Chemical Engineering Journal. Volumen 115, Issues 1–2. 15 diciembre de 2005, P. 113–120

ZEPEDA, José A.; AGUILAR, Aldo E.; HERNÁNDEZ, Álvaro; SALGADO, Maria C., Metodología para determinar la topología, patrones de parámetros eléctricos y entrenamiento de una Red Neuronal Artificial para un control inteligente. Revista Tecnología Digital Vol. 5 No. 1, 2015.

ANEXOS

Anexo A. Funciones típicas de transferencia o activación.

Figura 14. Funciones de activación habituales.

	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
Lineal a tramos	$y = \begin{cases} -1, & \text{si } x < -l \\ x, & \text{si } -l \leq x \leq +l \\ +1, & \text{si } x > +l \end{cases}$	$[-1, +1]$	
Sigmoidea	$y = \frac{1}{1 + e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = Ae^{-Bx^2}$	$[0, +1]$	
Sinusoidal	$y = A \text{sen}(\omega x + \varphi)$	$[-1, +1]$	

Fuente: (Martín del Brío, 2002).

Anexo B. Gráficas de dispersión de las variables del proceso de HDT.

Figura 15. Diagrama de dispersión para el flujo de ACPM (HC1).

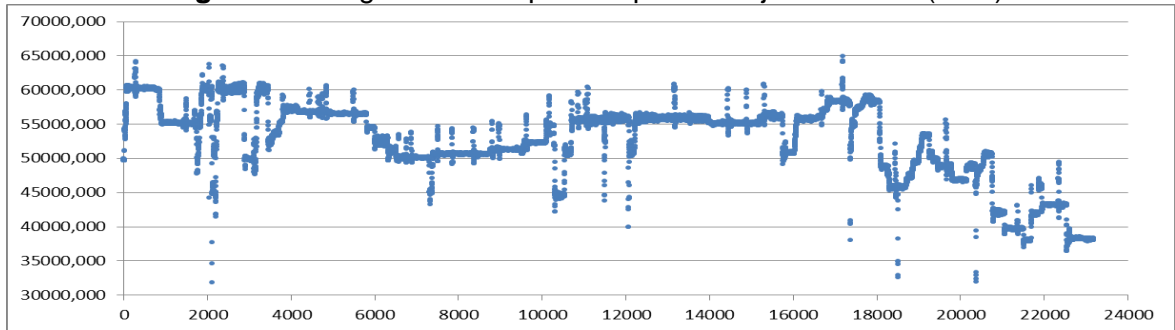


Figura 16. Diagrama de dispersión para el punto final de ebullición (PFB).

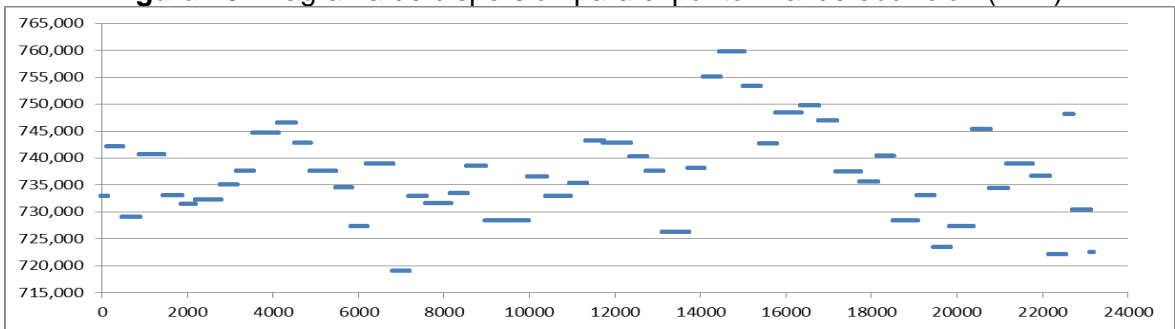


Figura 17. Diagrama de dispersión para el punto inicial de ebullición (PIB).

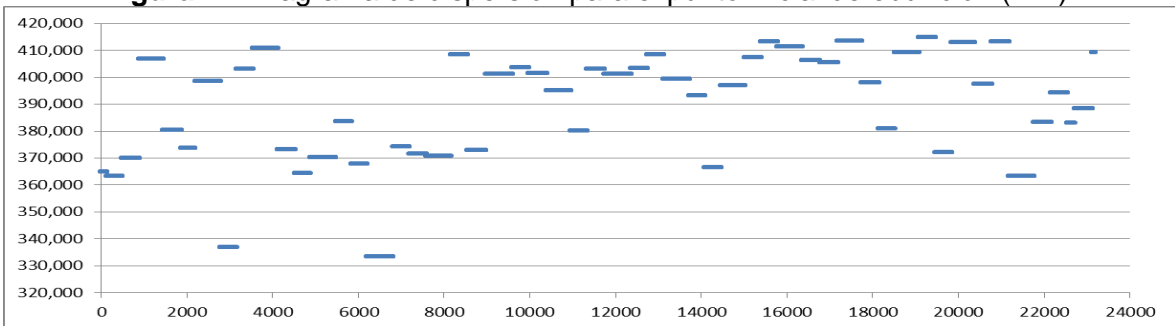


Figura 18. Diagrama de dispersión para grados API.

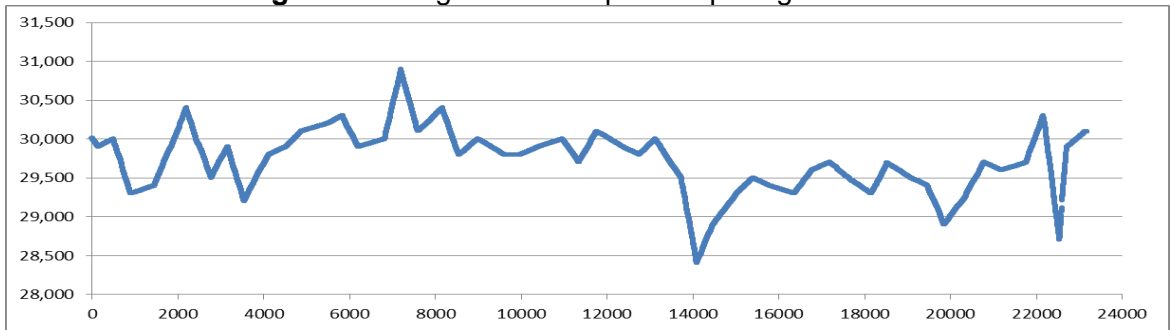


Figura 19. Diagrama de dispersión para el contenido de azufre en HC1 (S1).

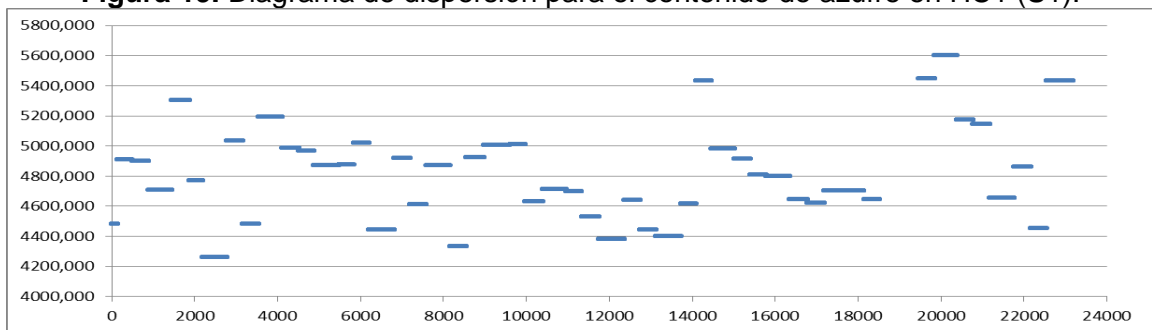


Figura 20. Diagrama de dispersión para el flujo de hidrógeno (H).

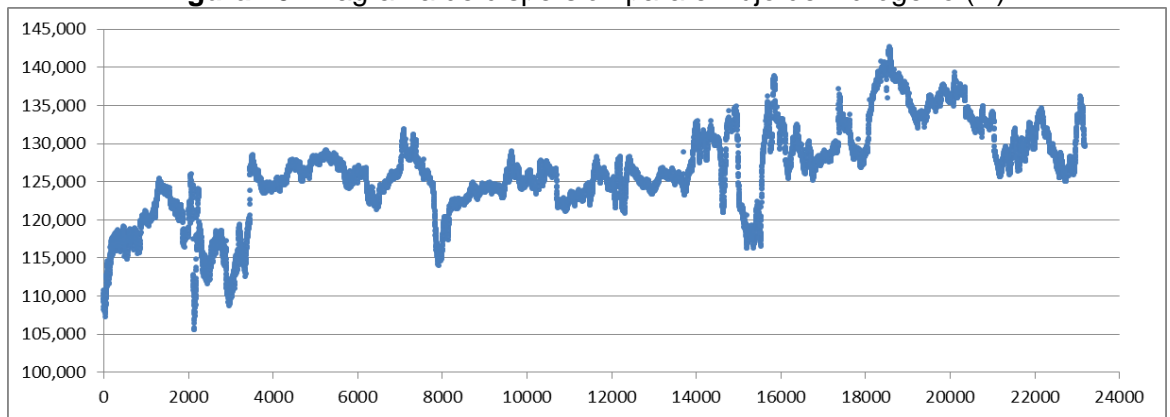


Figura 21. Diagrama de dispersión para la relación ACPM/hidrógeno (RR).

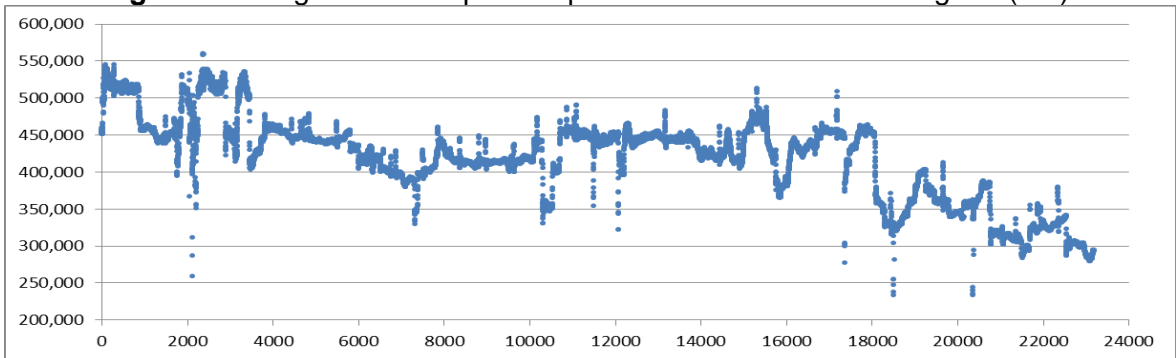


Figura 22. Diagrama de dispersión para la temperatura de entrada (T1).

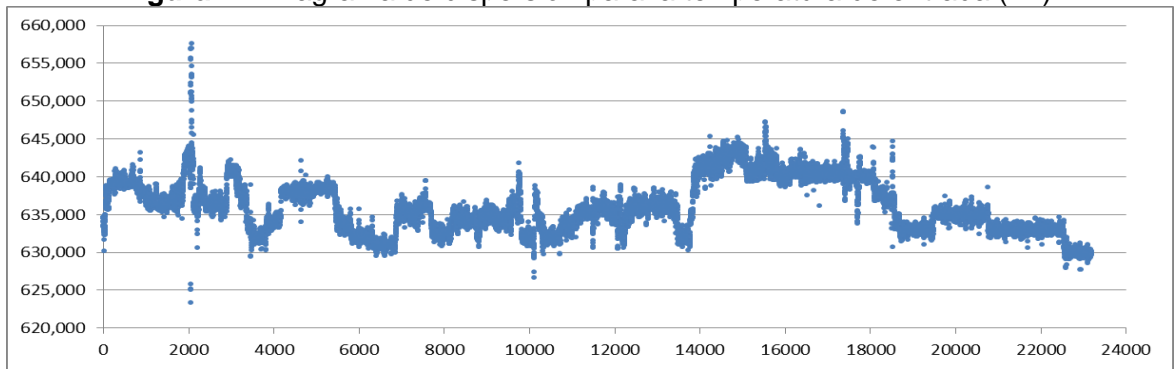


Figura 23. Diagrama de dispersión para el cambio de temperatura en la zona de reacción (dT).

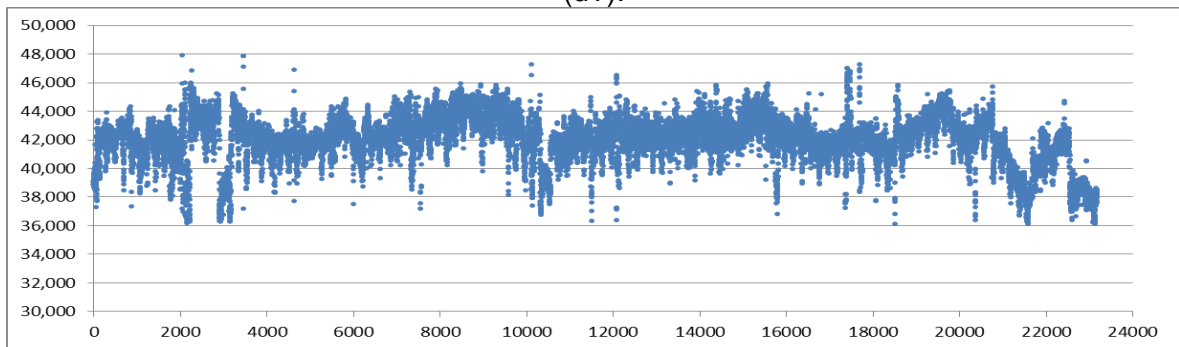


Figura 24. Diagrama de dispersión para la presión de operación en la zona de reacción (P).

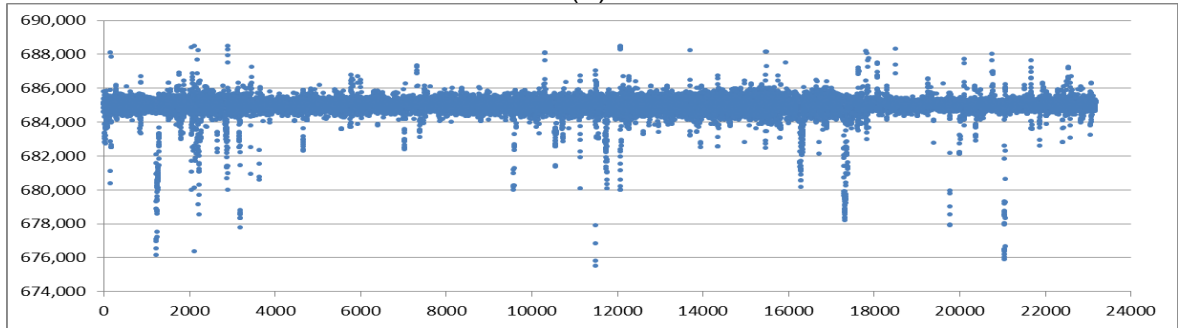


Figura 25. Diagrama de dispersión para el flujo de ACPM hidrotreatado (HC2).

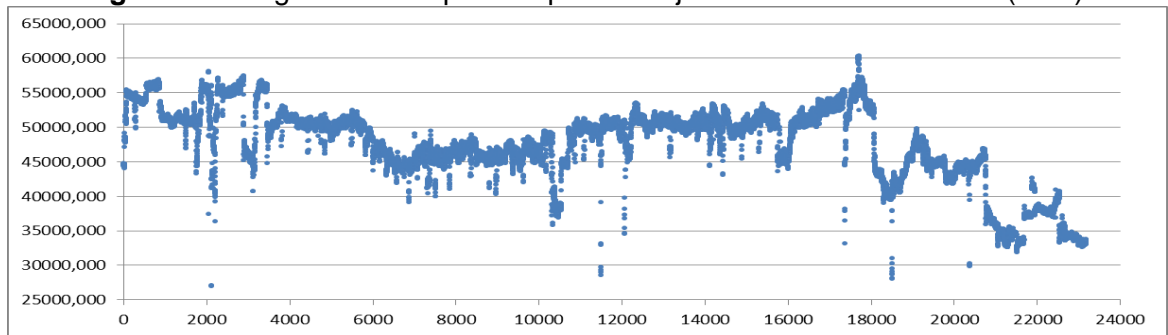
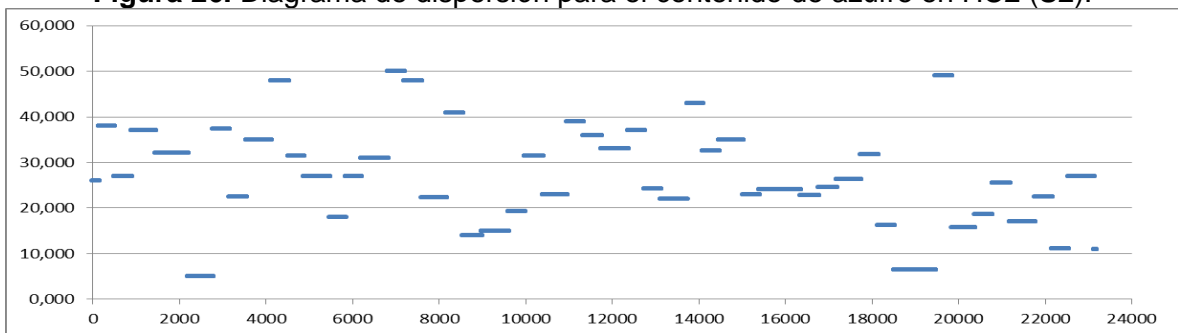


Figura 26. Diagrama de dispersión para el contenido de azufre en HC2 (S2).



Anexo C. Método de normalización Min-Max.

La normalización de los datos puede llevarse acabo de la siguiente forma:

$$X_i = \frac{(x_i - x_{min}) * (d_2 - d_1)}{(x_{max} - x_{min})} + d_1$$

Donde:

X_i es la variable normalizada.

x_i es el dato de la variable a normalizar.

$(x_i - x_{min})$ es el rango de la variable x .

$(d_2 - d_1)$ es el rango al cual será reducida la variable x .

Para obtener nuevamente el valor de los datos sin normalizar se realiza mediante la siguiente expresión:

$$x_i = \frac{(X_i - d_1) * (x_{max} - x_{min})}{(d_2 - d_1)} + x_{min}$$

Anexo D. Detalles calibración modelos de RLM.

Tabla 15. Resumen del ajuste de RLM para el contenido de azufre en el producto hidrotratado (S2). ($R^2= 0,2423$; $R^2_{ajustado}=0,2419$).

	coeficiente	Error estándar	Valor - T	Valor - P
Intercepto	0,15000	0,02800	5,36	0,000
HC1	0,37460	0,01570	23,92	0,000
PFB	0,18105	0,00957	18,93	0,000
PIB	-0,23548	0,00743	-31,68	0,000
API	0,21510	0,01340	16,06	0,000
S1	0,31320	0,00817	38,35	0,000
H	0,05900	0,01320	4,48	0,000
T1	0,05900	0,02320	2,54	0,011
dT	-0,08330	0,01390	-5,97	0,000
P	-0,14050	0,03080	-4,57	0,000

Tabla 16. Resumen del ajuste de RLM para el flujo de ACPM hidrotratado (HC2). ($R^2= 0,9426$; $R^2_{ajustado}=0,9426$).

	coeficiente	Error estándar	Valor - T	Valor - P
Intercepto	0,06596	0,00532	12,410	0,000
HC1	0,89495	0,00298	300,710	0,000
PFB	-0,03153	0,00182	-17,340	0,000
PIB	0,02720	0,00141	19,260	0,000
API	-0,06224	0,00254	-24,470	0,000
S1	-0,00385	0,00155	-2,480	0,013
H	-0,06048	0,00250	-24,200	0,000
T1	0,09446	0,00441	21,420	0,000
dT	0,01412	0,00265	5,330	0,000
P	0,01267	0,00585	2,170	0,030

Tabla 17. Resumen del ajuste de RLM para la relación ACPM/hidrógeno (RR).
($R^2= 0,9186$; $R^2_{ajustado}=0,9185$).

	coeficiente	Error estándar	Valor - T	Valor - P
Intercepto	0,03351	0,00613	5,460	0,000
HC1	1,02497	0,00322	317,970	0,000
PIB	-0,06048	0,00173	-34,980	0,000
API	0,08250	0,00294	28,050	0,000
S1	0,00802	0,00199	4,040	0,000
T1	-0,08754	0,00510	-17,150	0,000
dT	-0,08023	0,00317	-25,290	0,000
P	-0,03121	0,00717	-4,350	0,000
S2	-0,01942	0,00181	-10,730	0,000

Tabla 18. Resumen del ajuste de RLM para el flujo de hidrógeno (H).
($R^2= 0,3995$; $R^2_{ajustado}=0,3992$).

	coeficiente	Error estándar	Valor - T	Valor - P
Intercepto	0,69140	0,01580	43,750	0,000
HC1	-0,59731	0,00827	-72,190	0,000
PFB	-0,02653	0,00577	-4,600	0,000
PIB	0,14566	0,00443	32,900	0,000
API	-0,24312	0,00782	-31,070	0,000
S1	-0,02450	0,00509	-4,820	0,000
H	0,26690	0,01370	19,490	0,000
T1	0,20685	0,00817	25,310	0,000
dT	0,07850	0,01840	4,270	0,000
P	0,02100	0,00469	4,480	0,000

Tabla 19. Resumen del ajuste de RLM para el cambio de temperatura en la zona de reacción (dT). ($R^2= 0,2582$; $R^2_{ajustado}=0,2578$).

	coeficiente	Error estándar	Valor - T	Valor - P
Intercepto	0,23410	0,01510	15,510	0,000
HC1	0,53525	0,00766	69,840	0,000
PFB	-0,07027	0,00541	-12,990	0,000
PIB	0,02392	0,00426	5,610	0,000
API	-0,05905	0,00733	-8,060	0,000
H	0,18358	0,00726	25,300	0,000
T1	-0,33860	0,01270	-26,610	0,000
P	0,05090	0,01730	2,940	0,003
S2	-0,02528	0,00423	-5,980	0,000

Anexo E. Detalles calibración modelos de RLM en función de HC1.

Tabla 20. Resumen del ajuste de RLM para el flujo de ACPM hidrotratado (HC2) en función de HC1. ($R^2= 0,9352$; $R^2_{ajustado}=0,9352$).

	coeficiente	Error estándar	Valor - T	Valor - P
Intercepto	0.0217	0.0012	17.9111	0,000
HC1	0.9499	0.0020	483.7841	0,000

Tabla 21. Resumen del ajuste de RLM para la relación ACPM/hidrógeno (RR) en función de HC1. ($R^2= 0,8934$; $R^2_{ajustado}=0,8924$).

	coeficiente	Error estándar	Valor - T	Valor - P
Intercepto	-0.0227	0.0016	-14.1857	0.0000
HC1	0.9558	0.0026	368.6175	0.0000

Anexo F. Esquemas de RNA para la predicción de variables en conjunto.

Figura 27. Esquema RNA para la predicción del flujo de ACPM hidrotornado y su contenido de azufre.

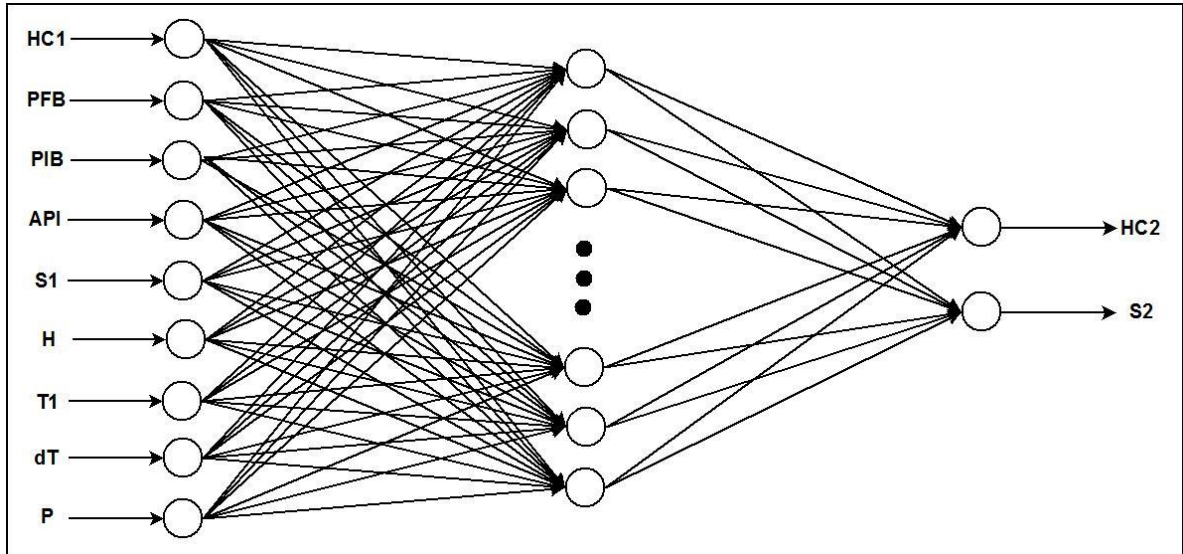


Figura 28. Esquema RNA para la predicción del flujo de hidrógeno, cambio de temperatura y presión en la zona de reacción.

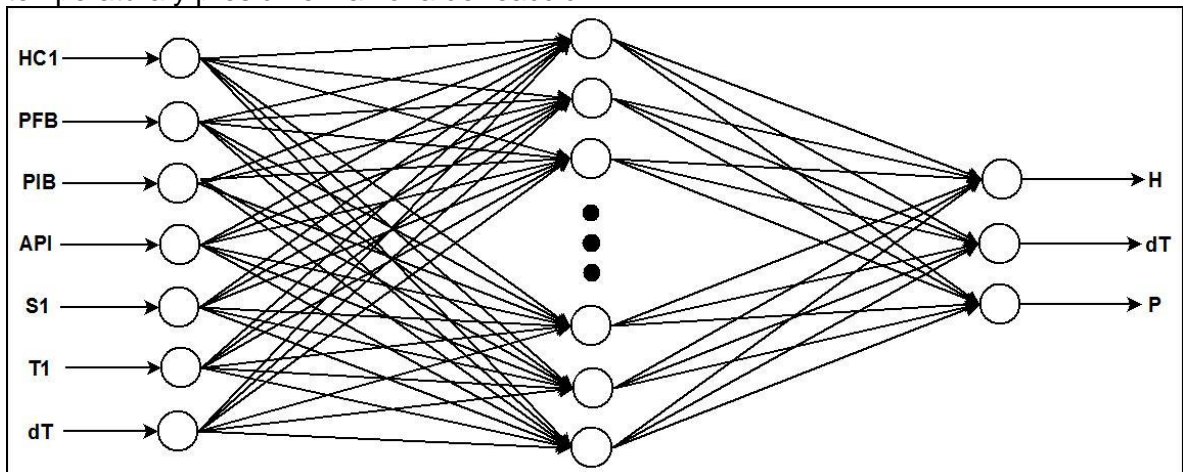
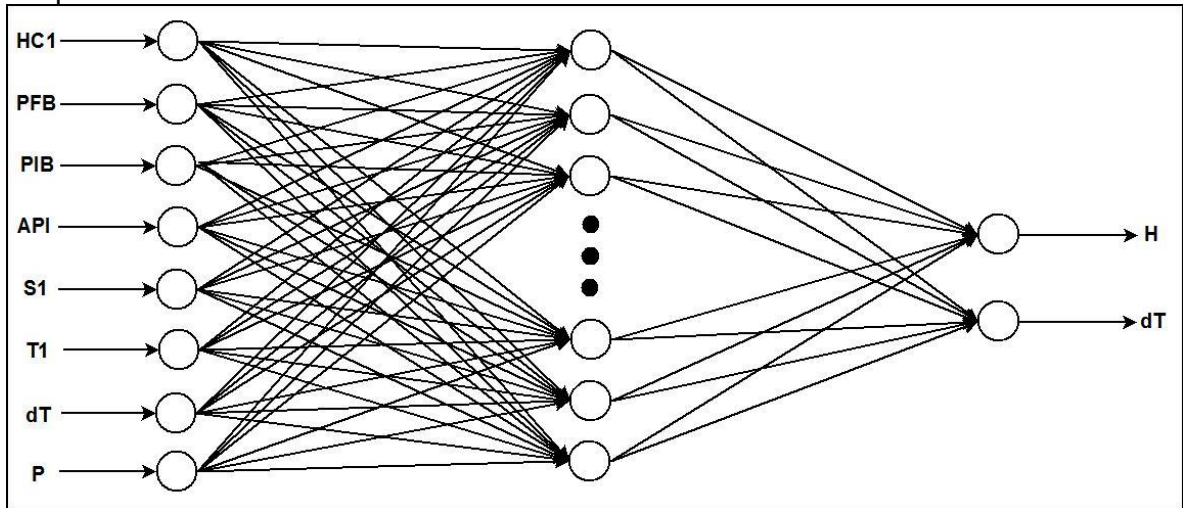


Figura 29. Esquema RNA para la predicción del flujo de hidrógeno y cambio de temperatura en la zona de reacción.



Anexo G. Resultados RNA para las condiciones de operación H y dT.

Tabla 22. Resultados modelos de redes neuronales con las funciones de activación FA I para la predicción de H y dT.

N	R^2	MSE de entrenamiento	MSE de validación
2	0,694	7,46E-03	7,51E-03
4	0,775	5,80E-03	5,85E-03
6	0,825	4,84E-03	4,85E-03
8	0,827	4,31E-03	4,28E-03
10	0,874	3,39E-03	3,40E-03
12	0,877	3,33E-03	3,34E-03
14	0,894	2,89E-03	2,96E-03
16	0,899	2,75E-03	2,72E-03
18	0,917	2,29E-03	2,35E-03
20	0,909	2,49E-03	2,53E-03
22	0,928	1,99E-03	2,02E-03

Tabla 23. Resultados modelos de redes neuronales con las funciones de activación FA II para la predicción de H y dT.

N	R^2	MSE de entrenamiento	MSE de validación
2	0,543	1,02E-02	1,04E-02
4	0,661	8,50E-03	8,63E-03
6	0,677	8,18E-03	8,29E-03
8	0,719	7,67E-03	7,80E-03
10	0,715	7,65E-03	7,83E-03
12	0,737	7,36E-03	7,48E-03
14	0,753	7,14E-03	7,28E-03
16	0,756	7,06E-03	7,19E-03
18	0,755	7,08E-03	7,22E-03
20	0,764	6,94E-03	7,07E-03
22	0,755	7,06E-03	7,21E-03

Figura 30. Gráfica del R^2 vs en número de neuronas en la capa oculta para la predicción del H y dT. Redes neuronales con función de activación FA I.

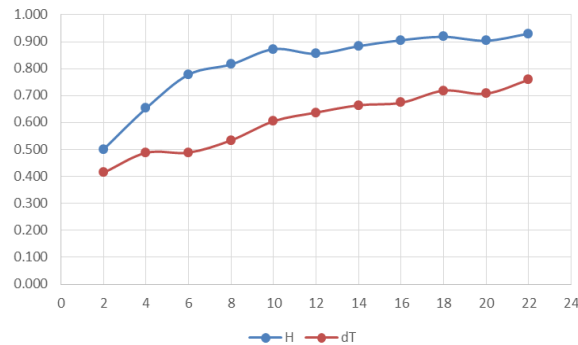


Figura 31. Gráfica del R^2 vs en número de neuronas en la capa oculta para la predicción el H y dT. Redes neuronales con función de activación FA II.

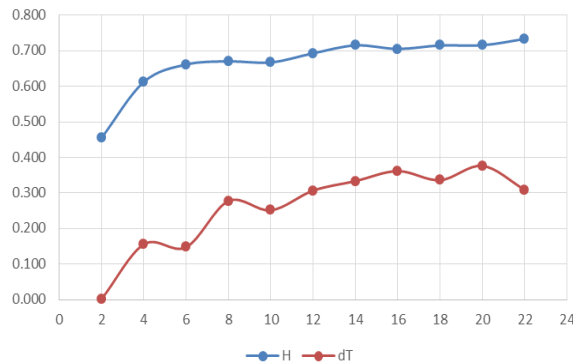


Figura 32. Gráfica del MSE vs en número de neuronas en la capa oculta para la predicción del H y dT. Redes neuronales con función de activación FA I.

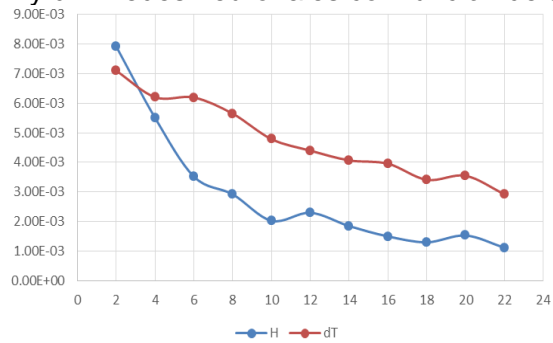
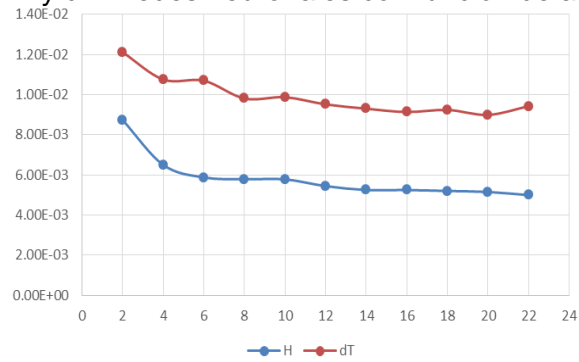


Figura 33. Gráfica del MSE vs en número de neuronas en la capa oculta para la predicción del H y dT. Redes neuronales con función de activación FA II.



Anexo H. Esquemas RNA para la predicción de variables individuales.

Figura 34. Esquema RNA para la predicción del flujo del contenido de azufre en HC2.

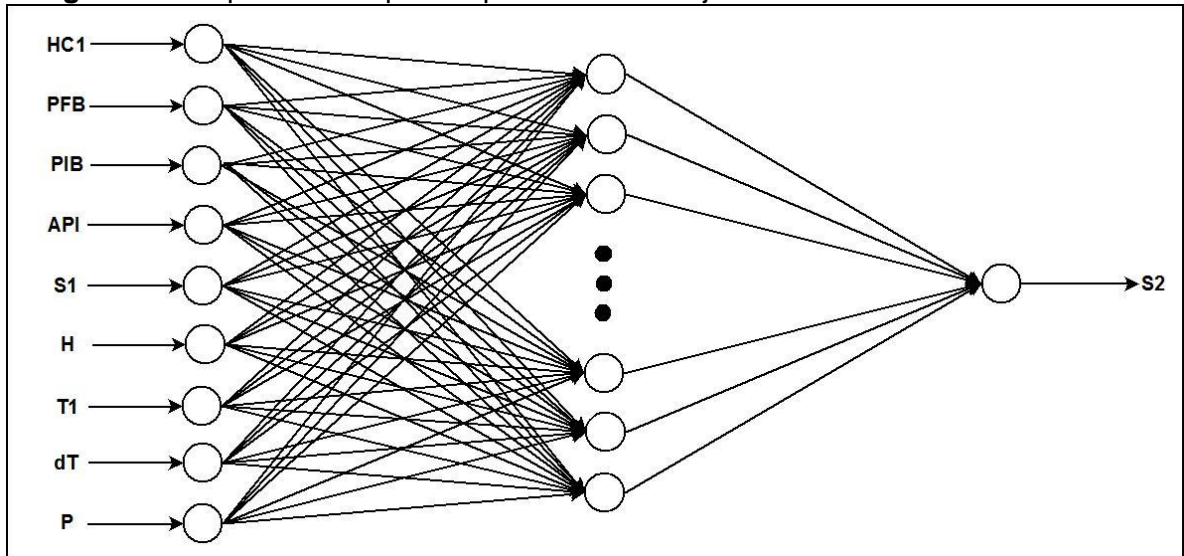


Figura 35. Esquema RNA para la predicción del flujo de ACPM hidrotratado.

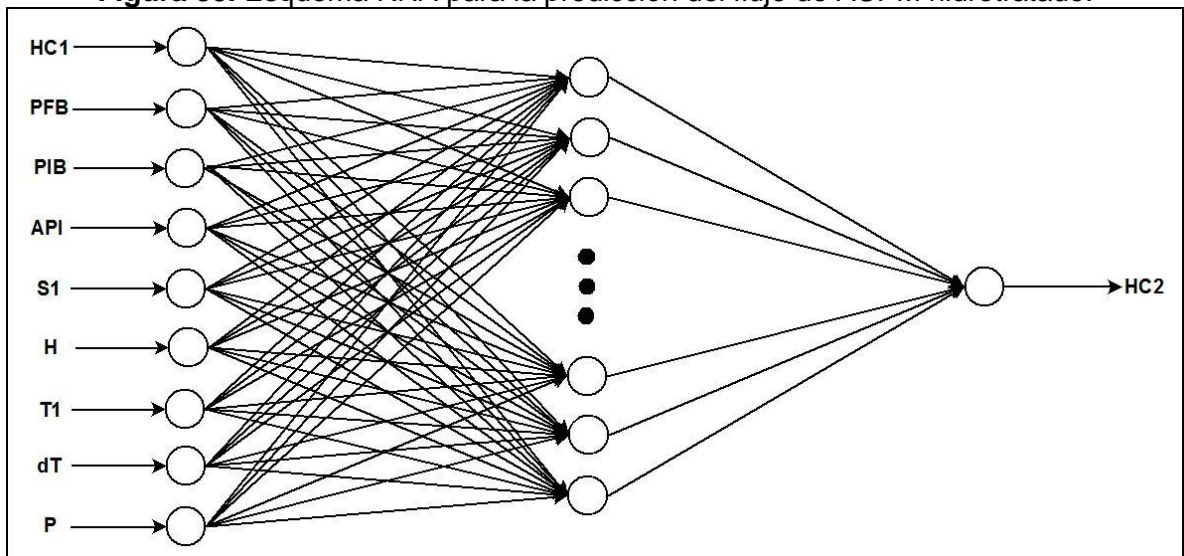


Figura 36. Esquema RNA para la predicción de la relación ACPM/hidrógeno.

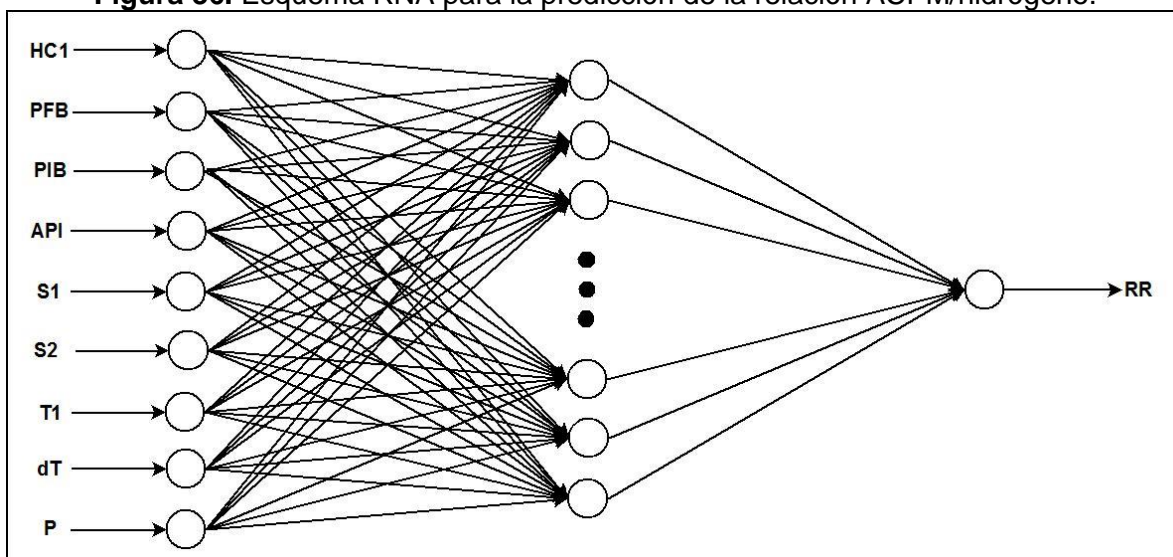


Figura 37. Esquema RNA para la predicción del flujo hidrógeno.

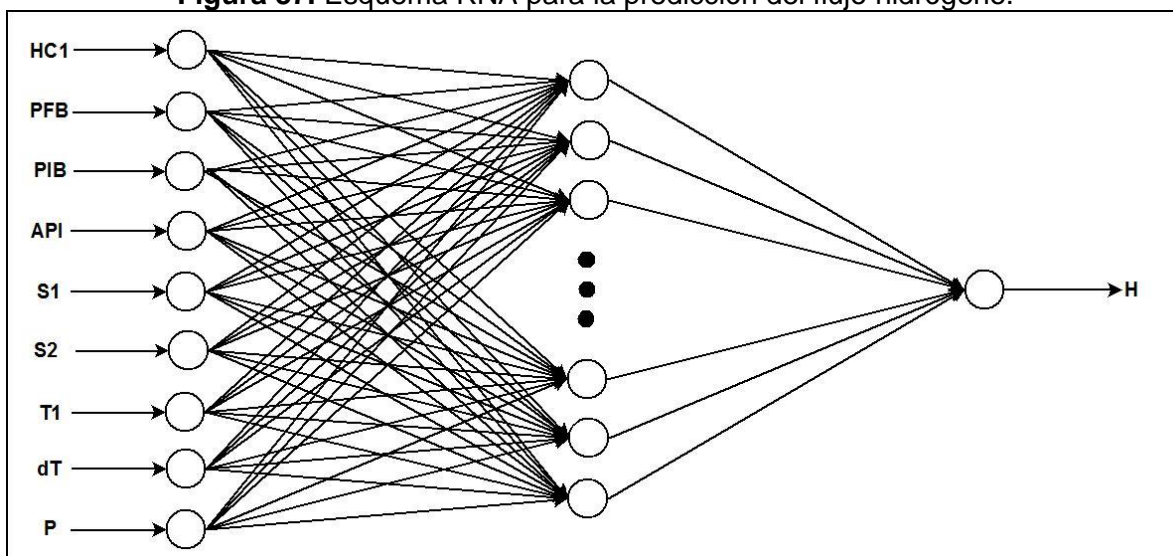


Figura 38. Esquema RNA para la predicción del cambio de temperatura en la zona de reacción.

