



## DESARROLLO DE UN PROGRAMA PARA ESTUDIAR EL COMPORTAMIENTO DE UNA COLUMNA DE FRACCIONAMIENTO ETANO/ETILENO DE UNA PLANTA DE OLEFINAS

(Development of a program for studying the behavior of an ethane - ethylene splitter of an olefins plant)

Recibido: 05/03/2012 Aceptado: 07/09/2012

**Gonz lez, Patricia**

Universidad del Zulia, Venezuela

[gonzalezpcj@gmail.com](mailto:gonzalezpcj@gmail.com)

**Alciaturi, Carlos**

Instituto Zuliano de Investigaciones Tecnol gicas, Venezuela

[alciaturi@hotmail.com](mailto:alciaturi@hotmail.com)

### RESUMEN

En este trabajo se presenta el desarrollo de programas para el estudio del comportamiento de una columna de fraccionamiento Etano/Etileno, perteneciente a una planta de Olefinas. Se utilizaron una red neuronal (RNA) del tipo multicapas con propagaci n hacia adelante y el algoritmo de m nimos cuadrados parciales (PLS) con este prop sito. El entrenamiento de la red neuronal y la obtenci n del modelo PLS se realizaron con datos experimentales de la planta. Estos datos experimentales se dividieron en un conjunto de prueba y un conjunto de validaci n. Se demostr  que la red neuronal fue muy superior al algoritmo PLS para la predicci n de la producci n de etileno a partir de par metros de funcionamiento de la planta. Adicionalmente, estos resultados fueron comparados con los resultados de la simulaci n arrojados por un paquete comercial, resultando en una buena correspondencia entre las predicciones del modelo de red neuronal y las del paquete comercial.

**Palabras clave:** Red neuronal artificial (RNA), M nimos cuadrados parciales (PLS), Columna de fraccionamiento, Etileno.

### ABSTRACT

This paper presents the development of programs for the study of the behavior of a fractionating column of Ethane/Ethylene, which belongs to an olefin plant. We used a neural network (ANN) multilayer type with forward propagation algorithm and partial least squares (PLS) for this purpose. The training of the neural network and the PLS model obtaining were performed with experimental data of the plant. These experimental data were divided into a test set and a validation set. It was shown that the neural network was much higher than PLS algorithm for predicting the production of ethylene from operating parameters of the plant. Additionally, these results were compared with simulation results achieved by a commercial package, resulting in a good agreement between predictions of the neural network model and the commercial package ones.



**Keywords:** Artificial neural network (ANN), Partial least squares (PLS), Fractionating column, Ethylene.

## INTRODUCCIÓN

Las plantas de Olefinas producen principalmente etileno, de gran importancia para la industria debido a que constituye la materia prima fundamental para la producción de diversos polímeros, el más conocido es el polietileno. La obtención final de etileno se lleva a cabo en la columna de fraccionamiento Etano/Etileno, que tiene como función principal lograr la separación del etileno como producto principal de los diversos subproductos que se forman en las diversas reacciones de craqueo en los hornos de pirólisis.

El desarrollo de procesos y el constante requerimiento de productividad llevan a incrementar la complejidad de las unidades industriales. En la industria petroquímica, es absolutamente necesario contar con modelos que permitan el estudio de los equipos involucrados en la planta, con la finalidad de incrementar la productividad de la planta, controlar los procesos que se ahí se llevan a cabo y detectar anomalías tan pronto como sea posible para prevenir riesgos y accidentes. Debido a lo antes mencionado es importante disponer de modelos que permitan estudiar el comportamiento de esta columna y evaluar su funcionamiento para implementar mejoras.

Diversos modelos empíricos han sido utilizados para el modelado de procesos, demostrando ser una herramienta valiosa para obtener de forma rápida y precisa estimados de variables y ofrecer una alternativa a los modelos matemáticos rigurosos. Entre los que se encuentra, el método de mínimos cuadrados parciales (AlGhazzawi y Lennox, 2009; Kano y otros 2000), y las redes neuronales artificiales, utilizadas ampliamente para la predicción de variables, control de procesos de sistemas no lineales (Oller, Giudici y Guardani, 2000; Motlaghi, Jalali y Ahmadabadi, 2008).

En este proyecto (González, 2009) se desarrollaron programas basados en los métodos empíricos mencionados anteriormente para estudiar el comportamiento de una columna de fraccionamiento Etano/Etileno de una planta de Olefinas ubicada en Venezuela, para realizar evaluaciones que mejoren el funcionamiento de esta columna.

## METODOLOGÍA

### PRETRATAMIENTO DE LA BASE DE DATOS

Se realizó una recopilación de datos históricos y una discriminación de los datos, omitiendo valores que se encuentran fuera del intervalo de medición del instrumento o para los cuales se considera que la planta está fuera de servicio. Para la selección de las variables independientes se utilizó el método del conocimiento previo del sistema que está siendo modelado. Las variables para llevar a cabo el entrenamiento de la red neuronal y la aplicación del algoritmo de mínimos cuadrados parciales se seleccionan considerando la influencia que tienen en el proceso, para que se cumpla de manera apropiada la estimación de la producción de etileno.



Con la finalidad de determinar el desfase en el tiempo (time lag) que mejor describiera el proceso, se desarrolló un algoritmo que permite calcular el coeficiente de correlación para distintos corrimientos de tiempo. Del total de los valores recolectados se eliminaron los valores que se encontraban aislados (outliers) aplicando el estadístico  $T^2$  de Hotelling (Macho, 2002).

La base de datos quedó estructurada en dos grupos, el primero conformado por el 85% de los valores totales para la etapa de entrenamiento y el segundo, con el 15% restante, para la etapa de validación. El método de Kennard-Stone fue utilizado para escoger los datos para cada fase. Este método selecciona un subgrupo de  $n$  objetos dimensionales, distribuidos uniformemente en el espacio experimental basado en la distancia entre muestras (Kennard y Stone, 1969).

### **DISEÑO Y ENTRENAMIENTO DE LA RED NEURONAL ARTIFICIAL**

Las Redes Neuronales Artificiales (RNA) son sistemas de procesamiento de la información cuya estructura y funcionamiento están inspirados en las redes neuronales biológicas. Consisten en un conjunto de elementos simples de procesamiento llamados neuronas conectadas entre sí por conexiones que tienen un valor numérico modificable llamado peso (Montaño, 2002).

En este paso se determinó la arquitectura de la red neuronal artificial que se ajusta satisfactoriamente a los datos de entrenamiento y validación. Para la realización del entrenamiento y posterior validación, se desarrolló un algoritmo. Este programa permite el entrenamiento y validación de redes neuronales artificiales, a la vez brinda la posibilidad de variar el número de capas ocultas, funciones de transferencia y número de neuronas en cada capa.

Para entrenar la red se utiliza el primer grupo de datos de entrada-salida (datos de referencia), fijando 50 corridas y un máximo de 100 épocas o ciclos de actualización de los pesos y bias, hasta encontrar el mejor ajuste entre la salida predicha por la red y los patrones de producción de Etileno gaseoso a la salida del sistema de fraccionamiento etano/etileno, utilizando para la comparación el método de la recta de  $45^\circ$  de los datos reales superpuestos a los valores calculados por la red.

La validación de la red se lleva a cabo con el conjunto de datos de validación, basada en los mismos métodos de comparación utilizados en la etapa de entrenamiento. La mejor estructura de la red para el proceso estudiado, se determina variando las funciones de transferencia y el número de neuronas en cada capa, hasta lograr el objetivo propuesto de la menor suma de los errores al cuadrado.

### **APLICACIÓN DEL ALGORITMO DE MÍNIMOS CUADRADOS PARCIALES**

El método de mínimos cuadrados parciales (PLS) es una extensión del modelo de regresión lineal múltiple y se encuentra relacionado con el análisis de componentes principales (PCA). En su forma más simple un modelo lineal especifica la relación (lineal) entre la variable dependiente  $Y$ , y un grupo de variables predictivas  $X$ .



En este paso se utiliz  el programa del algoritmo de m nimos cuadrados parciales desarrollado por Alciaturi y otros (2010). Este programa permite obtener el vector de regresi n  ptimo con los datos de referencia para luego validarlo con los datos de validaci n. Este programa permite encontrar el n mero de variables latentes que est n m s correlacionadas con la variable respuesta. El n mero  ptimo de variables latentes corresponde al que arroje el menor valor en la suma de los cuadrados del error general. Los resultados obtenidos corresponden a los calculados con el n mero  ptimo de variables latentes. Se utiliz  para la comparaci n el m todo de la recta de 45  y la desviaci n cuadr tica media.

### APLICACI N DEL PAQUETE COMERCIAL DE SIMULACI N COMO HERRAMIENTA COMPARATIVA

En este pas  se utiliz  una simulaci n de la columna de fraccionamiento etano/etileno realizada previamente en un paquete comercial de simulaci n. El resultado obtenido con los datos de la fase de validaci n, fue comparado con aquellos obtenidos con la red neuronal y los m nimos cuadrados parciales. Se utiliz  para la comparaci n el m todo de la recta de 45  y la desviaci n cuadr tica media.

### RESULTADOS

#### PRETRATAMIENTO DE LA BASE DE DATOS

✓ Las variables seleccionadas para estudiar el comportamiento de la columna de fraccionamiento Etano/Etileno se muestran en la Tabla 1.

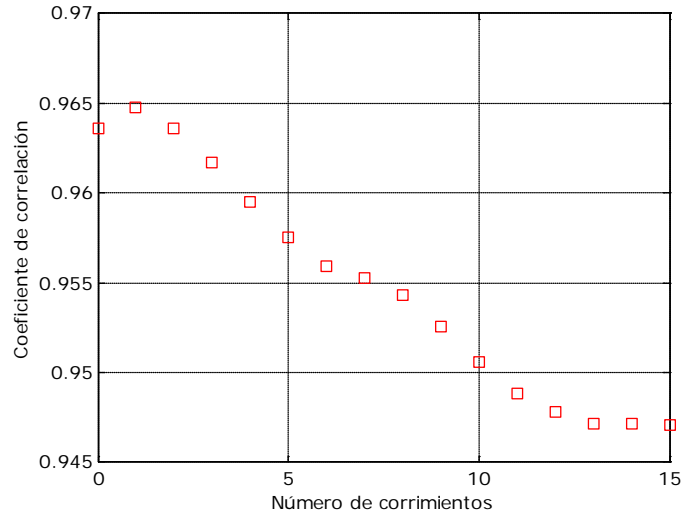
**Tabla 1. Variables seleccionadas para estudiar el comportamiento de la columna de fraccionamiento Etano/Etileno**

Variable	Unidad
Flujo m�sico de etileno gaseoso producido	Kg/s
Temperatura de alimentaci�n a la columna etano/etileno proveniente de la columna demetanzadora	K
Flujo m�sico de refrigerante del rehervidor de la columna etano/etileno	Kg/s
Flujo m�sico de entrada a la columna etano/etileno proveniente de la columna deetanizadora	Kg/s
Flujo m�sico de entrada a la columna etano/etileno proveniente de la columna demetanzadora	Kg/s
Flujo m�sico de refrigerante del rehervidor intermedio de la columna etano/etileno	Kg/s
Flujo m�sico del reflujo de la columna etano/etileno	Kg/s
Presi�n de tope de la columna etano/etileno	kPa

Fuente: elaboraci n propia.

- Para determinar el desfase en el tiempo (time lag), se calculó el coeficiente de correlación para varios corrimientos. Un corrimiento en el tiempo es igual al tiempo de la frecuencia de muestreo, para este estudio corresponde a 900 segundos. En la Figura 1 se observa que el mejor ajuste de los datos para la producción de etileno, correspondiente al máximo valor del coeficiente de correlación, se obtiene para un corrimiento de una unidad de tiempo (900 segundos).

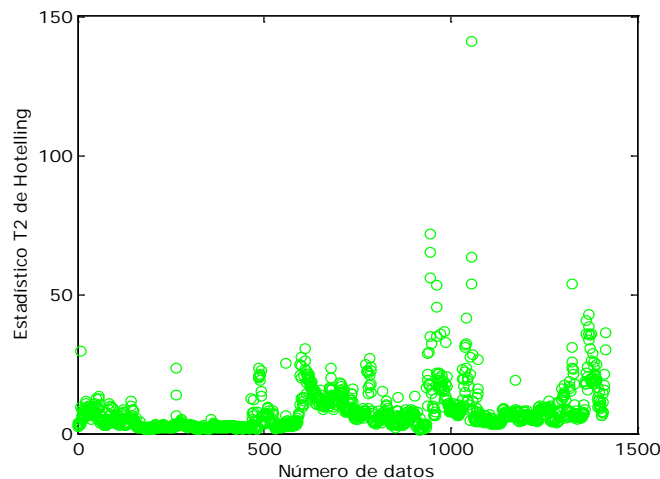
**Figura 1. Coeficiente de correlación para varios corrimientos en el tiempo**



Fuente: elaboración propia.

- La determinación de los datos fuera de rango (outliers) se realizó aplicando el estadístico T2 de Hotelling. Como se observa en la Figura 2, los datos que se encuentran más alejados representan los valores que se encuentran fuera de rango.

**Figura 2. Estadístico T2 de Hotelling de la base de datos**



Fuente: elaboración propia.

- Como último paso de la depuración de datos, se aplicó el método de Kennard-Stone para garantizar una buena distribución de los datos. La base de datos se dividió en dos fases, entrenamiento (85% de los datos) y validación (15% de los datos).

### RED NEURONAL ARTIFICIAL

El diseño de la red neuronal artificial se realizó a partir de una red neuronal multicapa de propagación hacia adelante con el método de aprendizaje de Levenberg-Marquardt (Singh, Gupta y Gupta, 2007). La estructura apropiada de la red se determinó a partir de pruebas de ensayo y error, la función de transferencia utilizada para cada capa se muestra en la Tabla 2.

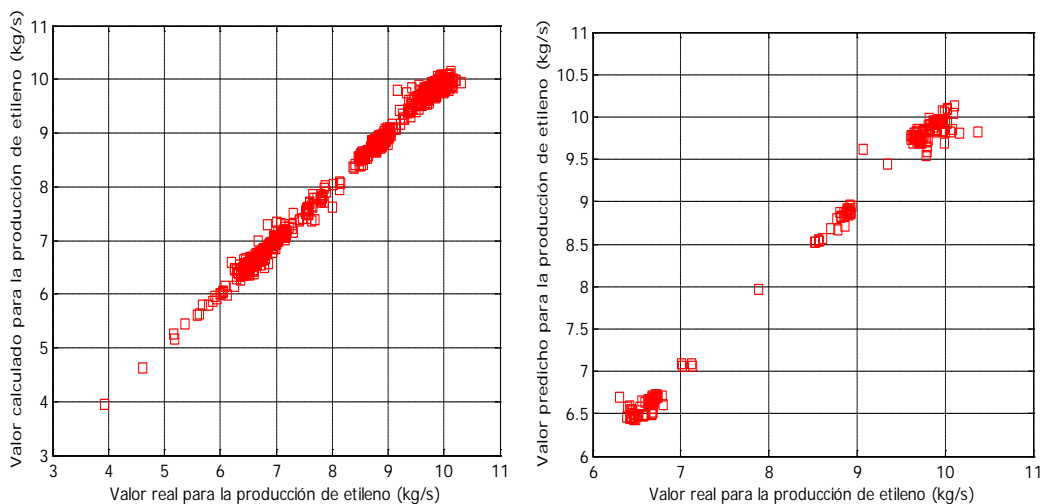
**Tabla 2. Función de transferencia para cada capa**

Capa	Función de Transferencia
Oculto 1	Tansig
Oculto 2	Tansig
Salida	Purelin

Fuente: elaboración propia.

El número de neuronas en la capa oculta 1 se tomó como 10, en la capa oculta 2 se tomó como 15 y en la capa de salida 1. En las Figuras 3 y 4 se muestra el valor real versus el valor calculado para los datos de referencia y el valor predicho para los datos de validación, respectivamente, observándose un buen ajuste en ambos casos demostrando así la potencialidad de las redes neuronales artificiales en la predicción de procesos industriales.

**Figura 3. Valor real versus valor calculado para la producción de etileno: datos de referencia [10,15,1] RNA, y datos de validación [10,15,1] RNA**

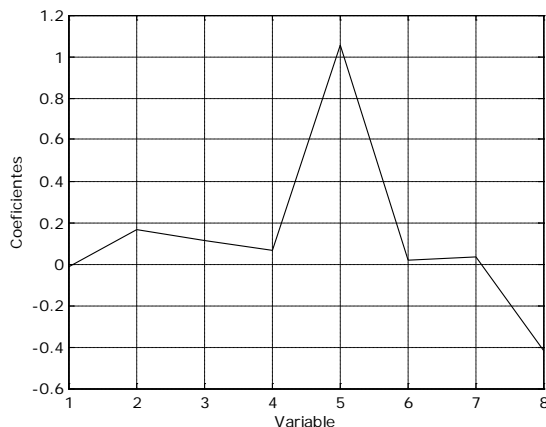


Fuente: elaboración propia.

## M NIMOS CUADRADOS PARCIALES

Se aplic  el algoritmo de los m nimos cuadrados parciales tomando los datos del grupo de referencia y luego se procedi  a validarlo con los datos del grupo de validaci n, en la Figura 5 se observa el vector de regresi n  ptimo obtenido.

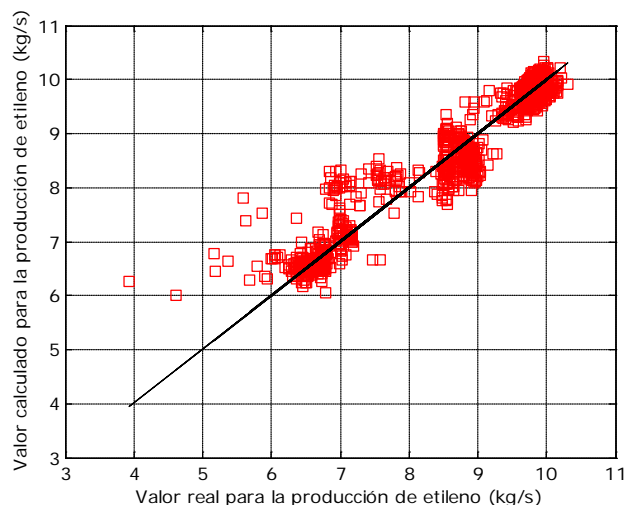
**Figura 5. Vector de regresi n  ptimo**



Fuente: elaboraci n propia.

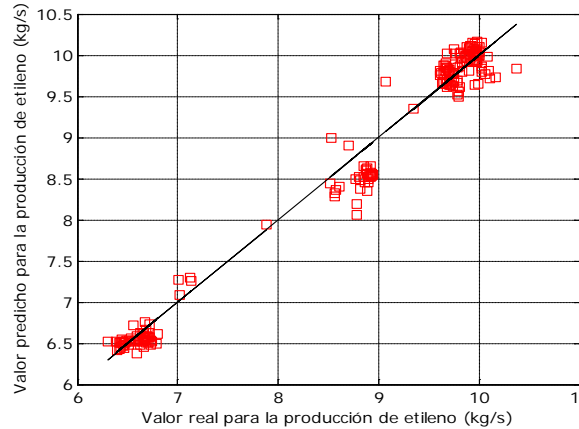
En las Figuras 6 y 7 se muestra el valor real versus el valor calculado para los datos de referencia y el valor predicho para los datos de validaci n, respectivamente. En ambas figuras se observa una desviaci n entre los valores reales y los calculados, haci ndose mayor para valores por debajo de 6.5 kg/s de producci n de etileno de los valores de referencia.

**Figura 6. Valor real versus valor calculado para la producci n de etileno. Datos de referencia. PLS**



Fuente: elaboraci n propia.

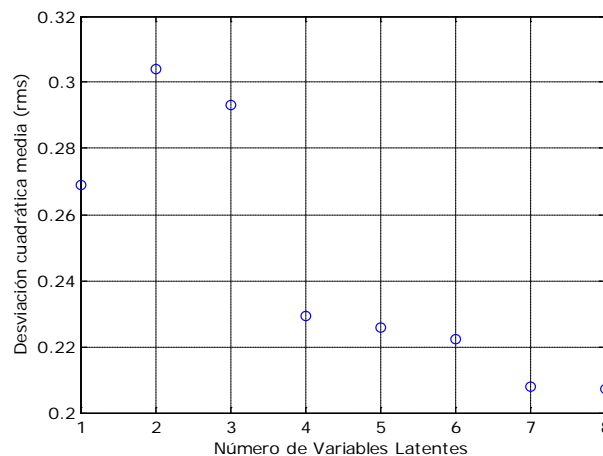
**Figura 7. Valor real versus valor predicho para la producción de etileno. Datos de validación. PLS**



Fuente: elaboración propia.

El número óptimo de variables latentes a utilizar se determinó por el método explicado anteriormente. El resultado para la producción de etileno se muestra en la Figura 8, donde el número óptimo de variables latentes es 8.

**Figura 8. Desviación cuadrática media versus el número de variables latentes para la producción de etileno predicho por el algoritmo de mínimos cuadrados parciales**



Fuente: elaboración propia.

### PAQUETE COMERCIAL DE SIMULACIÓN COMO HERRAMIENTA COMPARATIVA

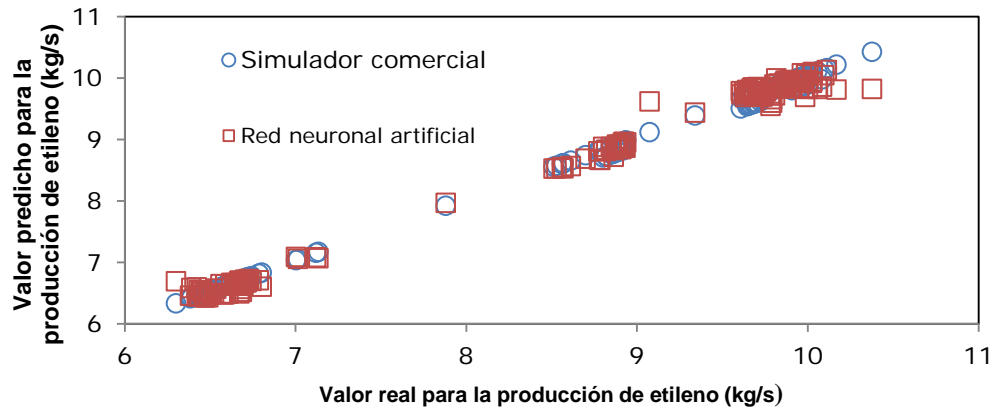
Se aplicó una simulación, previamente realizada con un paquete comercial de simulación, con la finalidad de comparar con los resultados obtenidos con redes neuronales y mínimos cuadrados parciales. La simulación fue obtenida empleando un modelo semiriguroso. En la Figura 9 y 10, se muestra la comparación de los resultados del simulador con las redes neuronales artificiales y los mínimos cuadrados parciales, respectivamente.



Se observa, como es de esperarse, los resultados obtenidos con la simulación predicen satisfactoriamente la producción de etileno. Los paquetes comerciales de simulación se basan en balances de energía, de masa, entre otros, con un enfoque riguroso y han sido utilizado por años para el modelado de procesos, con la desventaja que requiere de un buen conocimiento del proceso que se está modelando y el desarrollo de estos modelos puede requerir mucho tiempo y dinero.

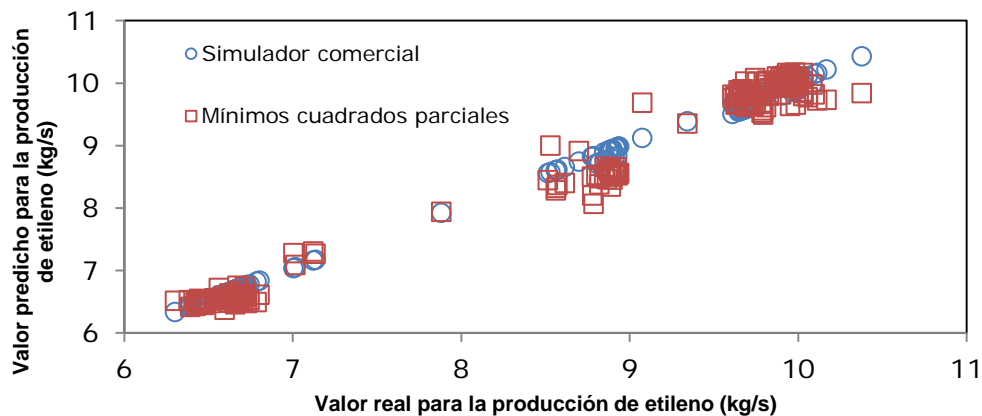
Por otro lado se observa que los resultados obtenidos con las redes neuronales se ajustan satisfactoriamente a los datos reales y a los del simulador, en cambio los resultados obtenidos con los mínimos cuadrados parciales presentan desviaciones con los datos de planta y con los obtenidos con el simulador.

**Figura 9. Valor real versus valor predicho para la producción de etileno. Datos de validación. Simulador comercial y Redes neuronales artificiales**



Fuente: elaboración propia.

**Figura 10. Valor real versus valor predicho para la producción de etileno. Datos de validación. Simulador comercial y Mínimos cuadrados parciales**



Fuente: elaboración propia.



En la Tabla 3 se muestra la desviaci n media cuadr tica obtenida con cada modelo para los datos del grupo de validaci n. Se observa que la mayor desviaci n se presenta para el modelo de m nimos cuadrados parciales, mientras que para la red neuronal y el paquete comercial de simulaci n es menor.

**Tabla 3. Desviaci n media cuadr tica obtenida por cada modelo**

Modelo	Desviaci�n media cuadr�tica (rms)
<b>Redes Neuronales Artificiales</b>	0.10499
<b>M�nimos Cuadrados Parciales</b>	0.2074
<b>Paquete Comercial de Simulaci�n</b>	0.0603

Fuente: elaboraci n propia.

## CONCLUSIONES

Despu s del estudio realizado y en funci n de los resultados obtenidos se llega a las siguientes conclusiones:

Se logr  describir de manera satisfactoria las variables involucradas en el proceso de fraccionamiento, identificando aquellas que contienen informaci n relevante para el proceso y eliminando aquellas que estaban correlacionadas, con ruido o no ten an una relaci n significativa con la producci n de etileno.

Despu s del proceso de entrenamiento, la red neuronal, fue capaz de capturar las caracter sticas operacionales de la columna de fraccionamiento para predecir de manera satisfactoria la producci n de etileno. Utilizando una red neuronal del tipo multicapas con propagaci n hacia adelante, con el m todo de aprendizaje de Levenberg-Marquardt, la estructura de la red qued  establecida con dos capas ocultas, la primera con 10 neuronas, mientras que la segunda capa oculta posee 15 neuronas. Por su parte, la capa de salida posee una sola neurona debido a que solo existe una salida.

El modelo obtenido con el m todo de m nimos cuadrados parciales present  ciertas desviaciones con respecto a los datos reales de la producci n de etileno, siendo mucho menos predictivo que el modelo de red neuronal.

Al comparar los resultados de la red neuronal y los m nimos cuadrados parciales, con el resultado obtenido con un paquete comercial de simulaci n, se concluy  que las redes neuronales representan una alternativa efectiva para la predicci n de la producci n de etileno, ya que el modelo se adapta de manera satisfactoria tanto a los datos reales como a los obtenidos por medio del paquete comercial de simulaci n.

## REFERENCIAS BIBLIOGR FICAS

Alciaturi, C.; Escobar, M.; De la Cruz, C. y Rinc n, C. (2003). Partial least squares (PLS) regression and its application to coal analysis. Rev. Tec. Univ. Zulia. Vol 26. No 23.



- AlGhazzawi, A. y Lennox, B. (2009). Model predictive control monitoring using multivariate statistics. *Journal of Process Control*, volume 19, issue 2, pages 314-327.
- González, P. (2009). Desarrollo de un programa para estudiar el comportamiento de una columna de fraccionamiento etano/etileno de una planta de olefinas. Trabajo Especial de grado. Universidad del Zulia. Venezuela.
- Kano, M.; Miyazaki, K.; Hasebe, S. y Hashimoto, I. (2000). Inferential control system of distillation compositions using dynamic partial least squares regression. *Journal of Process Control*, 10, 157-166.
- Kennard, R. y Stone, L. (1969). Computer aided design of experiments. *Technometrics*, 11:137-148.
- Macho, S. (2002). Metodologías analíticas basadas en espectroscopia de infrarrojo y calibración multivariante. Tesis doctoral. Departamento de Química Analítica I Química Orgánica. Universitat Rovirai Virgili. España.
- Montaño, J. (2002). Redes neuronales artificiales aplicadas al análisis de datos. Tesis Doctoral. Facultad de Psicología, Universitat de Les Illes Balears. España.
- Motlaghi, S.; Jalali, F. y Ahmadabadi, M. (2008). An expert system design for a crude oil distillation column with the neural networks model and the process optimization using genetic algorithm framework. *Expert Systems with Applications*, 35, 1540-1545.
- Oller, C.; Giudici, R. & Guardani, R. (2000). Neural network based approach for optimization of industrial chemical processes. *Computers and Chemical Engineering*, 24, 2303-2314.
- Singh, V.; Gupta, I. & Gupta, H. (2007). ANN-based estimator for distillation using Levenberg-Marquardt approach. *Engineering Applications of Artificial Intelligence*, 20, 249-259.