

## Modelación del equilibrio líquido-vapor a presión constante de mezclas etanol-agua utilizando redes neuronales artificiales

### Modeling of liquid-vapor equilibrium at constant pressure for ethanol-water systems, using artificial neural network

**Dra. Lourdes Zumalacárregui de Cárdenas, Dr. Osney Pérez Ones, Ing. Frank Abel Hernández Castellanos, Dr. Gil Cruz Lemus**

Grupo de Análisis de Procesos, Facultad de Ingeniería Química, Universidad Tecnológica de La Habana "José Antonio Echeverría" (CUJAE), La Habana, Cuba e-mail: lourdes@quimica.cujae.edu.cu

---

#### Resumen

En el presente trabajo se modeló el equilibrio líquido-vapor de la mezcla binaria etanol-agua a presión constante, utilizando datos a bajas, moderadas y elevadas presiones extraídas de la literatura, empleando redes neuronales artificiales con arquitectura multicapa perceptrón y algoritmo de aprendizaje "back propagation", implementadas en KNIME 3.1.1 y Matlab 2013. Para determinar la confiabilidad de los datos se examinó la presencia de errores no sistemáticos encontrando imprecisiones en seis puntos experimentales. La presencia de errores sistemáticos se comprobó mediante pruebas de consistencia termodinámica. Se aplicó la prueba de las áreas de Herington y la prueba punto a punto de Wisniak, resultando que los datos experimentales eran confiables para el modelado, excepto dos puntos a presiones bajas. Se obtuvieron topologías desde dos neuronas en la capa oculta hasta 10. Para la selección se aplicaron las pruebas no paramétricas de Friedman y Wilcoxon. La topología seleccionada fue la obtenida en Matlab 2013 de ocho neuronas en la capa oculta con una media del error cuadrático medio de 0,0054, una desviación del error cuadrático medio de 0,0006 y un coeficiente de correlación de 0,9729. De esta forma, con un solo modelo, es posible predecir el equilibrio líquido vapor para el intervalo de presión desde 6,6 kPa hasta 1 520 kPa.

**Palabras clave:** equilibrio líquido- vapor, modelación, redes neuronales artificial

---

## Abstract

The liquid-vapor equilibrium for the binary mixture ethanol-water at constant pressure, using low, moderate and high pressures from the literature was modeled. Artificial neuronal networks with multilayer architecture perceptron and back propagation learning algorithm, implemented in, KNIME 3.1.1 and Matlab 2013, were used. To determine the reliability of the data, the presence of non-systematic errors was examined finding fails in six experimental points. The presence of systematic errors was proven using thermodynamic consistency tests. Herington's areas test and Wisniak's point to point test were applied. Experimental data were reliable, except two points at low pressures. Topologies were obtained from two neurons in the hidden layer up to 10. For the selection, Friedman and Wilcoxon non parametric tests were applied. The selected topology was the one obtained in Matlab 2013 with eight neurons in the hidden layer with a mean of the mean square error of 0.0054, a deviation of the square means error of 0.0006 and a correlation coefficient of 0.9729. With this result, a single model can be used to predict the liquid- vapor equilibrium from 6.6 kPa to 1520 kPa.

**Keywords:** liquid-vapor equilibrium, modeling, artificial neural network

---

## INTRODUCCION

La destilación es uno de los métodos de separación más utilizados en la industria química [1]. Para el adecuado control de esta operación se requiere estimar los parámetros operacionales utilizando modelos cada vez más elaborados y que puedan ofrecer respuestas en tiempo real [2], [3]. En el modelado de procesos, diseño y simulación es esencial disponer de datos precisos sobre las propiedades termodinámicas del sistema estudiado, tales como datos de equilibrio líquido-vapor, propiedades termodinámicas, entre otros. La determinación de los datos del equilibrio líquido-vapor se realiza por diferentes métodos experimentales, como son el método de destilación, el método dinámico, entre otros. Estos métodos demandan equipos con un alto costo de adquisición producto de la necesidad de obtener puntos de equilibrio con la mayor precisión posible [4]. Otra forma de obtener dichos datos es mediante ecuaciones de estado, modelos termodinámicos y modelos estadísticos [5]. La mayoría de estos modelos tiene limitaciones para su uso debido a que no abarcan amplios intervalos de presión.

La calidad final de los resultados en la modelación de un proceso, independientemente del grado de sofisticación de las ecuaciones, depende de la calidad de las predicciones suministradas por el modelo termodinámico. Para la mezcla etanol-agua, que es uno de los sistemas binarios con más estudios reportados debido a su amplio empleo en la industria, la determinación de los datos del equilibrio líquido-vapor precisos, con rapidez y con el menor costo posible es de gran importancia para el diseño, análisis y control de las torres de destilación, parte esencial para la producción de las diferentes calidades de alcohol. [2]

Los modelos termodinámicos clásicos que pueden funcionar razonablemente para este sistema como por ejemplo UNIQUAC (Universal Quasi Chemical Theory), NRTL (Non-random Two Liquid) y Wilson como modelos clásicos de correlación y UNIFAC (Universal Functional Activity Coefficient) como modelo de predicción, requieren de una gran cantidad de parámetros, siendo su determinación, el principal problema para el correcto modelado de estos sistemas. Los parámetros de los modelos termodinámicos se calculan usualmente a partir de datos experimentales de equilibrio líquido-vapor, datos que están dispersos e incompletos en la literatura [6]. El sistema binario etanol-agua no está excluido de esta problemática.

En estos días han crecido rápidamente los métodos alternativos para la predicción de datos de ELV y se reportan investigaciones donde se aplican las redes neuronales artificial (RNA) como herramienta de cálculo [1], [7].

Las redes neuronales artificiales son un método sofisticado para procesar información, basado en la semejanza con las neuronas biológicas, en que se realiza el procesamiento de forma simultánea partiendo de la sincronización de las neuronas y no de forma secuencial. Una de sus principales ventajas es la capacidad para establecer relaciones no lineales entre variables independientes y dependientes, algo que es de mucha utilidad para añadir variables y proveer información [8]. Estas redes son capaces de aprender mediante conocimientos previos y de generalizar, a partir de casos precedentes, a casos nuevos.

La arquitectura del Perceptrón multicapa, muy común en las redes neuronales, se ha usado en la determinación del equilibrio líquido-vapor, para diferentes mezclas binarias y ternarias. Esta estructura también se aprovecha para la determinación de propiedades de mezclas, para la estimación de los coeficientes de actividad y la predicción de la temperatura de ebullición de una mezcla binaria de alcoholes y agua [9]. El Perceptrón multicapa es una red de alimentación hacia adelante compuesta por una capa de unidades de entrada (sensores), otra capa de unidades de salida y un número determinado de capas intermedias de unidades de proceso, conocidas como capas ocultas [10], [11].

Esta arquitectura se usó para para la predicción del ELV en soluciones acuosas de electrolitos [5], para correlacionar la fracción molar de dos hidrocarburos en la fase líquida [12], para la predicción de propiedades como la viscosidad, densidad, entre otras [13], [14], [15].

Para el entrenamiento de las redes neuronales, uno de los métodos más usados es el de propagación hacia atrás o "backpropagation". Este método se ha utilizado en el entrenamiento de redes para obtener la fracción molar en la fase líquida ( $x_1$ ) y en la vapor ( $y_1$ ) del CO<sub>2</sub> contenido en mezclas binarias [9], para predecir la densidad de las mezclas de líquidos iónicos y agua [16], en la estimación de la solubilidad del cloruro de sodio en mezclas de agua y etanol [17] y en la predicción de otras propiedades termodinámicas utilizando redes neuronales [18].

En el presente trabajo, se estableció como objetivo general estimar datos del equilibrio líquido-vapor para la mezcla binaria etanol-agua en un amplio intervalo de presiones, utilizando redes neuronales artificiales. Las variables independientes consideradas fueron la temperatura y la presión, siendo las dependientes la composición en la fase líquida y la fase vapor. Se estudiaron dos diferentes herramientas que aplican redes neuronales artificiales, KNIME 3.1.1 y Matlab 2013 determinando que con el uso de Matlab 2013 se obtienen los menores errores para las arquitecturas estudiadas.

## **MATERIALES Y MÉTODOS**

### **Comprobación de la calidad de los datos del ELV del sistema etanol-agua utilizados**

La base de datos utilizada contiene 534 juegos de valores de composición en fase líquida, composición en fase vapor y temperatura, agrupados a presión constante, que abarcan un intervalo desde las presiones subatmosféricas (6,6 kPa) hasta presiones elevadas (1519,87 kPa) [19].

### **Determinación de errores no sistemáticos**

Para determinar la presencia de errores no sistemáticos se representó la volatilidad relativa ( $a_{12}$ ) en función de las composiciones de la fase líquida ( $x_1$ ), siendo 1 el componente etanol y 2 el componente agua. La volatilidad relativa se calculó a partir de los datos experimentales según (1). Los puntos que no se encontraron sobre la curva se eliminaron con el objetivo de evitar futuros errores en el modelado.

$$\alpha_{12} = \frac{y_1 x_2}{y_2 x_1} \quad (1)$$

### Determinación de errores sistemáticos

Para comprobar la consistencia termodinámica de los datos experimentales extraídos de la literatura fue necesaria la búsqueda de propiedades termodinámicas, que se requerían para el desarrollo de las pruebas.

Las presiones de vapor se calcularon empleando la ecuación de Antoine de siete parámetros (2). En la tabla 1 se muestran los valores de las constantes para el etanol y el agua.

$$\ln(P^0 \text{ kPa}) = A + \frac{B}{T+C} + D \ln T + ET^F \quad (2)$$

Donde:

A, B, C, D, E y F son las constantes de la ecuación de Antoine extendida.

**Tabla 1.** Constantes de la ecuación de Antoine extendida según [20]

Componente	A	B	C	D	E	F
etanol	86,49	-7931	0	-10,25	6,389E-06	2
agua	65,93	-7228	0	-7,177	4,031E-06	2

Los coeficientes de actividad experimentales se evaluaron a partir de la expresión (3) de la ley de Raoult modificada.

$$y_i P = \gamma_i x_i P_i^0 \quad (3)$$

Se aplicaron dos pruebas de consistencia termodinámica, una global, la prueba de Herington [21] y una prueba punto a punto propuesta por Wisniak [22].

Según la prueba de Herington para un sistema binario a presión constante, si se grafica  $\ln(\gamma_1/\gamma_2)$  contra  $x_1$  las áreas generadas por encima y debajo de  $y=0$  permiten evaluar los parámetros  $D_{RK}$  y  $J$  según (4) y (5).

$$D_{RK} = \frac{(\text{área sobre el eje } x - |\text{área bajo el eje } x|)}{(\text{área sobre el eje } x + |\text{área bajo el eje } x|)} 100 \quad (4)$$

$$J = 150 \frac{T_{\text{máx}} - T_{\text{mín}}}{T_{\text{mín}}} \quad (5)$$

Donde  $T_{\text{máx}}$  son las temperaturas máximas y  $T_{\text{mín}}$  de ebullición de la mezcla en Kelvin (K).

Si  $D_{RK} < J$ , los datos son termodinámicamente consistentes, pero si  $D_{RK} > J$ , pero  $|D_{RK} - J| < 10$ , los datos son probablemente consistentes.

Con el objetivo de poder eliminar puntos que pueden causar errores en el aprendizaje, entrenamiento o validación se aplicó además un método para estimar la consistencia termodinámica punto a punto propuesto por Wisniak. Este método se basa en el cálculo de la relación  $L_i/W_i$  comprobándose la confiabilidad de los datos de equilibrio cuando esta función es = 2. Este método se basa en relacionar la función energía libre en exceso ( $G^E$ ) con el punto de ebullición de la

mezcla. Las expresiones de cálculo de los parámetros  $Li$  y  $Wi$  se deducen a partir de la expresión que describe el punto de burbuja de la mezcla. [22] (Wisniak, 1993).

$$Li = \sum T_i^0 x_i \Delta s_i^0 \Delta s - T \quad (6)$$

$$Wi = \frac{G^E}{\Delta s} - RT \frac{w}{\Delta s} \quad (7)$$

$$\Delta s = \sum x_i \Delta s_i^0 \quad (8)$$

$$w = \sum x_i \ln(y_i/x_i) \quad (9)$$

Donde:

$\Delta s_i^0$  Variación de entropía molar de vaporización del componente  $i$

$\Delta s$ : Variación de entropía molar de vaporización de la mezcla

Las entropías molares de vaporización se calcularon a partir de las variaciones de entalpías molares de vaporización ( $H_v$ ), utilizando (10). Los datos de las constantes se reportan en la Tabla 2.

$$H_v \left( \frac{J}{\text{kmol}} \right) = C_1 (1 - Tr)^{C_2 + C_3 Tr + C_4 Tr^2} \quad (10)$$

**Tabla 2.** Coeficientes para evaluar entalpías de vaporización según [23]

Componente	C 1	C 2	C 3	C 4
etanol	5,5789E+07	0,31245	0	0
agua	5,2053E+07	0,31990	-0,212	0,25795

### Modelación usando KNIME 3.1.1

Como las variables de entrada pueden tener diferencias de valores de varios órdenes de magnitud se requirió la normalización por el máximo quedando todas las variables entre los valores de 0 y 1. De lo contrario, el aprendizaje de la red se vería influido por el valor absoluto de la variable. Los datos se dividieron en 10 conjuntos iguales, usando un 90 % para el aprendizaje y un 10 % para la validación. Para la topología de la red se consideró una capa de entrada, una de salida y una capa oculta. Se varió la cantidad de neuronas en esta capa oculta entre 2 y 10. Se recogieron los resultados del error cuadrático medio, la desviación estándar del error cuadrático medio y del coeficiente de correlación para cada arquitectura.

El número de iteraciones en el entrenamiento se varió en el intervalo de 5 000 a 10 000. Este procedimiento se realizó para obtener dos modelos, uno para predecir la composición en la fase líquida ( $x_1$ ) y otro para la fase vapor ( $y_1$ ).

### Modelación utilizando Matlab 2013

Para la creación del modelo se utilizó la caja de herramientas "Neural Network Toolbox" de Matlab. En todos los casos las redes son unidireccionales (comando newff) y se realizó el entrenamiento con el algoritmo Levenberg-Marquardt de propagación hacia atrás o "backpropagation" (comando trainlm). Como funciones de transferencia se utilizaron la sigmoidea (tansig) en la capa oculta y la función lineal (purelin) para la capa de salida. Las variables se normalizaron en un intervalo entre 0 y 1. Para la topología de la red se consideró una capa de entrada, una de salida y una capa oculta. Las arquitecturas se crearon variando las neuronas de la capa oculta desde 2 hasta 10. Se

recogieron los resultados del error cuadrático medio, la desviación estándar del error cuadrático medio y del coeficiente de correlación para cada arquitectura.

### **Selección de la mejor topología**

Para la selección de la mejor topología se utilizó el error cuadrático medio. Se comprobó si existían diferencias estadísticamente significativas entre el error cuadrático medio de las diferentes topologías creadas en KNIME y Matlab a partir de la prueba no paramétrica de Friedman para la comparación de las medianas, en la herramienta Statgraphics Centurion 16.2. En caso de obtener un valor-P menor que 0,05 se concluyó que existen diferencias significativas, sin identificar cuáles son las topologías significativamente diferentes. En caso de que entre las muestras analizadas no existan diferencias significativas se analizaron otros parámetros como la complejidad de la topología, el coeficiente de correlación entre los valores estimados por la red y los valores reales de las variables de salida (composición en la fase líquida y vapor) en la validación y la desviación estándar del error cuadrático medio en las validaciones cruzadas (medida de la capacidad de generalización de la red neuronal).

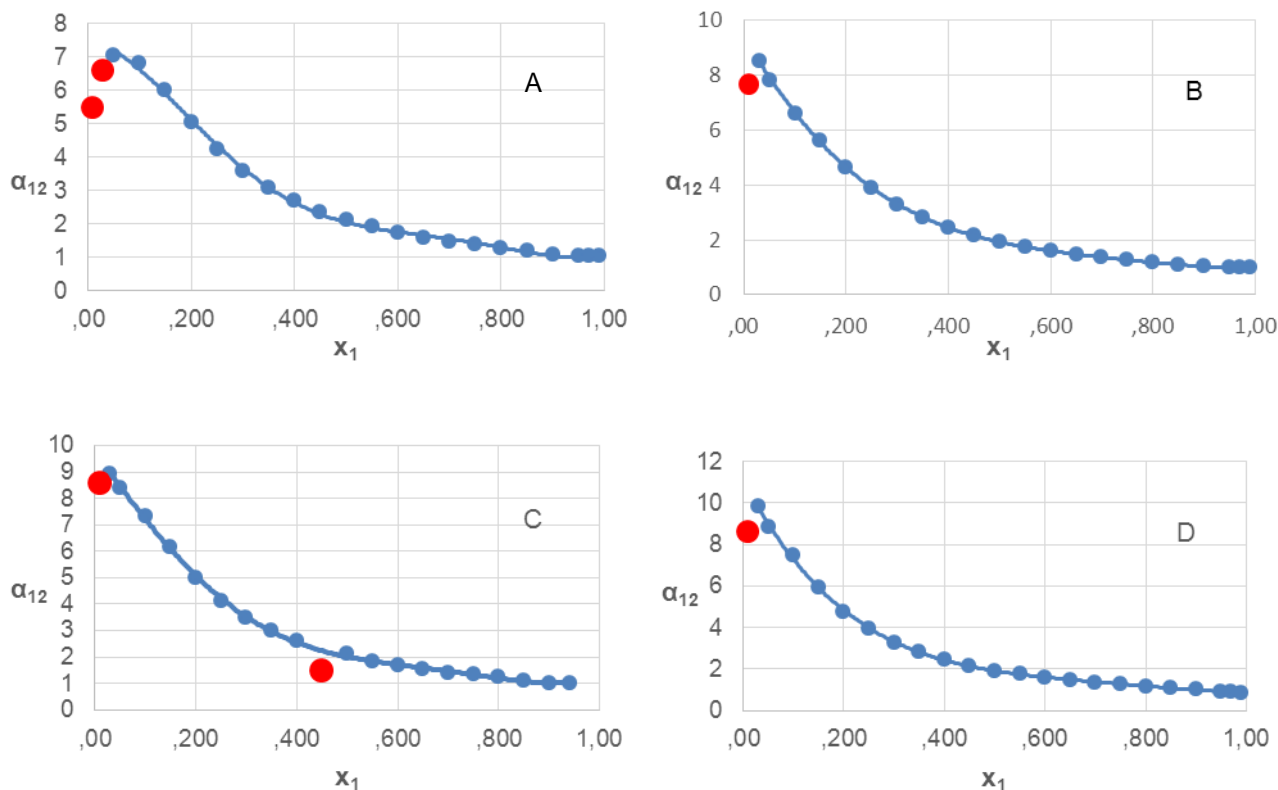
Cuando existieron diferencias significativas, para encontrar las que eran diferentes, se aplicó la prueba de Mann-Whitney (de Wilcoxon) en la herramienta Statgraphics Centurion 16.2. Cuando el valor-P fue menor que el valor de probabilidad establecido (0,05) se rechazó la hipótesis nula que plantea esta prueba estadística para la comparación de las medianas y por tanto se concluyó que existían diferencias significativas entre las dos topologías comparadas.

Después de realizar la prueba de Wilcoxon a todos los posibles pares agrupados en columnas se presentaron dos situaciones: que las topologías significativamente diferentes tengan los mayores errores cuadráticos medios y que las topologías significativamente diferentes tengan los menores errores cuadráticos medios. En la primera situación, se descartó esta topología y con las restantes se hizo un análisis entre la complejidad de la topología, el coeficiente de correlación y la desviación estándar del error cuadrático medio para llegar a una respuesta. En la segunda situación se analizó el coeficiente de correlación (mayor que 0,95) y la desviación estándar del error cuadrático medio (menor que 0,005).

## **RESULTADOS**

### **Determinación de errores no sistemáticos**

En la Figura 1 se ilustra el comportamiento del conjunto de valores a presión de 6,66 kPa, 12,66 kPa, 33,3 kPa y 66,6 kPa. Los puntos resaltados en rojo son los que se encuentran, como se aprecia, fuera de la tendencia global. Para las restantes presiones, no se apreciaron puntos discrepantes de la tendencia general. Por consiguiente de los 534 juegos de datos iniciales se desearon los 6 discrepantes, quedando un total de 528 juegos de datos.



**Figura 1.** Gráfico de relación de  $\alpha_{12}$  v/s  $x_1$  a diferentes presiones. A: 6,66 kPa; B:12,66 kPa; C:33,33 kPa; D:66,66 kPa.

### Determinación de los errores sistemáticos

#### Prueba de Herington

En las tablas 3, 4 y 5 se presenta el resumen de los resultados de la aplicación del método de Herington para las bajas, moderadas y altas presiones.

**Tabla 3.** Aplicación del método de Herington en el intervalo de bajas presiones

	Presiones kPa								
	6,66	12,66	13,33	25,33	33,33	50,66	66,66	98,65	101,32
<b>% D RK</b>	0,700	9,120	0,208	7,739	8,540	5,320	4,884	4,862	4,329
<b>J</b>	8,400	6,290	8,400	7,715	8,514	8,411	8,966	9,434	9,163

**Tabla 4.** Aplicación del método de Herington en el intervalo de moderadas presiones

	Presiones kPa						
	202,65	303,975	405,30	506,62	607,95	709,27	810,60
<b>% D RK</b>	7,56	15,71	18,51	17,71	18,80	14,94	16,16
<b>J</b>	10,13	10,47	10,74	10,77	10,87	10,82	10,68

**Tabla 5.** Aplicación del método de Herington en el intervalo de altas presiones

	Presiones kPa						
	0,911	1,013	1, 116	1, 217	1, 317	1, 418	1, 520
<b>% D<sub>RK</sub></b>	16,20	16,45	16,11	14,12	17,15	18,15	19,73
<b>J</b>	8,96	10,63	10,98	10,71	10,79	10,89	10,73

Al comparar ambos parámetros se aprecia que, en el intervalo de presiones estudiado, se cumple la condición de que el parámetro  $D_{RK} < J$  o que  $|D_{RK} - J| < 10$ . Por consiguiente, el conjunto de datos cumple la prueba de las áreas, por lo que es consistente termodinámicamente. A modo de confirmación, dado que esta prueba evalúa la consistencia global, se procedió a realizar la prueba punto a punto de Wisniak.

### Método punto a punto de Wisniak

De acuerdo a la prueba punto a punto de Wisniak la confiabilidad de los datos de equilibrio se tiene si la relación  $L_i/W_i = 2$ . El cálculo de la relación  $L_i/W_i$  fue inferior a 2 para todas las presiones y puntos experimentales con excepción del punto correspondiente a  $x_1 = 0,97$  a la presión de 6,66 kPa en que la relación fue 3,3399 y el punto correspondiente a  $x_1 = 0,99$  a la presión 101,3 kPa en que resultó 4,5478. Por consiguiente, esos dos puntos se eliminaron de los juegos de datos, quedando finalmente 526 juegos de datos válidos.

### Resultado del modelado con KNIME 3.1.1

Las topologías obtenidas en la predicción de la composición en la fase líquida  $x_1$  se muestran en la Tabla 6. Se trabajó con tres capas, la de entrada, la de salida y una capa oculta, a la cual se le fueron variando las neuronas desde 2 hasta 10. Se reporta la media del error cuadrático medio (MECM), la desviación estándar del error cuadrático medio (DEECM) y el coeficiente de correlación (R)

**Tabla 6.** Comportamiento de la topología en la predicción de  $x_1$  utilizando KNIME 3.1.1

Topologías	MECM	DEECM	R
2-2-1	0,0210	0,0088	0,855
2-3-1	0,0112	0,0056	0,898
2-4-1	0,0072	0,0020	0,922
2-5-1	0,0070	0,0039	0,903
2-6-1	0,0069	0,0023	0,915
2-7-1	0,0060	0,0038	0,929
2-8-1	0,0050	0,0025	0,959
2-9-1	0,0060	0,0023	0,937
2-10-1	0,0064	0,0036	0,968

Las arquitecturas obtenidas en la predicción de la composición en la fase vapor y  $x_1$ , presentaban tres capas, la de entrada, la de salida y una capa oculta, a la cual se le fueron variando las neuronas desde 2 hasta 10. La Tabla 7 lista el comportamiento de las topologías en cuanto a la MECM, DEECM y R.



**Tabla 7.** Comportamiento de la topología en la predicción de y 1 utilizando KNIME 3.1.1

<b>Topologías</b>	<b>MECM</b>	<b>DEECM</b>	<b>R</b>
2-2-1	0,0146	0,0041	0,843
2-3-1	0,0076	0,0037	0,885
2-4-1	0,0055	0,0009	0,928
2-5-1	0,0065	0,0022	0,894
2-6-1	0,0058	0,0010	0,921
2-7-1	0,0051	0,0014	0,965
2-8-1	0,0052	0,0013	0,953
2-9-1	0,0052	0,0011	0,937
2-10-1	0,0052	0,0012	0,958

### **Resultados de la modelación utilizando Matlab 2013**

En la Tabla 8 se presenta el resultado de la modelación utilizando Matlab 2013 en cuanto a la MECM, DEECM y R. En esta ocasión fue posible predecir con una sola red neuronal los valores de x 1 y y 1 , por lo que las topologías tienen la estructura 2-neuronas en la capa oculta.

**Tabla 8.** Comportamiento de la topología en la predicción de y 1 y x 1 en Matlab 2013

<b>Topologías</b>	<b>MECM</b>	<b>DEECM</b>	<b>R</b>
2-2-2	0,0077	0,0013	0,9062
2-3-2	0,0075	0,0019	0,9518
2-4-2	0,0071	0,0016	0,9686
2-5-2	0,0064	0,0023	0,9665
2-6-2	0,0061	0,0012	0,9550
2-7-2	0,0062	0,0015	0,9676
2-8-2	0,0054	0,0006	0,9729
2-9-2	0,0074	0,0015	0,9631
2-10-2	0,0052	0,0015	0,9766

## **ANÁLISIS DE LOS RESULTADOS**

### **Modelado con KNIME 3.1.1 para x 1**

Para comprobar si entre los errores cuadráticos medios existían diferencias significativas se realizó la prueba de Friedman. La prueba arrojó que se rechazara la hipótesis nula, por lo que existen diferencias estadísticamente significativas entre los errores cuadráticos medios de las topologías, siendo el valor-P igual a 0,0002, menor que 0,05.

Se aplicó la prueba de Wilcoxon a todas las posibles combinaciones de pares de topologías, empezando la comparación con la de mayor error hasta las de menor. Así se fueron eliminando los pares con criterio de igualdad, que se encontraron. La Tabla 9 muestra la comparación de 2-2-1 con las restantes topologías.

**Tabla 9.** Prueba de Wilcoxon entre pares de topologías en relación con 2-2-1 de KNIME 3.3.1 para x  
1

Pares	Valor-P	Comparación	a
2-2-1 & 2-3-1	0,1121	=	0,05
2-2-1 & 2-4-1	0,0006	=	0,05
2-2-1 & 2-5-1	0,0077	=	0,05
2-2-1 & 2-6-1	0,0006	=	0,05
2-2-1 & 2-7-1	0,0006	=	0,05
2-2-1 & 2-8-1	0,0003	=	0,05
2-2-1 & 2-9-1	0,0004	=	0,05
2-2-1 & 2-10-1	0,0015	=	0,05

Los resultados de comparar las topologías 2-2-1 y la 2-3-1 conducen a que no poseen diferencias significativas. Estas presentan los mayores valores de la MECM, se eliminaron de la comparación y se examinó si existen diferencias entre los restantes pares que poseen un valor-P menor que 0,05.

La Tabla 10 muestra la comparación por pares de la topología 2-4-1 con las restantes. Se observa que la 2-10-1 posee un valor-P menor que 0,05, por lo que se concluye que son estadísticamente diferentes y no se acepta la hipótesis nula. Esta topología 2-10-1 es la que posee el mayor valor de R y un valor de DEECM de 0,0036, por lo que se seleccionó, ya que presenta mejores ajustes que la 2-4-1.

**Tabla 10.** Prueba de Wilcoxon entre pares de topologías en relación con 2-4-1 de KNIME 3.3.1 para  
x 1

Pares	Valor-P	Comparación	a
2-4-1 & 2-5-1	0,6228	=	0,05
2-4-1 & 2-6-1	0,9698	=	0,05
2-4-1 & 2-7-1	0,4718	=	0,05
2-4-1 & 2-8-1	0,0958	=	0,05
2-4-1 & 2-9-1	0,0584	=	0,05
2-4-1 & 2-10-1	0,0451	=	0,05

### Modelado con KNIME 3.1.1 para y 1

Las arquitecturas obtenidas en la predicción de la composición en la fase vapor y 1, presentaban tres capas, la de entrada, la de salida y una capa oculta, a la cual se le fueron variando las neuronas desde 2 hasta 10. Se aplicó la prueba de Friedman para buscar similitudes o diferencias entre los valores del ECM al ser las MECM muy semejantes en cuanto al orden. El resultado permite concluir que existen diferencias estadísticamente significativas. Se aplicó la prueba de Wilcoxon comparando la topología 2-2-1 con las restantes al ser esta la de mayor valor MECM (Tabla 11).

**Tabla 11.** Prueba de Wilcoxon entre pares de topologías en relación con 2-2-1 de KNIME 3.3.1 para  
y 1

Pares	Valor-P	Comparación	a
2-2-1 & 2-3-1	0,0022	=	0,05

2-2-1 & 2-4-1	0,0002	=	0,05
2-2-1 & 2-5-1	0,0005	=	0,05
2-2-1 & 2-6-1	0,0002	=	0,05
2-2-1 & 2-7-1	0,0002	=	0,05
2-2-1 & 2-8-1	0,0003	=	0,05

La Tabla 11 demuestra que existen diferencias significativas entre todas las muestras comparadas, por lo que se aplicó nuevamente la prueba con el par siguiente en orden de mayor MECM.

La Tabla 12 muestra que solamente existe una diferencia estadísticamente significativa con la topología 2-10-1. Esta última posee similar valor de DEECM, pero es la de mayor correlación entre los datos experimentales y los predichos (Tabla 7)

**Tabla 12.** Prueba de Wilcoxon entre pares de topologías en relación con 2-3-1 de KNIME 3.3.1 para y 1

Pares	Valor-P	Comparación	a
2-3-1 & 2-4-1	0,0691	=	0,05
2-3-1 & 2-5-1	0,4267	=	0,05
2-3-1 & 2-6-1	0,1038	=	0,05
2-3-1 & 2-7-1	0,0815	=	0,05
2-3-1 & 2-8-1	0,0586	=	0,05
2-3-1 & 2-9-1	0,0637	=	0,05
2-3-1 & 2-10-1	0,0282	=	0,05

### Modelado con Matlab 2013

Las arquitecturas obtenidas en la predicción de la composición en la fase líquida y vapor presentaban tres capas, la de entrada, la de salida y una capa oculta, a la cual se le fueron variando las neuronas desde 2 hasta 10. El resultado de la Prueba de Friedman para los errores cuadrático medios aportó el valor-P de 0,0033 = 0,05, por lo que existían diferencias estadísticamente significativas. Se aplicó la prueba de Wilcoxon. En la Tabla 13 se muestra que existen diferencias estadísticamente significativas entre las medianas de los pares de muestras de las combinaciones 2-2-2 y las 2-6-2, 2-7-2, 2-8-2 y 2-10-2, puesto que el valor-P es menor que 0,05.

**Tabla 13.** Resultados de la prueba de Wilcoxon entre pares de topologías en relación con 2-2-2 de Matlab 2013

Pares	Valor-P	Comparación	a
2-2-2 & 2-3-2	0,7334	=	0,05
2-2-2 & 2-4-2	0,6227	=	0,05
2-2-2 & 2-5-2	0,4490	=	0,05
2-2-2 & 2-6-2	0,0125	=	0,05
2-2-2 & 2-7-2	0,0138	=	0,05
2-2-2 & 2-8-2	0,0004	=	0,05
2-2-2 & 2-9-2	0,3818	=	0,05
2-2-2 & 2-10-2	0,0040	=	0,05

La Tabla 14 muestra la búsqueda de las diferencias entre las cuatro topologías que resultaron diferentes estadísticamente.

**Tabla 14** . Prueba de Wilcoxon entre pares de topologías en relación con 2-7-2 de Matlab 2013

<b>Pares</b>	<b>Valor-P</b>	<b>Comparación</b>	<b>a</b>
2-7-2 & 2-6-2	0,9698	=	0,05
2-7-2 & 2-8-2	0,2892	=	0,05
2-7-2 & 2-10-2	0,1397	=	0,05

Los resultados anteriores demuestran que para la selección de la mejor topología es necesario recurrir a la comparación de otros parámetros estadísticos, ya que la prueba arrojó que no existen diferencias significativas entre las medianas de las topologías puesto que todas las comparaciones arrojaron que el valor-P es mayor que a.

Al revisar los restantes parámetros estadísticos se concluye que la arquitectura 2-8-2 es la que presenta el mejor comportamiento, ya que posee el menor valor de DEECM de las cuatro topologías, y por lo tanto tiene mayor capacidad de generalizar. El coeficiente de correlación de esta topología es el segundo en el orden, si se listaran de forma decreciente, por lo que también se le atribuyó una buena correlación entre las fracciones molares de ambas fases experimentales y las fracciones predichas por la red.

### **Comparación entre las redes neuronales obtenidas**

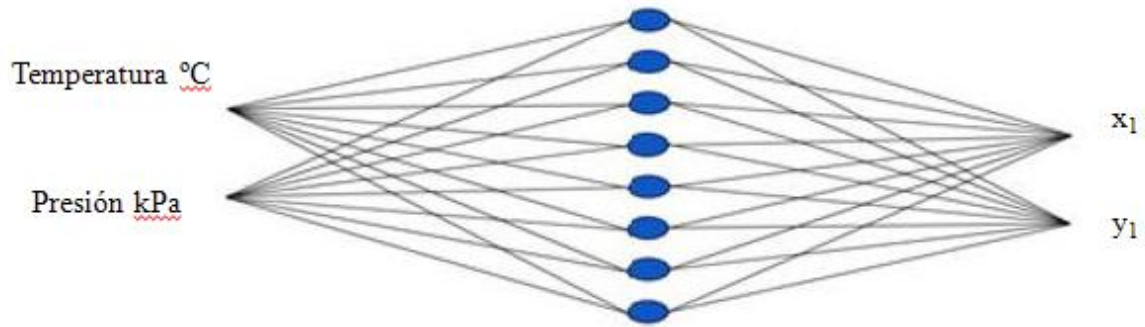
El modelo obtenido en Matlab 2013 tuvo la capacidad de predecir las dos composiciones, pero estadísticamente se comparó con los modelos obtenidos con la otra herramienta. Los resultados de la prueba de Wilcoxon se resumen en la Tabla 15.

**Tabla 15.** Resultados de la prueba de Wilcoxon para las dos herramientas

<b>Pares</b>	<b>Valor-P</b>	<b>Comparación</b>	<b>a</b>
KNIME (x 1 ) - Matlab	0,6229	=	0,05
KNIME (y 1 ) - Matlab	0,0002	=	0,05

Los resultados de la tabla anterior demuestran que no existen diferencias estadísticamente significativas entre las topologías obtenidas para la predicción de la fracción molar en la fase líquida en KNIME y el modelo de Matlab, de donde para la obtención de un valor de fracción molar en la fase líquida de la mezcla etanol-agua se puede emplear cualquiera de las dos topologías. El resultado de la comparación para la fase vapor entre KNIME y Matlab muestra diferencias estadísticas y como Matlab presenta mejor capacidad de generalizar al poseer una DECM (0,0006) menor, mejor relación entre los valores experimentales y los predichos ( $R = 0,9729$ ) y capacidad de predecir las dos fracciones en un solo modelo, es la mejor red neuronal obtenida para este sistema.

En la Figura 2 se presenta la estructura de la red neuronal seleccionada.



**Figura 2.** Topología de la red neuronal seleccionada.

### Comparación con otros modelos de la literatura

En la literatura [24], [25], [26] se encontraron modelos estadísticos de estimación de propiedades de la mezcla binaria etanol-agua a bajas, moderas y elevadas presiones. El uso de los modelos obtenidos está limitado al intervalo de presión para el que se desarrolló cada uno.

La Tabla 16 muestra un resumen del total de modelos que se encontraron en la literatura para diferentes presiones y el resultado obtenido en este trabajo.

**Tabla 16.** Comparación del modelo obtenido con los modelos estadísticos disponibles

Intervalo de presión (kPa)	Modelo	Cantidad	Fuente
6,7 – 98,7	$y_{\text{etanol}} = f(x_{\text{etanol}})$	50	[24]
	$y_{\text{etanol}} = f(T)$	25	
101,3	$y_{\text{etanol}} = f(T)$	2	
203 – 811	$y_{\text{etanol}} = f(x_{\text{etanol}})$	47	[25]
	$y_{\text{etanol}} = f(T)$	25	
912 – 1 520	$y_{\text{etanol}} = f(x_{\text{etanol}})$	57	[26]
	$y_{\text{etanol}} = f(T)$	22	
6,7 – 1 520	$x_{\text{etanol}}; y_{\text{etanol}} = f(T, P)$	1	Este trabajo (Matlab 2013)

La comparación anterior verifica la capacidad de generalización del modelo obtenido, eliminando la limitación de los intervalos de presiones y los 171 modelos estadísticos que se emplearían de utilizar los reportados en la literatura. Otra ventaja del modelo obtenido en este trabajo es su capacidad de predecir simultáneamente las fracciones molares en la fase líquida y la fase vapor en equilibrio.

## CONCLUSIONES

Se seleccionó el modelo obtenido en Matlab 2013 de tres capas, una de entrada, una de salida y una oculta con 8 neuronas, con una arquitectura de perceptrón multicapa, con un algoritmo de aprendizaje de propagación de errores y un algoritmo de entrenamiento de Levenberg-Marquardt, donde las funciones de activación empleadas fueron la tangente sigmoidea y la lineal.

El modelo se validó empleando la validación cruzada obteniendo una media del error cuadrático medio de 0,0054, un coeficiente de correlación entre los datos experimentales y los datos predichos por el modelo de 0,9729 y que posee la mejor capacidad de generalización al tener una desviación estándar del error cuadrático medio de 0,0006.

Se comparó el modelo obtenido con otros reportados por la literatura a partir de ajustes estadísticos y el obtenido en este trabajo posee mayor grado de generalización, pues se obtuvo un modelo que plantea la obtención de dos variables de salida ( $x_1$  ,  $y_1$  ) con dos variables de entrada (T, P), mientras que los reportados permitían calcular una sola variable y tenían intervalos estrechos de validez, lo que dificulta el trabajo.

## BIBLIOGRAFÍA

1. DANIYAN, A., ADEODU, A. O. & DANIYAN, O. L. "Prediction of Vapor-Liquid Equilibrium Data Using Neural Network for Hydrocarbon Ternary System (ethane-propane-n-butane)". *Journal of Computation in Biosciences and Engineering* . 2014, vol.1, Nº1, p. 1-7.
2. KARUNANITHI, B., SHRINIWASAN, S. & BOGESHWARAN.K. "Modeling of Vapor Liquid Equilibrium by Artificial Neural Networks ". *International Journal of Computational Engineering Research*. 2014, vol. 4, Nº6, p. 38-56.
3. LENIN, A., AUGUSTO, C., MANAYAY, D. & CASTILLO, W. "Diseño e implementación de un sistema de control difuso a una columna de destilación". *Ingeniería: Ciencia, Tecnología e Innovación*. 2015, vol. 2, Nº 1, p. 33-48.
4. AGUILAR, F. Puesta en marcha de una técnica de equilibrio líquido vapor isóbaro e investigación termodinámica de equilibrios binarios y ternarios de aditivos oxigenados en hidrocarburos de sustitución de gasolinas sin plomo . Tesis Doctoral, Universidad de Burgos, Burgos, España, 2010
5. GHAEMI, A., SHAHHOSEINI, S., MARAGEH, M. G., & FARROKHI, M. "Prediction of Vapor-Liquid Equilibrium of Aqueous Solutions of Electrolytes Using Artificial neural Network". *Journal of Applied Sciences* . 2008, vol. 8, Nº 4, p. 626-621.
6. VALDERRAMA, J. O. & FAÚNDEZ, C. A. "Modelado del equilibrio líquido-vapor en mezclas binarias y ternarias de interés en destilación vínica". *Información Tecnológica*. 2003, vol. 14 , Nº 1, p. 83-92.
7. RAJESH, R., CHATTOPADHYAY, S. & KUNDU, M. "Prediction of equilibrium solubility of CO<sub>2</sub> in aqueous alkanolamines through artificial neural network". *CHEMECA'06* , September 17, p. 1-7, 2006.
8. MOHANTY, S. "Estimation of vapor liquid equilibria of binary systems, carbon dioxide-ethyl caproate, ethyl caprylate and ethyl caprate using artificial neural networks". *Fluid Phase Equilibria*, 2005, vol. 235, Nº 1, p. 92.
9. ATASHROUZ, S. & MIRSHEKAR, H. "Phase equilibrium modeling for binary systems containing CO<sub>2</sub> using artificial neural networks. *Bulgarian Chemical Communications*", 2014, vol. 46, Nº 1, p. 104-116.

10. QUINTANA, A. "Implementación de un sistema de redes neuronales en matlab para la predicción del equilibrio líquido – vapor de mezclas binarias selectas". Tesis de grado, Universidad de San Buenaventura Seccional, Cartagena, Colombia, 2015.
11. TEIXEIRA, C. L. "Uma aplicação de redes neurais na instrumentação virtual do grau alcoólico no processo industrial de produção do etanol". Tesis de Maestría, Universidade de São Paulo, São Carlos. Brasil, 2014.
12. LV, H.C., & SHEN, Y.-H. "Correlation of Liquid-Liquid Equilibrium of Four Binary Hydrocarbon-Water Systems, Using an Improved Artificial Neural Network Model. Journal of the Korean Chemical Society". 2013, vol. 57, Nº 3, p. 370-376.
13. MOGHADASSI, A. R., PARVIZIAN, F., HOSSEINI, S. M., & FAZLALI, A. R. "A new approach for estimation of PVT properties of pure gases based on artificial neural network model". *Brazilian Journal of Chemical Engineering* . 2009, vol.26, Nº 1, p. 199-206.
14. VALDERRAMA, J. O., REÁTEGUI, A. & ROJAS, R. E. "Density of Ionic Liquids Using Group Contribution and Artificial Neural Networks". *Ind. Eng. Chem. Res* . 2009, vol. 48, Nº 6, p. 3254–3259.
15. ANDRADE, R. S., MAGALHAES, R. & IGLESIAS, M. "Artificial neural network model to predict thermodynamic properties of low molar mass protic ionic liquid". *World Wide Journal of Multidisciplinary Research and Development* . 2016, vol. 2, Nº 2, p. 1-6.
16. SHOJAEI, S. A., ZEINOLABEDINI, A., LASHKARBOLOOKI, M. & SHAFIPOUR, Z. S. (2014). "Prediction of the binary density of the ionic liquids + water using backpropagated feed forward artificial neural network". *Chemical Industry & Chemical Engineering Quarterly*. 2014, vol.20, Nº 3, p. 325-338.
17. JAFARI, D., & MOSTAFA, S. "Studies on the Applicability of Artificial Neural Network (ANN) in Prediction of Thermodynamic Behavior of Sodium Chloride Aqueous System Containing Non-Electrolytes ". *International Journal of Chemical, Molecular, Nuclear, Materials and Metallurgical Engineering* . 2015, vol. 9, Nº 1, p. 110-113.

Recibido: Marzo 2018

Aprobado: Julio 2018

*Dra. Lourdes Zumalacárregui de Cárdenas.* Universidad Tecnológica de La Habana "José Antonio Echeverría" (CUJAE)