

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”

DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.



VII Simposio Internacional de Química 2019

Modelación neuronal de la etapa de fermentación en la destilería Héctor Molina

Neural modeling of fermentation stage in the Hector Molina distillery

Ing. Luis Eduardo López de la Maza¹, Dr. Osney Pérez Ones¹, Dr. Lourdes Zumalacárregui de Cárdenas¹

1- Luis Eduardo López de la Maza. Universidad Tecnológica de La Habana “José Antonio Echeverría” CUJAE, Cuba. E-mail: mazacujae@gmail.com

2- Osney Pérez Ones. Universidad Tecnológica de La Habana “José Antonio Echeverría” CUJAE, Cuba. E-mail: osney@quimica.cujae.edu.cu

3- Lourdes Zumalacárregui de Cárdenas. Universidad Tecnológica de La Habana “José Antonio Echeverría” CUJAE, Cuba. E-mail: lourdes@quimica.cujae.edu.cu

Resumen:

La etapa de fermentación es la más crítica en la fabricación de bebidas alcohólicas pues determina la cantidad de etanol que se va a producir. Las técnicas de análisis multivariante como el análisis de componentes principales (ACP) permiten obtener una mejor comprensión de procesos complejos. La combinación del ACP y las redes neuronales conduce a la modelación robusta de la etapa de fermentación. Esta investigación tuvo como objetivo obtener un modelo para la estimación de la concentración de etanol una vez terminada la etapa de fermentación en función de variables operacionales. Para el estudio se recopilaron 249 instancias de las variables temperatura, concentración de células vivas en el inóculo, pH, sólidos aerométricos disueltos y grado alcohólico. Se utilizó el software Unscrambler v8.0 para el ACP y Matlab 2013 para la creación del modelo de redes neuronales. El primer componente principal explica el 94,9 % de la varianza del conjunto inicial de datos. Se comprobó que el intervalo en que se normalmente se trabajan las condiciones iniciales de las variables temperatura, pH, concentración de células vivas y concentración de sólidos aerométricos disueltos dan

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

lugar a fermentaciones con rendimiento medio y bajo. La mejor topología neuronal presenta cuatro variables de entrada y seis neuronas en la capa oculta con una media del error cuadrático medio de $4,34 \cdot 10^{-4}$ y un factor de correlación con los datos experimentales de 0,916. La correcta operación de los fermentadores a lo largo de la fermentación constituye un factor de altísima importancia en la obtención de altos rendimientos.

Abstract:

The fermentation stage is the most critical in the manufacture of alcoholic beverages because it determines the amount of ethanol that is going to be produced. Multivariate analysis techniques such as principal component analysis (PCA) allow a better understanding of complex processes. The combination of ACP and neural networks leads to robust modeling of the fermentation stage. The objective of this research was to obtain a model for estimating the ethanol concentration once the fermentation stage has been completed based on operational variables. For the study, 249 instances of temperature, concentration of living cells in the inoculum, pH, dissolved aerometric solids and alcoholic strength were collected. The Unscrambler v8.0 software was used for the PCA and Matlab 2013 for the creation of the neural network model. The first principal component explains 94.9% of the variance of the initial data set. Fermentation yield resulting for initial values of temperature, pH, biomass and dissolved aerometric solids may be low or normal. The best neuronal topology has four input variables and six neurons in the hidden layer with a mean square error of $4.34 \cdot 10^{-4}$ and a correlation factor with the experimental data of 0.916. The correct operation of the fermenters throughout the fermentation is a very important factor in obtaining high yields.

Palabras Clave: Fermentación alcohólica; Etanol; Modelo

Keywords: Alcoholic fermentation, Ethanol, Model

1. Introducción

Los conocimientos acerca de la fermentación alcohólica se remontan a las primeras civilizaciones que llegaron a alcanzar un grado de desarrollo considerable. La fermentación alcohólica es un proceso

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**

**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**



llevado a cabo por microorganismos en condiciones anaerobias, en el cual se degradan carbohidratos para obtener etanol y dióxido de carbono como productos principales (Epifanio, 2005).

Para la fabricación de bebidas alcohólicas a escala industrial se han desarrollado dos tecnologías fundamentales de fermentación: la fermentación continua y la fermentación discontinua. En las destilerías cubanas se utiliza la fermentación discontinua y las mieles finales del proceso de fabricación de azúcar como materia prima, pues esta constituye un medio de cultivo idóneo para la obtención de etanol (Menéndez, 2010).

Actualmente existe marcada tendencia a la simulación y optimización de procesos, para contribuir a la disminución de los tiempos de operación y al incremento de los rendimientos en el caso de la fermentación alcohólica. Los modelos matemáticos constituyen valiosas herramientas para lograr este objetivo pues permiten simular diversas condiciones operacionales y diseñar equipamiento para lograr volúmenes de producción deseados (Menéndez, 2010; Lin & Tanaka, 2006).

El proceso de fermentación alcohólica tiene parámetros tanto no lineales como dinámicos dada su naturaleza biológica, por tanto, su modelación resulta difícil y desafiante por la gran cantidad de variables y condiciones de operación que influyen en la misma.

Se han desarrollado modelos de crecimiento de microorganismos que toman en cuenta varios efectos inhibitorios de manera simultánea. Sin embargo, no existe un modelo en el que se tengan en cuenta todos los factores inhibitorios, medioambientales y de nutrientes. Atendiendo al variado número de factores que se pueden considerar, pudiera resultar un modelo tan complejo que no sería viable. Los modelos a escala de reactor, que difieren entre sí por el modo de operación, deben tener en cuenta necesariamente un modelo de crecimiento microbiano, que por demás no describe totalmente las condiciones particulares de una operación que no es netamente de reproducción (Psychogios & Ungar, 1992).

La tecnología existente en la destilería “Héctor Molina” permite que allí se cuente con grandes volúmenes de información acerca de las variables y parámetros del proceso de fermentación alcohólica, pero su procesamiento no se explota con la finalidad de optimizar los rendimientos. Toda esta memoria almacenada constituye una valiosa fuente de información que puede ser útil en la comprensión del presente y en la predicción del futuro. Por otra parte, se conoce que se presentan actualmente bajas

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**

**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**



concentraciones de etanol al finalizar la etapa de fermentación, lo que arroja rendimientos inferiores a los indicadores reportados en la bibliografía.

La minería de datos (MD), proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos; permite realizar tareas de predicción, clasificación, asociación, agrupamiento y correlación apoyándose de técnicas estadísticas como el análisis de componentes principales y computacionales como las redes neuronales artificiales. Actualmente la MD se ha popularizado debido al incremento en la capacidad de cálculo de los ordenadores y su abaratamiento, aparejado al incremento en la capacidad de almacenamiento de datos y en la calidad de los mismos.

El análisis de componentes principales, una de las técnicas de MD, es un método de análisis multivariante que permite la interpretación más simple del problema estudiado, la extracción de la información importante, la identificación de datos discrepantes y la reducción de la matriz de datos.

Las redes neuronales artificiales constituyen una herramienta computacional que imita el funcionamiento del cerebro humano pues tiene la capacidad de aprender patrones o comportamientos de una base de datos. La obtención de modelos predictivos a partir del entrenamiento, que se desarrolla presentándole una matriz de entrada y su correspondiente salida, ha permitido la modelación del proceso de fermentación alcohólica. La capacidad de aprender procesos no lineales y dinámicos, así como constituir modelos de caja negra, son dos de las principales ventajas de las redes neuronales artificiales (Assidjo & Yao, 2006). Estas características han permitido que la aplicación de redes neuronales artificiales se haya extendido incluso a la fermentación alcohólica (Assidjo, 2015; Assidjo, 2009; Esfahanian et. al., 2009). La integración de los resultados del análisis de componentes principales a las redes neuronales resulta prometedora.

Los modelos predictivos obtenidos por técnicas de MD constituyen una alternativa a los modelos matemáticos y a la vez una herramienta para el análisis de la información almacenada de los procesos de fermentación alcohólica en la destilería “Héctor Molina”.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**



**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**

El objetivo de este estudio es obtener un modelo, que teniendo en cuenta los valores iniciales de las variables que mayor información encierren, permita estimar la concentración de etanol una vez terminada la etapa de fermentación, utilizando redes neuronales artificiales.

2. Metodología

La construcción de un modelo neuronal con fines de regresión transcurre por varias etapas. El presente trabajo se divide en 4 etapas.

Etapas 1. Recopilación de datos

Los valores de las variables temperatura (T), concentración de células vivas en el inóculo (X_v), pH, sólidos aerométricos disueltos (⁰Brix) y grado alcohólico (cEt) utilizados en el análisis se extrajeron de las hojas de control en la etapa de fermentación que se registraron en la destilería “Héctor Molina”. Los valores de T, X_v, pH y ⁰Brix corresponden a condiciones iniciales de la etapa de fermentación y la cEt se determina una vez concluida esta etapa. Se pudieron recopilar datos de 249 fermentadores en este intervalo de tiempo.

Durante la fermentación se toman muestras en los fermentadores y se realizan mediciones de las variables pH, X_v y ⁰Brix de acuerdo a las normas cubanas aprobadas (NC 709, 2009; NC 711, 2009; Vázquez et al., 2012). La temperatura se mide mediante un termómetro Alla France de acero inoxidable (± 1 °C) directamente en el fermentador. El valor de cEt solo se mide cuando se ha detenido la fermentación (NC 290, 2007).

Se decidió añadir a la matriz inicial de datos la variable categórica “rendimiento” con los niveles medio y bajo asociados a la cEt y a indicadores productivos de procesos de fermentación alcohólica en destilerías que utilizan miel final como sustrato sin sistema de enfriamiento (Pérez, 2011). La adición de esta variable permite la identificación de condiciones iniciales favorables para alcanzar altos rendimientos en la etapa de fermentación.

Etapas 2. Transformación de los datos

La transformación previa de los datos consistió en el centrado y el escalado de los datos ya que las variables tienen diferentes magnitudes, con el propósito de que todas las columnas tengan media cero (centrado) y la misma varianza (escalado). Para el centrado, el programa utiliza la ecuación 1 y para el

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

escalado, la ecuación 2. Con el centrado la información sobre el origen se pierde, pero las distancias entre los puntos permanecen inalteradas. El autoescalado elimina cualquier peso indeseado debido a la diferencia de magnitud entre las variables o de las muestras dándoles a todas las mismas posibilidades de competir; además garantiza que el ACP se realice con la matriz de correlaciones (Jolliffe & Cadima, 2016).

$$x_{(i,k)c} = x_{i,k} - |\bar{x}_i| \quad (1)$$

$$x_{(i,k)e} = \frac{x_{i,k}}{\sigma} \quad (2)$$

Etapa 3. Análisis de componentes principales

Se utilizó el ACP para el conjunto de variables X_v , pH, T y $^0\text{Brix}$ para la descripción del proceso fermentativo, la reducción de la dimensionalidad y limpieza de la matriz de datos, y la selección de las variables para el modelado. El ACP se realizó con el software Unscrambler v8.0 por sus facilidades en el uso y excelentes capacidades de representación gráfica.

Etapa 4. Creación y entrenamiento de la red neuronal

La red neuronal utilizada para la modelación de la operación de los fermentadores de la destilería Héctor Molina fue perceptrón multicapa con una capa oculta. Esta elección está sustentada sobre la base de que este tipo de redes neuronales es fácil de utilizar y permite la modelación de funciones complejas (Assidjo, 2015); además varios autores la han utilizado para un fin muy similar al de este trabajo (Assidjo, 2009; Esfahanian et. al., 2009; Ahmadian et. al., 2006).

La función de activación empleada para la capa oculta fue la tangente sigmoidea (tansig), mientras que para la capa de salida fue seleccionada la función lineal (purelin).

El número de neuronas de la capa oculta se varió desde 4 hasta 10. Para el entrenamiento de la red se empleó una técnica de propagación hacia atrás, el algoritmo Levenberg-Marquardt. Las variables de entrada a la red fueron: T, X_v , pH y $^0\text{Brix}$, mientras que la variable de salida fue cEt.

Se procesaron 248 pares de datos de entrada/salida. Los datos se dividieron en 5 subgrupos (pliegues) de entrada/salida de forma tal que se pudiera realizar la validación cruzada del modelo de modo que, en

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**



**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**

cada caso, la red se entrenó con el 80% de los datos y se validó con el 20%. El error de validación se chequeó durante el entrenamiento. Inicialmente este error va decreciendo, pero cuando la red comienza a sobreentrenarse este comienza a aumentar. El incremento en 5 iteraciones consecutivas del error de validación y 1 000 como número máximo de ciclo de entrenamiento fueron los criterios utilizados para detener el entrenamiento.

Antes de realizar el entrenamiento los datos se normalizaron en el intervalo de 0 a 1. Tanto la creación de la red neuronal como su entrenamiento se realizaron en Matlab 2013.

Para la selección de la cantidad de neuronas óptima en la capa oculta se tuvieron en cuenta dos criterios: el error cuadrático medio en la validación del modelo y el coeficiente de correlación entre los valores de concentración de etanol estimados por el modelo neuronal y los valores reales.

El error cuadrático se muestra a continuación (Esfahanian et. al., 2009):

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_e - y_c)^2 \quad (3)$$

Donde y_e y y_c son el valor real y el valor estimado de la concentración de etanol para $i = 1, \dots, N$, siendo N el número de muestras con que se trabajó.

Métodos estadísticos

Coeficiente de correlación

El coeficiente de correlación es muy utilizado para evaluar el comportamiento de modelos neuronales. Se obtiene por regresión lineal entre los valores estimados por la red y los valores reales. A continuación, se muestra cómo calcularlo (Assidjo, 2015).

$$R = \frac{\sum_{i=1}^N (y_e - \bar{y}_e)(y_c - \bar{y}_c)}{\sqrt{\sum_{i=1}^N (y_e - \bar{y}_e)^2} \sqrt{\sum_{i=1}^N (y_c - \bar{y}_c)^2}} \quad (4)$$

Prueba de Friedman

La prueba de Friedman, que permite la comparación de varias muestras, se utilizó para la selección de la arquitectura neuronal de mejor comportamiento de acuerdo al criterio del error cuadrático medio, ya

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

que los resultados parecían muy similares y la decisión se dificultaba. Esta prueba tiene dos hipótesis: h_0 (hipótesis nula), que plantea la igualdad entre todas las medianas de un grupo, y h_1 (hipótesis alternativa), que plantea la no igualdad entre todas las medianas del grupo. Es una prueba no paramétrica.

Prueba de Wilcoxon

Cuando la prueba de Friedman revela que existen diferencias estadísticamente significativas entre las medianas del grupo es necesario realizar la prueba de Wilcoxon para seleccionar la arquitectura neuronal de mejor comportamiento. La prueba de Wilcoxon (no paramétrica) permite la comparación de pares de muestras. Tiene dos hipótesis: h_0 (hipótesis nula), que plantea la igualdad entre dos medianas, y h_1 (hipótesis alternativa), que plantea la no igualdad entre dos medianas

3. Resultados y discusión

El ACP se realizó partiendo de un fichero de datos para 249 muestras, con las variables temperatura (T), concentración de iones hidrógeno (pH), concentración de células vivas (Xv) y °Brix (Brix). El gráfico de la influencia se muestra en la figura 1.

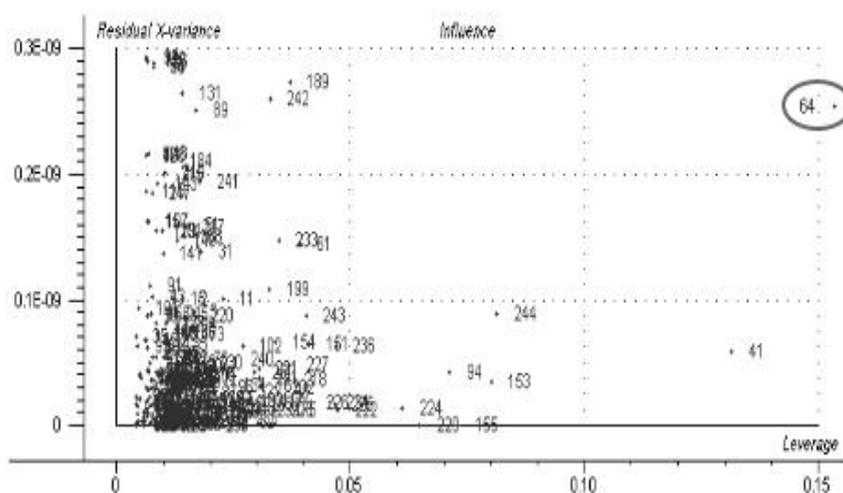


Figura 1. Gráfico de la influencia (Elaboración propia)

La figura 1 permite apreciar que la muestra 64 tiene elevados valores de error residual y distanciamiento y debe por tanto ser eliminada del análisis. En esta muestra la concentración de células vivas (Xv) presenta un valor bajo. La Xv es una variable que se determina mediante análisis de laboratorio y este comportamiento puede estar relacionado con un error en su medición o con el pase del pie de

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

fermentación al fermentador antes de tiempo. Los operadores de la planta cuentan con elevada capacitación y experiencia, y no es usual errores operacionales, por lo que la condición inicial representada por la muestra 64 no es de interés que se refleje en el modelo de ACP. Una vez eliminada la muestra 64 se repitió el ACP.

Dos componentes principales logran explicar el 98,55 % de la varianza de los datos iniciales (ver figura 2). El primer componente explica el 94,92 % de la varianza mientras que el segundo explica 3,63 % de la varianza.

Las barras que representan la varianza explicada de calibración y validación tienen un comportamiento muy parecido por lo que este modelo no solo describe bien los datos de calibración, sino que también describe bien nuevos datos.

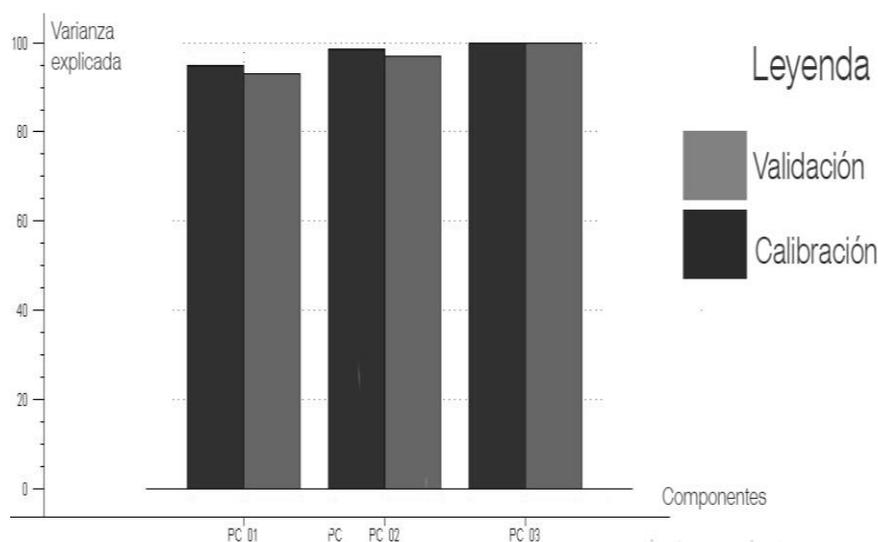


Figura 2. Gráfico de la varianza explicada (Elaboración propia)

Como se puede observar en la figura 3 todas las variables se encuentran entre las dos elipses, muy cercanas a la elipse exterior, por lo que todas son significativas ya que tienen un elevado aporte al primer componente que logra explicar el 94,92 % de la varianza total inicial. Por tanto, las cuatro variables (pH, T, °Brix y Xv) se deben tomar en cuenta en cualquier descripción del proceso que se pretenda realizar.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

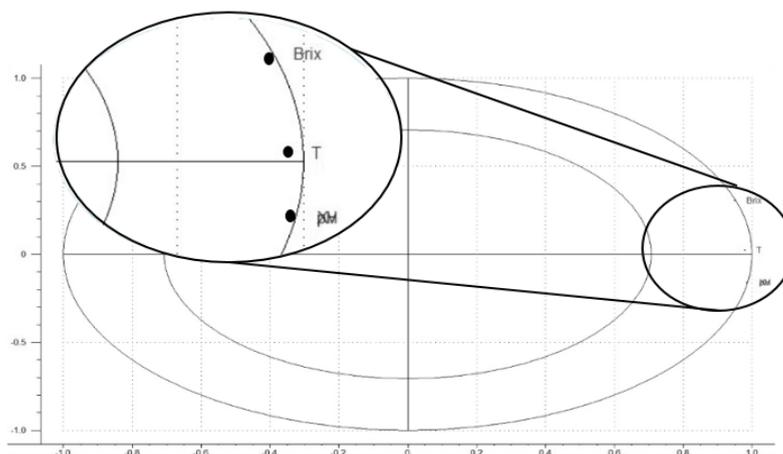


Figura 3. Gráfico de los pesos correlacionados (Elaboración propia)

El hecho de que con dos componentes se explique el 98,55 % de la varianza demuestra que la proyección en dos dimensiones de los datos iniciales (gráfico de las puntuaciones) refleja en gran medida la distribución real de los datos.

En la tabla 1 se encuentran los pesos de cada una de las variables en los dos primeros componentes principales.

Tabla 1. Pesos de las variables en los dos primeros componentes principales

Variables	CP1	CP2
Brix	0,487	0,803
T	0,502	0,065
pH	0,506	-0,419
Xv	0,506	-0,419

El primer componente que permite explicar un 94,92 % de la varianza total puede representarse por la ecuación 5. Todos los pesos del primer componente tienen el mismo signo y por tanto representa un promedio ponderado de todas las variables.

$$CP1 = 0,487Brix + 0,502T + 0,506pH + 0,506Xv \quad (5)$$

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTIFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

La ecuación se pudiera utilizar para la estimación de la concentración de etanol dados los valores iniciales de las variables Brix, T, pH y Xv mediante modelos de regresión simple o técnicas de inteligencia artificial.

Se utilizó el gráfico de las puntuaciones con el objetivo de identificar qué condiciones iniciales de fermentación garantizaban obtener una concentración de etanol al final de dicha etapa acorde a indicadores operacionales propuestos por Pérez (Pérez, 2011). Para destilerías sin sistema de enfriamiento en los fermentadores el grado alcohólico al final de la etapa de fermentación debe estar en el intervalo de 5,5 a 6,5% (v/v). Como puede observarse en la figura 4 se logran rendimientos bajos y en correspondencia con los indicadores para todas las combinaciones de condiciones iniciales.

En este trabajo no se tuvo en cuenta la variación de las variables durante el tiempo. Las variaciones de las variables durante el transcurso del proceso de fermentación reflejan los procedimientos operacionales como el momento de realizar los incrementos de melaza al fermentador y la adición de nitrógeno, factores de gran importancia. Por tanto, se reafirma la necesidad de la correcta actuación de los operadores de la planta para lograr alcanzar concentraciones de etanol acordes a los indicadores productivos para destilerías con estas características.

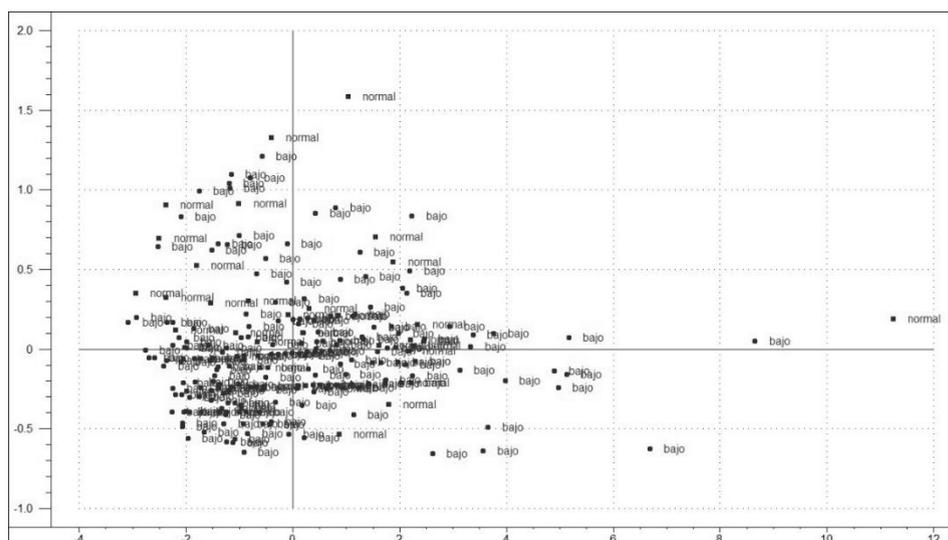


Figura 4. Gráfico de las puntuaciones (Elaboración propia)

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

Topología de la red neuronal

El número de neuronas en la capa oculta se varió de 4 a 10 y se comparó el comportamiento de todas las topologías. El coeficiente de correlación de la concentración de etanol y el error cuadrático medio para cada una de las topologías se muestra en la tabla 2.

Tabla 2. Coeficientes de correlación de concentración de etanol para las diferentes topologías neuronales

Neuronas en la capa oculta	Coeficiente de correlación	Error cuadrático medio
	cEt	
4	0,712	$9,54 \cdot 10^{-4}$
5	0,829	$5,21 \cdot 10^{-4}$
6	0,916	$4,34 \cdot 10^{-4}$
7	0,856	$5,57 \cdot 10^{-4}$
8	0,691	$5,72 \cdot 10^{-4}$
9	0,866	$6,29 \cdot 10^{-4}$
10	0,839	$4,63 \cdot 10^{-4}$

El valor subrayado corresponde al mayor coeficiente de correlación

A partir del análisis de la tabla 2 se muestra que los valores de coeficiente de correlación están en el intervalo de 0,691 a 0,916. También se observa que el mayor coeficiente de correlación se alcanzó para la topología con 6 neuronas en la capa oculta. Atendiendo al criterio del coeficiente de correlación la red neuronal con mejor comportamiento tiene la estructura 4-6-1.

En la tabla 2 se observa la similitud en el orden de error cuadrático medio para cada topología por lo que fue necesario realizar la prueba de Friedman para determinar si existían diferencias estadísticamente significativas.

El resultado del valor-P para la prueba de Friedman fue de 0,00979 ($< 0,05$). Por tanto, existían diferencias significativas entre los comportamientos de las diferentes topologías y debía realizarse la

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**

**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**



prueba de Wilcoxon para definir la mejor topología neuronal atendiendo al criterio del error cuadrático medio.

Solo existen diferencias estadísticamente significativas respecto al error cuadrático medio entre la topología de 6 neuronas en la capa oculta y la de 4 neuronas en la capa oculta. Por tanto, pudiera considerarse como buena cualquier topología excepto la de 4 neuronas en la capa oculta teniendo en cuenta el criterio del error cuadrático medio.

Tomando como bases los resultados del mejor modelo de acuerdo con la topología de la red neuronal, según los criterios del error cuadrático medio y el coeficiente de correlación por separado, se decidió que el modelo neuronal que mejor describe la etapa de fermentación de la destilería Héctor Molina es el 4-6-1. El mejor coeficiente de correlación entre todas las topologías, una moderada complejidad estructural que permite un ahorro en cálculos al software Matlab y bajos valores de error cuadrático medio, son razones que justifican la decisión anterior.

Simulación

La topología 4-6-1 se utilizó para simular 16 condiciones iniciales de fermentaciones que no fueron empleadas para el entrenamiento y la validación. La calidad del modelo se puede observar en la figura 5, donde se muestran los valores reales y los estimados por el modelo neuronal para las diferentes condiciones iniciales. Se puede observar que los valores estimados están muy cerca de los valores reales.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCION CIENTIFICA INTERNACIONAL
“II CCI UCLV 2019”

DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

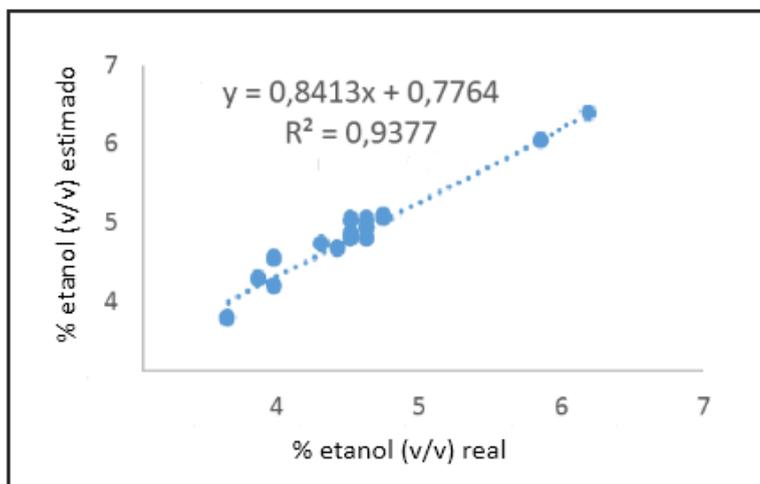


Figura 5. Simulación para condiciones iniciales diferentes (Elaboración propia)

4. Conclusiones

La aplicación del análisis de componentes principales al proceso fermentativo permitió evidenciar la necesidad de tener en cuenta los valores de las variables T, pH, Brix y Xv para la descripción de las condiciones iniciales de fermentación en la destilería “Héctor Molina”.

El primer componente principal, que explica el 94,92 % de la varianza total se puede utilizar en la construcción de un modelo de estimación de la concentración de etanol en función de las variables T, pH, Brix y Xv.

La correcta operación de la etapa de fermentación (tiempo de realización de los incrementos de melaza en el fermentador y la adición de nitrógeno) constituye un factor esencial para alcanzar rendimientos acordes a las características de la destilería ya que las condiciones iniciales en las que se mueven las variables Xv, pH, T y ⁰Brix pueden resultar en cEt en el intervalo de 5,5 a 6,5% (v/v) e inferiores.

La red neuronal del tipo perceptrón multicapa con topología 4-6-1 se utilizó para la modelación de fermentadores en la destilería Héctor Molina y demostró su capacidad de estimar satisfactoriamente la concentración de etanol al final de la etapa de fermentación. El alto valor del coeficiente de determinación (0,9377) entre los valores simulados y los valores reales, y los bajos valores de error

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”



DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.

cuadrático medio en la validación del modelo indican la conveniencia de la utilización de redes neuronales en la modelación de la fermentación alcohólica.

5. Referencias bibliográficas

1. Ahmadian-Moghadam, H., Elegado, F., & Nayve, R. (2013). Prediction of ethanol concentration in biofuel production using artificial neural networks. *American Journal of Modeling and Optimization*, 1(3), 31-35.
2. Assidjo, E., & Yao, B. (2006). Industrial Brewery Modeling by Using Artificial Neural Network. *Journal of Applied Sciences*, 6(8), 1858-2006.
3. Assidjo, E. (2009). A hybrid neural network approach for batch fermentation simulation. *Australian Journal of basic and applied sciences*, 3(4), 3930-3936.
4. Assidjo, E. (2015). Batch fermentation process of sorghum wort modeling by artificial neural network. *European Scientific Journal*, ESJ, 11(3).
5. Epifanio, S. (2005). *Influencia de la tecnología de vinificación en la microbiología y el desarrollo de la fermentación alcohólica*. Tesis Doctoral. Universidad de la Rioja, España.
6. Esfahanian, M., Nikzad, M., Najafpour, G., & Ghoreyshi, A. A. (2013). Modeling and optimization of ethanol fermentation using *Saccharomyces cerevisiae*: Response surface methodology and artificial neural network. *Chemical Industry and Chemical Engineering Quarterly/CICEQ*, 19(2), 241-252.
7. Jolliffe, I., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions*. doi: <http://dx.doi.org/10.1098/rsta.2015.0202>
8. Lin, Y., & Tanaka, S. (2006). Ethanol fermentation from biomass resources: current state and prospects. *Applied microbiology and biotechnology*, 69(6), 627-642.
9. Menéndez, Z. (2010). *Desarrollo de módulos de cálculo para los procesos de fermentación alcohólica*. (Tesis de maestría en Análisis y Control de Procesos), Universidad Tecnológica de La Habana “José Antonio Echeverría” CUJAE.
10. NC 290: 2007: Norma Cubana Bebidas alcohólicas- Determinación del grado alcohólico en alcoholes, bebidas alcohólicas destiladas, vinos, licores, bebidas alcohólicas preparadas, cocteles y extractos hidroalcohólicos. Oficina Nacional de Normalización. 2007.

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu

**PLANTILLA OFICIAL PARA LA PRESENTACIÓN DE TRABAJOS
II CONVENCIÓN CIENTÍFICA INTERNACIONAL
“II CCI UCLV 2019”**

**DEL 23 AL 30 DE JUNIO DEL 2019.
CAYOS DE VILLA CLARA. CUBA.**



11. NC 709: 2009: Norma Cubana MIEL FINAL — Determinación de sólidos aerométricos disueltos. Oficina Nacional de Normalización. 2009.
12. NC 711: 2009: Determinación potenciométrica del pH. Oficina Nacional de Normalización. 2009.
13. Pérez, O. (2011). *Modelación, simulación y análisis con fines energéticos de destilerías de etanol hidratado*. (Tesis de doctorado en Ingeniería Química), Universidad Tecnológica de La Habana "José Antonio Echeverría" CUJAE.
14. Psychogios, D. C., & Ungar, L. H. (1992). A hybrid neural network-first principles approach to process modeling. *AChE Journal*, 38(10), 1499-1511.
15. Vázquez, M., Cuellar, H., García, R., Díaz, M. E., Reyes, A., González, M. D., Ibáñez, M.-L. (2012). Técnicas microbiológicas *Manual de técnicas analíticas para destilerías*. Editorial ICIDCA. (pp. 78-85).

Información de contacto
convencionuclv@uclv.cu
www.uclv.edu.cu