

Sensor Virtual Adaptable de Concentración de Etanol para Fermentadores Industriales

Boris Martínez * Francisco Herrera * Luis Peralta **

* Dpto. de Automática y Sistemas Computacionales

** Dpto. de Ingeniería Química

Universidad Central de Las Villas (UCLV)

Carretera a Camajuaní Km. 5,5; Santa Clara, Cuba.

e-mail: (boris, herrera, lperalta)@uclv.edu.cu

Resumen: Los sistemas de control emplean sensores para observar el estado del proceso y tomar decisiones. En ocasiones, se necesita estimar las variables del proceso pues el sensor adecuado no existe, es prohibitivamente costoso o las mediciones son difíciles de realizar. Una solución consiste en inferir las variables no medidas a partir de otras variables mediante sensores virtuales o sensores por software (*soft-sensors*). En los procesos de fermentación alcohólica, la medición de la concentración del etanol es esencial. Sin embargo, no existen sensores baratos y confiables para medirla en línea ni existe una solución aceptada por todos del modelado de dicha variable. Además, las fermentaciones nunca son iguales pues los microorganismos son muy sensibles a pequeñas desviaciones en las variables involucradas. Por tanto, estos procesos requieren un sistema de estimación adaptable y altamente robusto. En este trabajo se presenta un sensor virtual adaptable para un proceso fermentativo de bioetanol empleando un modelo borroso evolutivo a partir de datos del proceso. Además, el modelo obtenido es compacto y presenta una estructura adecuada para su aplicación futura en estrategias de control, en aras de optimizar la productividad del proceso y disminuir los costos de producción. Copyright © 2009 CEA.

Palabras Clave: bioetanol, procesos fermentativos, sensores virtuales o sensores software, sistemas adaptables, sistemas borrosos.

1. INTRODUCCIÓN

La industria biotecnológica ha tenido un desarrollo creciente y un alto impacto en la sociedad actual. Por otro lado, también ha generado importantes retos a los ingenieros y tecnólogos dada la complejidad de los procesos biotecnológicos. Estos procesos se ven afectados por infinidad de variables que tienen una influencia directa en el metabolismo de las células, tales como: temperatura, grado de acidez (pH), oxígeno disuelto, concentración de sustrato, velocidad de agitación, etc. Para optimizar la producción se requiere controlar las variables importantes, muchas de las cuales pueden ser medidas en línea, por ejemplo: temperatura, pH, velocidad de agitación, entre otras. Sin embargo, variables influyentes como las concentraciones de biomasa y de producto, raramente son medidas en línea, ya sea porque no existen dispositivos adecuados o estos son muy caros.

En contraste con el ambiente académico, las políticas de control de bioprocesos industriales se basan mayoritariamente en el empleo de mediciones del proceso fuera de línea. Como estos análisis de laboratorio son costosos, se reduce al máximo posible su frecuencia, lo cual conlleva a una pobre regulación del proceso ante cambios en el mismo o ante perturbaciones.

Dado que el objetivo fundamental de una fermentación es la producción de biomasa y/o producto, es necesario medir en línea las concentraciones de estas variables para conocer y controlar el desempeño del proceso. Esta necesidad viene dada porque si fuese inadecuado el comportamiento de la fermentación, la acción correctiva pudiera resultar tardía y desastro-

sa utilizando los métodos tradicionales de medición fuera de línea (Reynaga and Leal, 2000). La poca rentabilidad de los equipos adecuados y la ausencia de métodos apropiados para la obtención de lecturas de todas las variables del proceso en línea son fuertes obstáculos en la tarea de conducir adecuadamente una fermentación (Franco, 2002; Muñoz and Angulo, 2006; Sundstrom and Enfors, 2008).

A esta situación se agrega que las fermentaciones nunca son iguales, aun con los más estrictos controles de condiciones iniciales, sustratos e instrumentación, pues el metabolismo de los microorganismos es muy sensible al conjunto de pequeñas variaciones en las distintas variables involucradas en el proceso. Por tanto, estos procesos requieren de un sistema de estimación altamente robusto y adaptable que brinde soluciones adecuadas a estos problemas.

La determinación exacta, continua y en línea de las concentraciones de biomasa, sustrato y producto es uno de los principales sueños en la biotecnología (Reynaga and Leal, 2000; Franco, 2002). En muchos casos, la información sobre estas variables existe pero en forma de relaciones desconocidas, complejas y altamente no lineales. Una posibilidad de dar cumplimiento a este sueño consiste en inferir estas concentraciones a partir de variables medidas empleando los llamados 'sensores virtuales' o 'sensores por software' (*soft-sensors*).

Los sensores virtuales, también conocidos como 'observadores' o 'estimadores de estado' en la ingeniería de control, surgen basados en la idea de la inferencia. Dichos sistemas se centran en el proceso de estimación de alguna variable del proceso

o cualidad de un producto empleando modelos matemáticos que sustituyen a algunos sensores físicos y que emplean datos adquiridos por otros sensores disponibles, cuando no se tiene un modelo adecuado o cuando tal modelo es casi imposible de formular (Fortuna *et al.*, 2007).

Los sensores virtuales han encontrado un amplio campo de acción en la ingeniería química. En algunos procesos, como el caso de reactores bioquímicos, su uso no es sólo cuestión de economía sino también de factibilidad, ya que resulta prácticamente imposible contar con sensores que midan en línea todas las variables del proceso (Salazar, 2008).

Un prerrequisito para cualquier técnica de estimación basada en modelo es el desarrollo de un modelo matemático balanceado del sistema bajo estudio. El modelo debe tener un balance entre la complejidad dinámica y la habilidad para obtener parámetros a partir de datos. Varias técnicas para la estimación de variables importantes de los procesos biotecnológicos han sido sugeridas para el desarrollo de observadores en línea así como el de controladores del bioproceso (Bastin and Dochain, 1990; Babuska, 1998; Roubos *et al.*, 2000; Reynaga and Leal, 2000; Franco, 2002; Muñoz and Angulo, 2006; Sundstrom and Enfors, 2008). Estas se basan en combinar la estructura del modelo obtenido con ecuaciones de balance con la estimación de parámetros de dicho modelo a partir de datos, combinación conocida como modelado de caja gris o modelado híbrido. Sin embargo, no todas estas técnicas han sido aplicables industrialmente ni han sido ampliamente aceptadas en la comunicad industrial

Por otro lado, el modelado borroso ofrece una forma muy estructurada de formulación del conocimiento, capaz de proporcionarlo a partir de datos y de incorporarlo a partir de expertos. Otra propiedad importante de los modelos borrosos es su capacidad de representar sistemas no lineales continuos (Babuska, 1998).

En los últimos años, se ha visto un gran impulso a los métodos de construcción de modelos borrosos a partir de datos. Muchas de estos se caracterizan por emplear técnicas de agrupamiento para dividir los datos analizados en subconjuntos y por construir modelos del tipo Takagi-Sugeno (TS) (Takagi and Sugeno, 1985) a partir de dichas particiones, lo cual ha sido establecido como una de las tendencias actuales en la literatura de identificación y control inteligente (Babuska, 1998; Angelov and Filev, 2004; Díez *et al.*, 2004; Lughofer and Klement, 2005; Díez *et al.*, 2007; Bouchachia and Mittermeir, 2007; Martínez, 2007). Dicha técnica se ha convertido en una herramienta práctica para el modelado de sistemas complejos debido a sus buenos resultados, a su sencillez matemática y a que puede describir sistemas altamente no lineales utilizando pocas reglas.

Como resultado del desarrollo de algoritmos en línea prácticos que trabajan en tiempo real, recientemente han surgido los llamados sistemas inteligentes evolutivos (*evolving Intelligent Systems-eIS*), y dentro de esta clasificación, los sistemas borrosos evolutivos (*evolving Fuzzy Systems-eFS*), (Kasabov and Song, 2002; Angelov and Filev, 2004; Lughofer and Klement, 2005; Bouchachia and Mittermeir, 2007; Martínez *et al.*, 2008). Estos sistemas poseen una adaptación de alto nivel que se refiere tanto a la estructura del mismo (p. ej., la base de reglas) como a sus parámetros, lo cual forma la base conceptual para el desarrollo de sistemas verdaderamente inteligentes. Si se compara con los algoritmos evolutivos convencionales (p. ej., los algoritmos genéticos) que imitan el proceso evolutivo de una población de individuos (filogénesis), este nuevo paradigma

evolutivo imita el desarrollo de los individuos en la naturaleza (ontogénesis) (Angelov and Kasabov, 2006).

En este trabajo se analiza la obtención de un sensor virtual basado en un modelo borroso adaptable para la estimación en línea de la concentración de etanol en un proceso fermentativo. Para ello se utilizan sistemas borrosos Takagi-Sugeno evolutivos, siguiendo el procedimiento expuesto en Martínez (2007).

A continuación, el trabajo se organiza de la siguiente forma. En la sección 2 se presentan las características del proceso de fermentación alcohólica bajo estudio. La sección 3 describe la metodología utilizada para la obtención de modelos borrosos Takagi-Sugeno evolutivos. En la sección 4 se analiza el sensor virtual obtenido y, finalmente, en la sección 5 se presentan las conclusiones.

2. FERMENTACIÓN ALCOHÓLICA

Antiguamente, la mayor parte del etanol se consumía en forma de bebidas alcohólicas, pero en la actualidad esta sustancia ha adquirido una extraordinaria importancia comercial e industrial. La fabricación de alcohol como combustible alternativo para sortear los altos precios del petróleo, ha despertado recientemente un gran interés, ya que este es más barato y compatible con el medio ambiente, unido a los excelentes resultados obtenidos en varios países con Brasil a la cabeza. Además, son numerosas las otras aplicaciones en las que se involucra el etanol o alcohol etílico. Entre ellas se destacan la utilización como desinfectante, disolvente y como materia prima en la industria química para la obtención de una gran variedad de compuestos como acetaldehído, ácido acético, vinagre, clorofórmio, éter, etileno, bromuro y yoduro de etilo, acetato de etilo, medicamentos, perfumes, etc.

Por otro lado, el desarrollo ambientalmente compatible de la industria azucarera como productora no sólo de azúcar, sino también de materias primas y portadores energéticos como bioetanol y electricidad, es un imperativo de estos tiempos. La posibilidad de lograr, en determinadas instalaciones de la industria azucarera, niveles de generación de etanol que satisfagan varias necesidades, tiene una relevancia económica estratégica. El bioetanol, además de ser una apuesta fuerte para el futuro de los combustibles, se puede considerar un tema económico, científico, medio ambiental y social, es decir, una primera plana en el conocimiento actual.

Actualmente se conoce una gran variedad de especies capaces de producir bioetanol, principalmente levaduras y bacterias. Entre las levaduras ha adquirido gran importancia la especie *Saccharomyces cerevisiae* y entre las bacterias se destaca especialmente *Zymomonas mobilis* por sus altos rendimientos.

La etapa fundamental de producción de etanol es la fermentación. En Cuba, el alcohol etílico para las licorerías y el uso industrial (etanol 95 %) se produce por fermentación de levaduras. Una vez terminada la fermentación se separan las levaduras del sustrato fermentado, el alcohol se destila y se somete a un proceso de rectificación (Fabelo, 1999).

2.1 Dinámica del proceso

Un *biorreactor* o *fermentador* es un tanque acondicionado para que se produzcan simultáneamente varias reacciones bioquímicas mientras ocurre el crecimiento de la masa de microorganismos o células (*biomasa*), dado por el consumo de nutrientes

apropiados (*sustratos*) en condiciones ambientales favorables, mientras se obtienen nuevas sustancias (*productos*). Las tres formas básicas de operación de un fermentador son: la fermentación discontinua o en lotes (*batch*), la fermentación en lotes alimentada (*fed-batch*) y la fermentación continua. En la industria, la fermentación más común es la fermentación en lotes alimentada (*fed-batch*).

En los cultivos microbianos, el crecimiento de los microorganismos generalmente se formula en términos de rendimiento de crecimiento, velocidad específica de crecimiento y concentración del sustrato (Franco, 2002; Berovic and Kieran, 2006). La teoría matemática empleada fue tomada por analogía con los procesos químicos, pero estas ecuaciones no abarcan a los procesos biológicos en su totalidad. Las diferencias parten de que en las reacciones químicas los iones o moléculas se consideran similares, en cambio, en los sistemas biológicos las células tienen distintas 'edades': mientras unas están 'naciendo', otras están en 'máxima actividad' o 'muriendo' (Garro, 1993).

Empleando esta teoría se han propuesto muchos modelos matemáticos para describir la producción de etanol. La dinámica de los biorreactores generalmente se describe por modelos cinéticos en los cuales las variables de estado son las concentraciones de biomasa (X), de sustrato (S) y del producto (P), y el modelado de la velocidad específica de crecimiento de la biomasa y de la velocidad de producción de etanol son los pasos más importantes. A pesar de que se han obtenido varios modelos, aún no existe una solución aceptada por todos (Berovic and Kieran, 2006).

En el presente trabajo se analiza el proceso fermentativo de producción de etanol que ocurre en la destilería anexada al Complejo Agro-Industrial George Washington de la provincia de Villa Clara, Cuba. La dinámica del proceso bajo estudio está dada por las siguientes ecuaciones de balance de masa:

$$\frac{dX}{dt} = \mu X - KmX - \frac{Fin}{V}X, \quad (1)$$

$$\frac{dS}{dt} = -\frac{\mu}{Y_{XS}}X + \frac{Fin}{V}(S_{in} - S), \quad (2)$$

$$\frac{dP}{dt} = \frac{\mu}{Y_{XP}}X - \frac{Fin}{V}P, \quad (3)$$

$$\frac{dV}{dt} = Fin, \quad (4)$$

donde X , S y P son las concentraciones de biomasa, de sustrato y de etanol (kg/m^3); V es el volumen del medio de cultivo (m^3); Fin es el flujo de alimentación (m^3/h); S_{in} es la concentración de sustrato en el flujo de entrada (kg/m^3); μ es la velocidad específica de crecimiento ($1/h$); Km es un parámetro de mortandad ($1/h$); Y_{XS} es el rendimiento celular y Y_{XP} es la productividad de etanol. Estos últimos parámetros son de difícil obtención y cambiantes con cada fermentación (Peralta *et al.*, 2000).

2.2 Procedimiento industrial

La etapa fermentativa de la destilería anexa al C.A.I. George Washington cuenta con 12 fermentadores de $100 m^3$ de capacidad, 3 prefermentadores de $10 m^3$ y utiliza la levadura *Saccharomyces cerevisiae* de uso industrial. El procedimiento de trabajo de cada fermentador es sencillo. Primeramente se añade miel diluida (16–20 °Bx) hasta, aproximadamente, la cuarta parte de su volumen total. Simultáneamente se adiciona

un prefermento obtenido con anterioridad en los prefermentadores; de tal forma se alcanza, aproximadamente, un tercio del volumen del tanque. Luego, se deja en reposo y al término de media hora o una hora, se observa que la fermentación ha comenzado y adquiere vigor.

En tales condiciones se procede a llenar el fermentador con miel diluida. Este llenado se puede hacer de tres maneras: de modo lento y continuo, en etapas o refrescos, o de modo rápido y continuo. En las destilerías cubanas se utiliza, preferentemente, el sistema de refrescos por su sencillez e independencia de factores mecánicos como capacidad de bombas, operación del disolutor de miel, etc. Sin embargo, este procedimiento es altamente dependiente de la experiencia y habilidades del operador.

La marcha de la fermentación se controla por medición del Brix cada una hora. Este desciende progresivamente mientras se agota el azúcar fermentable existente en el biorreactor. En la práctica, una vez llenado el fermentador, cuando en dos mediciones horarias se repite el mismo valor del Brix, se da por terminada la fermentación. El ciclo de la fermentación fluctúa entre 24 y 32 horas, dependiendo de la cantidad y actividad del prefermento, la cantidad de azúcar fermentable existente y ciertas características de la miel que influyen sobre la velocidad y eficiencia de la fermentación (Peralta *et al.*, 2000).

El estudio del proceso se realizó tomando los datos de ocho experimentos realizados en fermentadores industriales en operación en lotes alimentada, específicamente, en fermentación discontinua con refrescos. Estos refrescos son añadidos por el operador cuando se considera que la temperatura del fermentador puede ser perjudicial para la producción de etanol (aproximadamente entre 30 y 34 °C). Durante las fermentaciones, cada una hora se realizaron mediciones del pH, la temperatura, el Brix y el porcentaje alcohólico, obteniéndose un total de 240 valores de cada variable. Los valores de las dos últimas variables se realizaron en el laboratorio.

3. MODELADO BORROSO EVOLUTIVO

3.1 Modelo borroso Takagi-Sugeno (TS)

El sistema borroso utilizado es el conocido sistema de inferencia Takagi-Sugeno (Takagi and Sugeno, 1985). Dicho sistema está compuesto por K reglas borrosas de la siguiente forma:

$$\begin{aligned} \mathfrak{R}_i : & \text{if } x_1 \text{ is } A_{i1} \text{ and } \dots \text{ and } x_r \text{ is } A_{ir} \\ & \text{then } y_i = a_{i0} + a_{i1}x_1 + \dots + a_{ir}x_r, \end{aligned} \quad (5)$$

donde x_j , $j = 1, \dots, r$, son las variables de entrada definidas en los universos de discurso X_j y A_{ij} son los conjuntos borrosos definidos por sus funciones de pertenencia $\mu_{A_{ij}} : X_j \rightarrow [0, 1]$. En la parte del consecuente, y_i es la salida de la regla i y a_{ij} son escalares.

Para un vector de entrada $\mathbf{x} = [x_1, x_2, \dots, x_r]^T$, cada función del consecuente puede expresarse como:

$$y_i = \mathbf{a}_i^T \mathbf{x}_e, \quad (6)$$

donde $\mathbf{x}_e = [1, \mathbf{x}^T]^T$ y $\mathbf{a}_i = [a_{i0}, a_{i1}, \dots, a_{ir}]^T$.

El resultado de la inferencia, la salida del sistema y , es el promedio ponderado de la salida de cada regla y_i , como se indica a continuación:

$$y = \frac{\sum_{i=1}^K w_i y_i}{\sum_{i=1}^K w_i} = \frac{\sum_{i=1}^K w_i \mathbf{a}_i^T \mathbf{x}_e}{\sum_{i=1}^K w_i}, \quad (7)$$

donde

$$w_i = \prod_{j=1}^r \mu_{A_{ij}}(x_j) \quad (8)$$

es el nivel de impacto de la regla i .

La ecuación 7 puede reescribirse como:

$$y = \sum_{i=1}^K \tau_i y_i = \sum_{i=1}^K \tau_i \mathbf{a}_i^T \mathbf{x}_e, \quad (9)$$

donde

$$\tau_i = \frac{w_i}{\sum_{i=1}^K w_i} \quad (10)$$

es el nivel de impacto normalizado de la regla i .

Finalmente, todas las funciones de pertenencia utilizadas son del tipo *gaussianas*:

$$\mu(x_d, c_d, \sigma) = \exp\left(-\left(\frac{x_d - c_d}{2\sigma}\right)^2\right), \quad (11)$$

donde c_d es el valor del centro de la clase en la dimensión x_d , y σ es proporcional al radio de atracción de dicha clase.

3.2 Algoritmo de aprendizaje

El algoritmo empleado para la generación de modelos TS a partir de datos consta de dos pasos:

1. *Determinación de la estructura de la base de reglas.* En este paso se divide el espacio de los datos utilizando el método de agrupamiento en línea MAEA (Martínez *et al.*, 2008), el cual es simple, rápido y utiliza dos parámetros definidos por el usuario: un valor de radio umbral de las clases, $Rthr$, y un valor de similitud umbral entre clases, $Sthr$. Una vez halladas las clases, estas se proyectan para obtener el antecedente de cada regla.
2. *Determinación de los parámetros del consecuente de cada regla.* Manteniendo el antecedente de las reglas fijo, los parámetros del consecuente se determinan utilizando algoritmos de mínimos cuadrados.

En la identificación en línea, continuamente arriban nuevos datos y las clases deberán cambiar de acuerdo a estos nuevos datos. Si el dato no pertenece a una clase existente, se crea una nueva clase. Si el nuevo dato está muy cercano a una clase existente, dicha clase es actualizada. Para la determinación de cual de estas acciones realizar se emplea el radio umbral ($Rthr$). Si durante la ejecución de este proceso, dos clases de acercan mucho, estas clases redundantes se unen. Para ello se emplea el parámetro similitud umbral ($Sthr$).

Una nueva regla es generada solo si hay información significativa presente en los datos. Para esto se considera el peso de la clase (W_i), el cual está dado por la cantidad de datos que pertenecen a la misma. El procedimiento verifica si el valor del peso es apreciable ($W_i > \epsilon$) antes de generar una nueva regla. Una vez realizada la identificación de la estructura, los resultados del agrupamiento en línea son utilizados para ajustar los centros y anchos de las funciones de pertenencia según (11).

Para crear y actualizar las funciones lineales de los consecuentes se utilizan estimadores de mínimos cuadrados. Para ello se construye una matriz de composición \mathbf{X}' a partir de los niveles de impacto normalizados τ_i y de los m datos iniciales del proceso (Martínez *et al.*, 2008):

$$\mathbf{X}' = [(\mathbf{T}_1 \mathbf{X}_e), (\mathbf{T}_2 \mathbf{X}_e), \dots, (\mathbf{T}_K \mathbf{X}_e)] , \quad (12)$$

donde la matriz $\mathbf{X}_e = [\mathbf{1}, \mathbf{X}]$ está conformada por las filas $\mathbf{x}_{ek}^T = [1, \mathbf{x}_k^T]$ y el k -ésimo elemento de cada matriz diagonal \mathbf{T}_i ($i = 1, \dots, K$) es el valor de τ_i obtenido de (10).

La expresión

$$\mathbf{a} = (\mathbf{X}'^T \mathbf{X}')^{-1} \mathbf{X}'^T \mathbf{Y}, \quad (13)$$

es utilizada para obtener la matriz inicial de los parámetros del consecuente $\mathbf{a} = [\mathbf{a}_1^T \mathbf{a}_2^T \dots \mathbf{a}_K^T]^T$.

Nótese que (13) se puede reescribir como:

$$\mathbf{P} = (\mathbf{X}'^T \mathbf{X}')^{-1}, \quad (14)$$

$$\mathbf{a} = \mathbf{P} \mathbf{X}'^T \mathbf{Y}.$$

Para la actualización de la matriz de parámetros \mathbf{a} se utiliza un estimador de mínimos cuadrados recursivos (RLS) con un factor de olvido. Sea \mathbf{x}'_k el k -ésimo vector fila de la matriz \mathbf{X}' y \mathbf{y}_k el k -ésimo elemento de \mathbf{Y} , entonces \mathbf{a} puede calcularse iterativamente por:

$$\mathbf{P}_{k+1} = \frac{1}{\lambda} \left(\mathbf{P}_k - \frac{\mathbf{P}_k \mathbf{x}'_{k+1} \mathbf{x}'_{k+1}^T \mathbf{P}_k}{\lambda + \mathbf{x}'_{k+1}^T \mathbf{P}_k \mathbf{x}'_{k+1}} \right), \quad (15)$$

$$\mathbf{a}_{k+1} = \mathbf{a}_k + \mathbf{P}_{k+1} \mathbf{x}'_{k+1} (\mathbf{y}_{k+1} - \mathbf{x}'_{k+1}^T \mathbf{a}_k),$$

donde λ es un factor de olvido constante, con valores típicos entre 0.8 y 1, utilizado para posibilitar un mejor seguimiento a sistemas variantes en el tiempo. Los valores iniciales, $\mathbf{P}(0)$ y $\mathbf{a}(0)$, son calculados utilizando (14).

El procedimiento recursivo para la actualización en línea de modelos borrosos TS evolutivos utilizado en este trabajo se expresa en los siguientes pasos (Martínez *et al.*, 2008):

Algoritmo de aprendizaje.

1. Inicializar el modelo borroso. Para ello:
 - (a) Tomar los primeros m datos del conjunto de datos.
 - (b) Aplicar el método de agrupamiento MAEA.
 - (c) Obtener una regla borrosa por cada clase con información significativa, creando el antecedente con (11) e utilizando (14) para obtener los valores iniciales de \mathbf{P} y \mathbf{a} .
2. Tomar el próximo dato.
3. Actualizar recursivamente los centros de clase utilizando el algoritmo MAEA.
4. Posible modificación de la base de reglas. Una nueva regla borrosa es creada si hay información significativa presente en la nueva clase creada por MAEA, para ello:

SI el peso de la clase no es despreciable ($W_i > \epsilon$),
ENTONCES una nueva regla borrosa es creada.
5. Actualizar los antecedentes mediante (11)
6. Calcular recursivamente los parámetros de los consecuentes mediante (15)
7. Estimar la salida mediante el modelo borroso TS 2.

La ejecución del algoritmo continúa en el próximo instante de tiempo desde el paso 2.

4. SENSOR VIRTUAL ADAPTABLE

En Peralta *et al* (2000) se presenta un modelo analítico para el proceso fermentativo de producción de etanol (1–4) y se

obtienen los parámetros optimizados para cada fermentación. Sin embargo, los resultados de la estimación correspondientes a la concentración de etanol no son adecuados, como se muestra más adelante en la Tabla 1. Una variante entonces sería determinar los parámetros de un modelo parcial del proceso fuera de línea y emplear este modelo en unión de estimadores en línea de la concentración de etanol, resultando en un observador del proceso parcialmente adaptable. Con el propósito de obtener un sensor virtual para la estimación del porcentaje alcohólico (Pp), por ser esta una variable de gran interés en el proceso, se construye un modelo borroso TS evolutivo empleando el algoritmo expuesto en la sección anterior (Epígrafe 3.2).

Considerando las condiciones reales en el proceso, los estudios previos de esta fermentación alcohólica (Fabelo, 1999; Peralta *et al.*, 2000; Sotolongo, 2001) y el análisis de cada variable, se obtiene que las variables más influyentes son el Brix de la batición (Bx), la temperatura (T) y el grado de acidez (pH). Por tanto:

$$\begin{aligned} y &= \{Pp_k\}, \\ \mathbf{x} &= \{Bx_k, pH_k, T_k\}, \end{aligned} \quad (16)$$

donde y representa la salida y \mathbf{x} representa el vector de entrada del estimador. Se hace notar que Pp es proporcional a P y que Bx lo es a S . El modelo de estimación del proceso fermentativo de producción de etanol queda entonces:

$$\begin{aligned} \mathcal{R}_i : & \text{if } Bx_k \text{ is } A_{i1} \text{ and } pH_k \text{ is } A_{i2} \text{ and } T_k \text{ is } A_{i3} \\ & \text{then } Pp_{ki} = a_{i0} + a_{i1}Bx_k + a_{i2}pH_k + a_{i3}T_k, \end{aligned} \quad (17)$$

El procedimiento para la construcción del estimador adaptable es implementado en Matlab. Para obtener el modelo borroso inicial se toman los primeros cinco datos ($m = 5$) y el resto de los datos de los ocho experimentos se toma para el aprendizaje. Luego, con radio umbral $Rthr = 0,16$ y similitud umbral $Sthr = 0,08$, el modelo evoluciona continuamente desde una regla hasta resultar en un modelo borroso evolutivo (eFS) de seis reglas (Fig 1). En la Figura 2 se muestran los resultados gráficos de la estimación del modelo borroso. Se hace notar que se trabaja con los datos de cada experimento una hora consecutiva, correspondiendo el instante de muestro a una hora (tiempo entre mediciones de las variables, como se expuso en el Epígrafe 2.2).

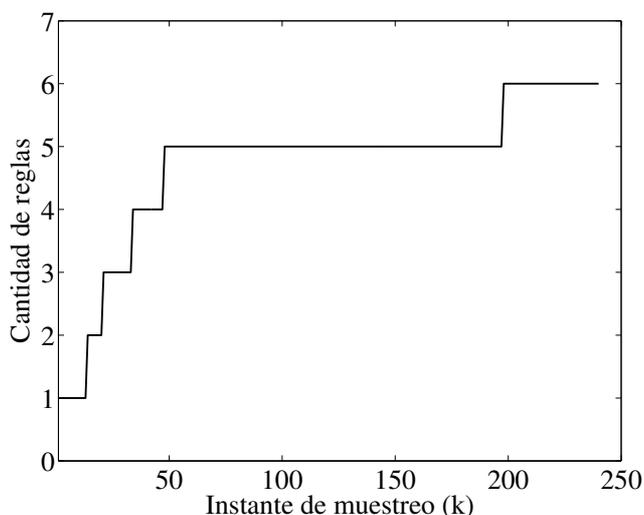


Figura 1. Evolución del número de reglas

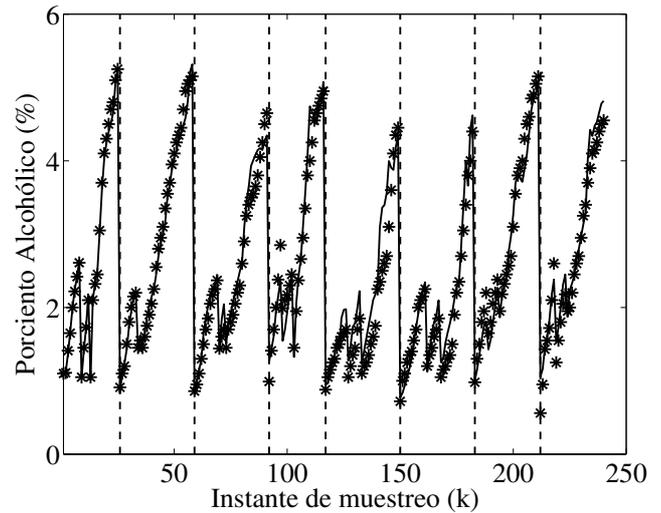


Figura 2. Estimación en línea del porcentaje alcohólico, datos reales (*), estimación del sensor virtual (—). Las líneas verticales discontinuas indican las fronteras entre experimentos.

Con el objetivo de realizar comparaciones, también se obtiene un modelo fuera de línea utilizando el popular método de agrupamiento sustractivo (Chiu, 1994) combinado con mínimos cuadrados, el cual se ajusta con el método ANFIS implementado en Matlab utilizando 20 épocas de entrenamiento. En este caso se obtiene un modelo de ocho reglas con $radii = 0,29$.

Los índices de funcionamiento para la validación empleados aquí son el error cuadrático medio (MSE), la raíz del error cuadrático medio ($RMSE$) y una medida porcentual de la coincidencia entre la salida del sistema y la salida del modelo, el índice VAF (*Variance Accounted For*). Los valores de estos índices numéricos que miden cuán exitoso es el ajuste de ambos modelos borrosos, así como el error máximo ($MaxError$) de cada uno, son mostrados en la Tabla 1. En la misma tabla aparecen los índices correspondientes al modelo analítico (MA) descrito por (1)–(4), con parámetros optimizados para cada fermentación (Peralta *et al.*, 2000). Como se aprecia, sin optimización y con menos reglas, el modelo borroso evolutivo presenta los mejores índices, así como una menor desviación máxima. Se hace notar que los errores de estimación obtenidos por dicho modelo evolutivo son comparables a los de equipos especializados de elevados costos.

Tabla 1. índices de funcionamiento.

índices	eFS	$ANFIS$	MA
MSE	0.058	0.071	1.17
$RMSE$	0.24	0.27	1.08
$MaxError$	0.90	1.29	4.03
VAF	96.39	94.99	61.49

Una característica importante del modelo evolutivo es que constantemente se adapta a las variaciones del proceso. Como ejemplo, la evolución de los parámetros de las dos primeras reglas a medida que llegan nuevos datos se muestra en la Figura 3, donde c_{ij} es el valor del centro de la función de pertenencia $\mu_{A_{ij}}$, σ_i es el ancho de las funciones y a_{ij} son los parámetros del consecuente para la regla $i = 1, \dots, 2$ y la entrada $j = 1, \dots, 3$. Note que cada escalar a_{i0} representa al término independiente de la i -ésima regla.

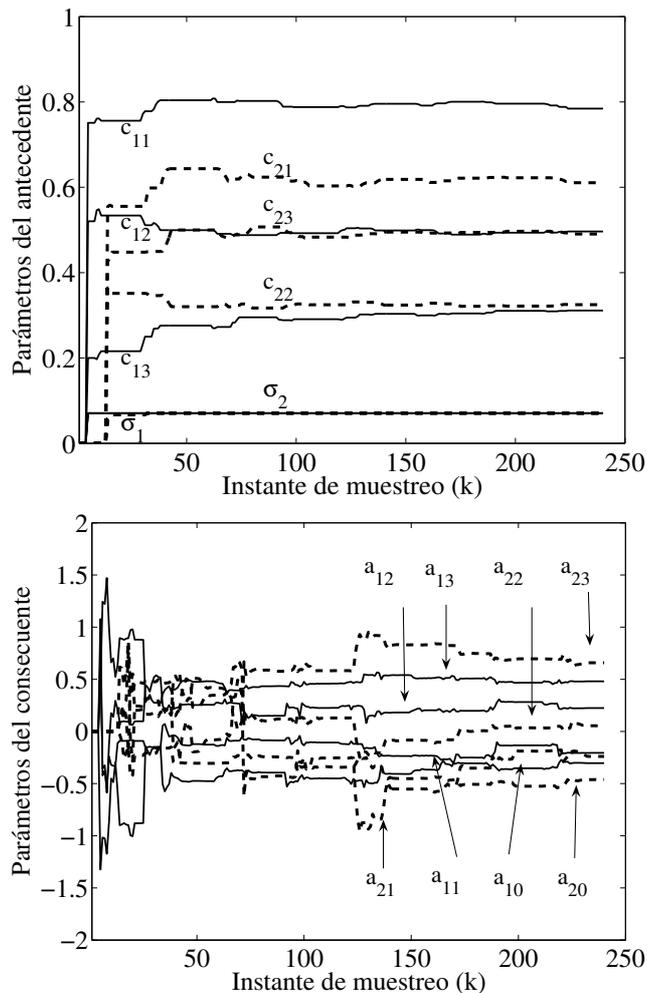


Figura 3. Evolución de los parámetros de las dos primeras reglas del modelo borroso evolutivo empleado como sensor virtual adaptable: (arriba) antecedente, (debajo) consecuente; regla 1 (—), regla 2 (---)

Las características detalladas del modelo al finalizar el aprendizaje se brindan en la Figura 4 y la Tabla 2. La Figura 4 muestra las funciones de pertenencia de cada una de las entradas del sistema. Se hace notar que las funciones de pertenencia 5 y 6 ($mf5$ y $mf6$) en la Entrada 1 (Bx) se funden en una ($mf56$), al emplear el método de simplificación de modelos borrosos para el aprendizaje en línea expuesto en (Victor and Dourado, 2006) con un umbral de similitud $\gamma = 0,9$. De igual forma se unen las funciones de pertenencia 3 con 5 y 2 con 6 en la Entrada 2 (pH), así como las funciones 3 y 5 en la Entrada 3 (T).

Tabla 2. Parámetros del modelo borroso.

R_i	c_{Bx}	c_{pH}	c_T	Pp_i
R_1	0.78	0.50	0.31	$-0.30-0.21Bx+0.22pH+0.48T$
R_2	0.61	0.27	0.50	$-0.46-0.24Bx+0.05pH+0.64T$
R_3	0.16	0.13	0.88	$-0.24-0.71Bx-0.36pH+1.25T$
R_4	0.92	0.66	0.16	$-0.06-0.02Bx+0.29pH+0.14T$
R_5	0.41	0.13	0.88	$+0.17+0.94Bx+1.69pH-1.12T$
R_6	0.41	0.27	0.68	$-0.44-0.41Bx+0.50pH+0.42T$

En la Tabla 2 se brinda información numérica de los antecedentes (centros de las funciones gaussianas) y los consecuentes

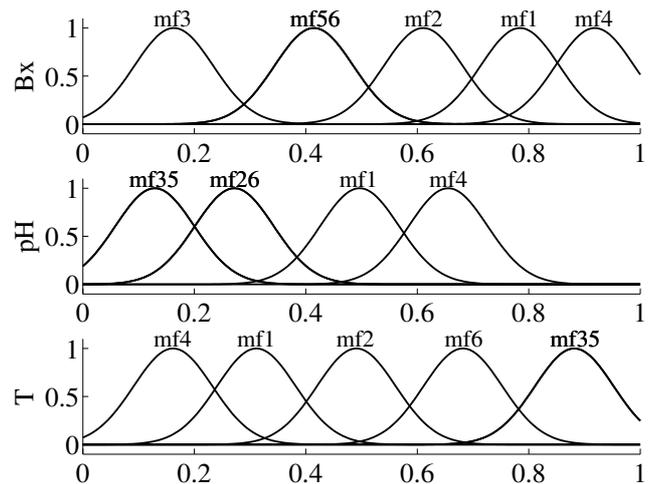


Figura 4. Funciones de pertenencia del sensor virtual

(subsistemas lineales) del sensor virtual basado en el modelo borroso evolutivo.

5. CONCLUSIONES

En este trabajo se ha presentado la obtención de un modelo borroso adaptable para la identificación del proceso fermentativo de etanol que tiene lugar en una de las destilerías de la región central de Cuba, región azucarera por excelencia. Se hace notar que modelos con características similares no fueron encontradas para este proceso en particular.

Los resultados permiten comprobar que el modelo desarrollado es capaz de aprender las relaciones internas entre ciertas variables del proceso fermentativo estudiado y la concentración de etanol en el biorreactor, proceso este que requiere adaptación en tiempo real. Empleando el algoritmo expuesto, se obtiene un modelo compacto, con buenas características de estimación, adaptable y de mejor desempeño cuando se compara con una técnica clásica de modelado borroso fuera de línea. Por tanto, el modelo constituye un sensor virtual útil para estimar la concentración alcohólica en un proceso real de producción de bioetanol, proceso complejo y de característica no lineal.

REFERENCIAS

- Angelov, P. and D. Filev (2004). An approach to online identification of Takagi-Sugeno fuzzy models. *IEEE Trans. on Syst. Man, and Cyb.-Part B* **34**(1), 484–498.
- Angelov, P. and N. Kasabov (2006). Evolving intelligent systems, eIS. *IEEE SMC eNewsLetter* (15), 1–13.
- Babuska, R. (1998). *Fuzzy Modeling for Control*. Kluwer Acad. Pub.. Boston, USA.
- Bastin, G. and D. Dochain (1990). *On-line estimation and adaptive control of bioreactors*. Elsevier Science Publishing. Amsterdam.
- Berovic, M. and Kieran, P., Eds.) (2006). *Bioprocess engineering: doctoral/post-doctoral level*. Ljubljana: Faculty of Chemistry and Chemical Technology.
- Bouchachia, A. and R. Mittermeir (2007). Towards incremental fuzzy classifiers. *Soft Comput.* **11**(2), 193–207.
- Chiu, S.L. (1994). Fuzzy model identification based on cluster estimation. *J. Intell. Fuzzy Syst.* **2**(3), 267–278.
- Díez, J.L., J.L. Navarro and A. Sala (2004). Algoritmos de agrupamiento en la identificación de modelos borrosos. *RIAI* **1**(2), 32–41.

- Díez, J.L., J.L. Navarro and A. Sala (2007). A fuzzy clustering algorithm enhancing local model interpretability. *Soft Comput.* **11**, 973–983.
- Fabelo, J.A. (1999). Estudio de la etapa de fermentación alcohólica utilizando mezcla de diferentes sustratos. Tesis doctoral. Dpto. de Ing. Química, UCLV. Santa Clara, Cuba.
- Fortuna, L., A Rizzo, S. Graziani and M.G. Xibilia (2007). *Soft sensors for monitoring and control of industrial processes*. Advances in Industrial Control. Springer-Verlag London Limited. London, UK.
- Franco, E. (2002). Framework for online modeling, optimization and monitoring of bioprocesses. Tesis doctoral. MLU. Halle-Wittenberg, Alemania.
- Garro, O.A. (1993). Fermentación alcohólica con *Zymomonas* sp. Estudio de modelos matemáticos y su verificación mediante ensayos de fermentación. Tesis doctoral. Universidad Nacional de Tucumán, Argentina.
- Kasabov, N. and Q. Song (2002). DENFIS: Dynamic evolving neural-fuzzy inference system and its application for time-series prediction. *IEEE Trans. on Fuzzy Syst.* **10**(2), 144–154.
- Lughofer, E. and E.P. Klement (2005). FLEXFIS: A variant for incremental learning of Takagi-Sugeno fuzzy systems. In: *FUZZ-IEEE 2005*. Reno, USA. pp. 915–920.
- Martínez, B., F. Herrera, J. Fernández and E. Marichal (2008). Método de agrupamiento en línea para la identificación de modelos borrosos Takagi-Sugeno. *RIAI* **5**(3), 63–69.
- Martínez, B.L. (2007). Identificación borrosa de sistemas no lineales mediante algoritmo de agrupamiento incremental con aplicación a un proceso fermentativo de bioetanol. Tesis doctoral. Dpto. de Automática, UCLV. Santa Clara, Cuba.
- Muñoz, R. and F. Angulo (2006). Aproximación de estimación de estados en un reactor UASB. *Revista Colombiana de Tecnologías de Avanzada* **1**, 27–33.
- Peralta, L.M., M.E. O’Farril and Y. Olivert (2000). Modelación, optimización y control del proceso de fermentación de la producción de alcohol. In: *Iera Conferencia Internacional de Química de la UCLV*. Santa Clara, Cuba.
- Reynaga, C.F. and R.R. Leal (2000). Estimación de biomasa y pigmento en línea para una fermentación tipo fed-batch utilizando redes neuronales artificiales. In: *Congreso SOMI XV*. Guadalajara, Jalisco.
- Roubos, J.A., R. Babuska, P. Krabben and J.J. Heijnen (2000). Hybrid modeling of fed-batch bioprocesses: combination of physical equations with metabolic networks and black-box kinetics. *Journal A* **41**(3), 17–23. ISBN: 0771-1107.
- Salazar, R. (2008). Sensores virtuales: Una revolución en ingeniería de control, (en línea). <http://www.ingenieriaquimica.org/articulos/sensores-virtuales> (Acceso: 15 Junio 2008).
- Sotolongo, E. (2001). Estudio de la preservación del jugo de caña como sustrato en la fermentación alcohólica. Tesis doctoral. Fac. de Ing. Química, ISPJAE. La Habana, Cuba.
- Sundstrom, H. and S.-O. Enfors (2008). Software sensors for fermentation processes. *Bioprocess Biosyst Eng.*
- Takagi, T. and M. Sugeno (1985). Fuzzy identification of systems and its applications to modelling and control. *IEEE Trans. on Syst. Man, and Cyb.* **15**(1), 116–132.
- Victor, J. and A. Dourado (2006). Pruning for interpretability of large spanned eTS. In: *EFS’06*. Lake District, UK. pp. 55–60.