

Estimación de Estados Funcionales en Procesos Complejos con Base en Agrupamiento Difuso

Henry O. Sarmiento^(1,2,*), Claudia V. Isaza⁽³⁾, Tatiana Kempowsky-Hamon^(4,5) y Marie-Veronique LeLann^(4,5)

(1) Grupo de Investigación en Control Automático y Robótica (ICARO), Cra 48 No. 7-151 (P19-145), Politécnico Colombiano Jaime Isaza Cadavid, Medellín, Colombia. (e-mail: hosarmiento@elpoli.edu.co).

(2) Grupo de Electrónica de Potencia Automatización y Robótica (GEPAR), Dpto. Ing. Electrónica, Calle 67 No. 53 - 108 (B20-305), Universidad de Antioquia, Medellín, Colombia.

(3) Grupo de Investigación en Sistemas Embebidos e Inteligencia Computacional (SISTEMIC), Dpto. Ing. Electrónica, Calle 67 No. 53 - 108 (B19-426), Universidad de Antioquia, Medellín, Colombia.

(e-mail: cisaza@udea.edu.co)

(4) CNRS, LAAS, 7 avenue du colonel Roche, F-31400 Toulouse, France.

(5) Univ de Toulouse, INSA, LAAS, F-31400 Toulouse, France.

(e-mail: mvlelann@laas.fr, tkempows@laas.fr)

* Autor a quien debe ser dirigida la correspondencia

Recibido Sep. 12, 2012; Aceptado Nov. 06, 2012; Versión final recibida Dic. 15, 2012

Resumen

Este artículo presenta una metodología para predecir estados funcionales en procesos complejos a partir de la estimación de grados de pertenencia difusos. La propuesta integra una medida estática como es el resultado de un clasificador difuso entrenado con los datos históricos del proceso y un algoritmo de estimación basado en la teoría de Markov para eventos discretos. La propuesta, que puede ser integrada a un sistema de monitoreo de sistemas complejos, comprende dos etapas: una etapa de entrenamiento fuera de línea para definir el clasificador difuso y el estimador; y una etapa en línea donde se realizan la clasificación de la situación actual del proceso y la estimación del estado funcional para el siguiente tiempo de muestreo. La propuesta desarrollada para la estimación de estados funcionales permite utilizar cualquier método de agrupamiento difuso que suministre la información base que requiere la metodología. La metodología fue probada con éxito en un sistema de monitoreo para una línea de transmisión de energía y en el monitoreo de un sistema de caldera.

Palabras clave: predicción de estados funcionales, clasificador difuso, agrupamiento difuso, cadenas de Markov

Estimation of Functional States in Complex Processes Based on Fuzzy Clustering

Abstract

This paper presents a methodology to predict functional states in complex processes from the estimation of fuzzy membership degrees. The proposal integrates a static measure, such as the result of a fuzzy classifier trained with historical process data, and an estimation algorithm based on Markov theory for discrete events. The proposal, which can be integrated to the monitoring of complex systems, provides two stages: an off-line training stage to define the fuzzy classifier and the estimator; and an online stage where the classification of the current process situation and the estimation of the next functional state are performed. The proposal for the estimation of functional states allows using any fuzzy clustering method that provides the information required by the methodology. The proposed methodology was successfully tested on a monitoring system for a power transmission line and in the monitoring of a boiler system.

Keywords: functional state prediction, fuzzy classifier, fuzzy clustering, Markov's chains

INTRODUCCIÓN

Disponer de la información necesaria para la toma de decisiones en un proceso industrial permite reducir pérdidas económicas, ambientales y humanas. El monitoreo, como parte fundamental en los sistemas de supervisión y diagnóstico de fallas, da al operador del proceso los mecanismos necesarios para reconocer el estado funcional actual, la posibilidad de interacción amigable con el proceso y el registro del comportamiento histórico (Colomer et al., 2000). Una de las técnicas - basadas en datos- utilizadas para el monitoreo de procesos complejos es la de clasificadores (Cuadrado, 2002). Los clasificadores entrenados con la información histórica tienen la tarea de asociar los valores de las variables del proceso (en cada tiempo de muestreo) a una clase o estado funcional del proceso (e.g. funcionamiento normal, estado de alarma, estado de falla, etc.). La información de la clase o estado del proceso es suministrada al operador como elemento de análisis para la toma de decisiones requerida.

Clasificadores concretos y difusos en sistemas de monitoreo han sido reportados en la literatura (De Silva et al., 2008; Teppola et al., (1998a, 1998b); Jyoti y Singh, 2011). El clasificador concreto asocia el vector de datos en cada instante de muestreo a una sola clase, mientras que el clasificador difuso lo asocia a cada una de las clases con un cierto grado de pertenencia. La información de los grados de pertenencia permite tener un conocimiento más detallado de la situación actual del proceso por parte del operador.

Para obtener las clases de los sistemas clasificadores se pueden usar algoritmos de agrupamiento difuso. Algunos de los algoritmos de agrupamiento difuso utilizados para el entrenamiento de los clasificadores difusos son Fuzzy C-means (FCM) (Bezdek, 1981), GK-means (GKM) (Gustafson y Kessel, 1978), Learning Algorithm for Multivariate Data Analysis (LAMDA) (Aguilar-Martin y Lopez de Mantaras, 1982), entre otros. Todos estos algoritmos hacen un análisis de relación n-dimensional independiente del tiempo y han presentado un muy buen desempeño en sistemas de monitoreo a nivel industrial en la identificación de estados (Teppola et al. (1998a, 1998b); Kemposky et al., 2006). Los clasificadores hacen un análisis multivariado en línea (durante el funcionamiento del proceso a partir del vector de datos en el instante t) lo que permite identificar el estado actual (de falla o no) del sistema.

El análisis independiente del tiempo (análisis estático), que es característico de los clasificadores basados en datos, no tiene en cuenta la naturaleza dinámica de los procesos reales. Para tener en cuenta la dinámica del proceso y mejorar los sistemas de monitoreo y supervisión de procesos, algoritmos de estimación o predicción han sido utilizados (Isermann y Ballé, 1997; Isermann, 1984).

En general, diversos algoritmos son utilizados para la estimación de estados funcionales en forma directa e indirecta. La forma indirecta permite estimar el estado funcional del proceso como consecuencia de estimar variables y/o parámetros del proceso (Weber y Gentil, 1997). La estimación directa del estado funcional es obtenida con base en el conocimiento que se puede extraer de la información histórica y actual del proceso.

Desde hace varias décadas y con diferentes técnicas, se ha registrado la estimación de variables y/o parámetros en sistemas donde se puede disponer de un modelo matemático aproximado. La estimación de estados (variables de estado) y la estimación de parámetros ha sido utilizada en forma efectiva en los sistemas de supervisión basados en detección de fallos (Zapata et al., 2011). Trabajos que incluyen estimación de estados han sido ampliamente reportados en diversas áreas, entre ellos, los sistemas eléctricos de potencia (Zhang et al., 2011) donde los Filtros de Kalman Extendidos -EKF(Extended Kalman Filter)- (Van Der Merwe y Wan; 2001) y los Cuadrados Mínimos Pesados -WLS(Weighted Least Square)- (Nguyen y Shin, 2010) son utilizados como algoritmos de estimación. En procesos químicos, la estimación de estado puede ser realizada por medio de EKFs, o una amplia variedad de Observadores de estado (Botero y Álvarez, 2008). En sistemas de control de tráfico los EKFs, los filtros UKF(Unscented Kalman Filter), filtros de partículas PF(Particles Filters), filtros duales o filtros conjuntos son los algoritmos utilizados (Hegyí et al., 2006). En sistemas climatológicos (estimación de lluvia y clima) han sido utilizados EKF, UKF y filtros de Kalman conjuntos (Infante et al., 2012).

En el trabajo de Prakash et al. (2011) es presentada una categorización para los métodos de estimación (basados en modelos) de estados en sistemas no lineales: los estimadores no lineales desarrollados bajo un marco de referencia determinístico, y los estimadores no lineales desarrollados bajo un marco de referencia Bayesiano. En general, los autores concluyen que los métodos de estimación que han sido desarrollados durante los últimos 40 años, son computacionalmente intensivos y requieren todavía de mucho refinamiento.

Dentro de las técnicas de estimación o predicción de estados funcionales en forma directa, aplicadas al monitoreo de procesos complejos, están los Modelos de Markov Ocultos - HMM (Hidden Markov Models)- (Daidone et al., 2006), y en años más recientes la utilización de técnicas de Inteligencia Artificial como Redes Neuronales Artificiales -ANN (Artificial Neural Networks)-, Máquinas de Soporte Vectorial -SVM

(Support Vector Machines)- y Logica Difusa -FL(Fuzzy Logic)- entre otros (Roth et al., 2010). En los HMMs y las ANNs(Redes Recurrentes) la estimación es del tipo probabilística, y la dependencia de los datos históricos para definir sus estructuras es considerada una debilidad potencial. Ejemplos específicos de estimación de estados con HMMs para sistemas mecánicos son detallados por Roth et al. (2010), y Tobon-Mejía et al. (2011).

Trabajos de predicción de estados funcionales basándose en los resultados de algoritmos de clasificación difusa no se han reportado en la literatura. Sin embargo, hay desarrollos que se pueden considerar como antecedentes; estos están relacionados con el establecimiento de lazos de conexión entre estados funcionales para conformar un autómata (Klir and Yuan, 1995) direccionado hacia la tarea de predicción (determinar los posibles cambios durante el comportamiento del proceso). Autómatas difusos han sido construidos por medio de inferencia basada en reglas difusas (Chang and Chen, 2011) y a partir de los grados de membrecía difusos (Waissman et al., 2000 y 2005; Kempowsky et al., 2006; Botia et al., 2012), pero sin asociar a los lazos de conexión un peso, fuerza o valor que pueda dar indicación de la evolución de los estados.

Otro antecedente relevante, es el trabajo de Du y Yeung (2004), donde basándose en un esquema de supervisión con clasificadores, propusieron una metodología para el cálculo de Probabilidades de Transición Difusas (PTD). La propuesta permite estimar la probabilidad que el sistema permanezca en el estado actual y evolucione hacia los demás estados. En esta propuesta, para incluir la información temporal, la formulación matemática de las Cadenas de Markov para eventos discretos es integrada. Los autores (Du y Yeung, 2004) proponen utilizar en lugar de probabilidades clásicas, las Probabilidades Difusas (PD) (Zadeh, 1984), que son obtenidas mediante un clasificador desarrollado por los autores. La propuesta se limita a procesos con estados no renovables, es decir que el sistema no retorna a estados que ya se presentaron durante el funcionamiento del proceso. Un ejemplo de procesos no renovables son los procesos degradantes, como el proceso natural de desgaste de piezas mecánicas que es tratado por Du y Yeung (2006) para mostrar la efectividad de su propuesta.

En el campo de procesos complejos, como son la mayoría de procesos industriales, el sistema puede -en su normal evolución- retornar entre estados, e.g. el efecto de la acción de un controlador de nivel, presión, entre otras, puede llevar al proceso a restablecer estados. Una propuesta que permita -estando en el instante t - establecer la probabilidad del próximo estado posible en $t+1$, que se base en clasificadores difusos, y que permita trabajar con procesos renovables, constituye una herramienta valiosa para el operador del proceso como apoyo en la toma de decisiones.

En este artículo se propone un nuevo método para la estimación en línea de estados funcionales en procesos complejos. La propuesta integra las ventajas de los clasificadores difusos, la tarea de estimación y la característica de ser aplicable a procesos renovables. La parte esencial de la metodología es la integración de la información difusa -obtenida con el agrupamiento- a la formulación matemática de las Cadenas de Markov para hacer posible la tarea de estimación.

En la primera sección se presentan generalidades del agrupamiento difuso y las características que lo hacen útil en un esquema de monitoreo de procesos. En la segunda sección, la fundamentación teórica y el desarrollo de la propuesta son presentados. La tercera sección describe cómo la propuesta de estimación de estados es integrada a un esquema de monitoreo basado en agrupamiento difuso, y detalla la forma en que son realizadas las etapas de entrenamiento y monitoreo para un ejemplo sencillo. La cuarta sección describe los casos de estudio; el primero de los cuales (sistema de monitoreo térmico para una línea de transporte de energía) permite conocer de manera detallada la aplicación de la propuesta realizada, y presenta los resultados obtenidos para un segundo caso (subsistema de caldera). Las conclusiones y trabajos futuros están incluidas en la quinta y última sección.

MONITOREO DE PROCESOS BASADO EN AGRUPAMIENTO DIFUSO

Los algoritmos de agrupamiento difuso son utilizados en la tarea de identificar clases que representen los estados funcionales presentes en un sistema o proceso a partir de sus datos históricos. La tarea de asociación de las clases obtenidas a estados funcionales es realizada con el conocimiento del experto en el proceso. Las dos tareas anteriores constituyen la etapa de aprendizaje del clasificador difuso. En la etapa de monitoreo, ya con el proceso en línea, se toman datos del proceso (en cada instante de muestreo) que son reconocidos y asociados a los estados funcionales por el sistema clasificador. Entonces, teniendo en cuenta que un Estado funcional está definido como la caracterización de la situación en que se encuentra un proceso en función de su pasado, de las funciones que debe desempeñar y de las posibles transiciones hacia nuevas situaciones (Aguilar-Martín, 2007), el clasificador es una herramienta de soporte valiosa en la toma de decisiones para el operador del proceso.

Agrupamiento Difuso

Los métodos de agrupamiento se encargan de agrupar datos en diferentes clases a partir de una medida de similitud (Omran et al., 2007). En los procesos los datos corresponden a las muestras que son tomadas por medio de los sistemas SCADA (*Supervisory Control And Data Acquisition*), y los grupos pueden ser asociados a los estados funcionales. En el caso de agrupamiento concreto, cada muestra es comparada con los centros de clase mediante una medida de similitud para determinar la pertenencia o no a un grupo y, en el caso de agrupamiento difuso, la comparación se hace para determinar el grado de pertenencia a cada clase.

El grupo de n vectores de datos $X=[x_1, x_2, \dots, x_n]$ con a atributos cada uno, donde n corresponde al número de muestras (tiempos de muestreo), es dividido (usando un método de agrupamiento) en m grupos o clases. El agrupamiento difuso permite obtener la matriz de grados de pertenencia $U=[\mu_{jf}]_{m \times n}$, donde μ_{jf} representa el grado de pertenencia difusa de la muestra f al j -ésimo grupo. A partir del análisis de los grados de pertenencia difusa por cada muestra f , se puede determinar la clase en la que es ubicada la muestra. En general, el mayor grado de pertenencia determina la clase que es asignada a la muestra (ver (1)).

$$C_f = j, \quad \text{donde } j \text{ corresponde a la clase en la cual } \mu_{jf} \text{ es máximo } (\max(\mu_{j\Box})); j = 1, 2, \dots, m., \text{ para cada } f \quad (1)$$

Para el agrupamiento difuso se han propuesto diferentes métodos, entre los cuales los más conocidos son los basados en distancia. Dentro de los métodos basados en distancia se incluyen los algoritmos Fuzzy C-means (FCM) (Bezdek, 1981) y Gustafson-Kessel Means (GKM) (Gustafson, 1978) que utilizan un criterio de optimización J_b (ver ec. (2)) que hace posible agrupar los datos de acuerdo con la similitud entre ellos. El valor del subíndice b regula la 'difusividad' de la partición.

$$J_b(U, v) = \sum_{f=1}^n \sum_{j=1}^m (\mu_{jf})^b (d_{jf})^2 \quad (2)$$

En estos algoritmos, la similitud es evaluada por medio de la función de distancia d_{jf} (3), que se mide con respecto a unos centroides o prototipos $v=\{v_1, v_2, \dots, v_m\}$.

$$d_{jf} = (x_f - v_j)^T H_j (x_f - v_j) \quad (3)$$

En el caso de FCM, $H_j=1$ y d_{jf} es la distancia Euclidiana que permite hacer agrupamiento de datos en hiperesferas. Para GKM, H_j está definido de acuerdo con la ecuación (4), donde δ_j es el índice volumétrico del grupo j , F_j es la matriz de covarianza difusa del grupo j , y la agrupación de datos corresponde a hiperelipsoides que permiten una mayor posibilidad de adaptación a las formas de agrupamiento que las hiperesferas. Los algoritmos FCM y GKM requieren conocer de antemano el número de clases.

$$H_j = [\delta_j \det(F_j)]^{1/n} (F_j)^{-1} \quad (4)$$

Learning Algorithm for Multivariate Data Analysis (LAMDA) (Aguilar-Martín y Lopez de Mantaras, 1982) está basado en encontrar los grados de adecuación global (GAD) de la muestra a cada una de las clases, estos GADs son los grados de pertenencia difusos. El método puede trabajar en forma supervisada donde requiere conocer datos históricos con su respectiva asociación de clases, y no supervisada donde naturalmente identifica los grupos. Los GADs son el resultado de realizar operaciones difusas (operadores de agregación difusos) con los grados de adecuación marginal (MAD), donde los MAD son los valores que representan las contribuciones de cada descriptor de la muestra (vector x) a cada clase. LAMDA permite procesar información cuantitativa, cualitativa e intervalos (Hedhazi et al., 2011), y no requiere conocer el número de clases *a priori*.

Una característica de las técnicas de agrupamiento difuso es la independencia con el tiempo. El análisis de agrupamiento se hace a partir de la similitud en el espacio de datos (vector x), donde el tiempo no se tiene en cuenta. Las clases obtenidas con el agrupamiento permiten definir un clasificador que utiliza la misma medida de similitud empleada en el entrenamiento (distancia en el caso de GKM y FCM, y MAD en el caso de LAMDA) para, ya en línea (durante el funcionamiento del proceso), estimar los grados de pertenencia del vector de datos en un instante t a las clases creadas.

Los clasificadores difusos han sido integrados con éxito en esquemas de monitoreo y supervisión de procesos (Jyoti y Singh, 2011; Aguado y Rosen, 2008; Kempowsky et al., 2006). En estos esquemas la información obtenida con los clasificadores difusos constituye una herramienta de trabajo para el operador del proceso.

Monitoreo y supervisión basado en Agrupamiento Difuso

El clasificador entrenado con técnicas de agrupamiento difuso es la base del esquema de monitoreo mostrado en la Fig. 1.

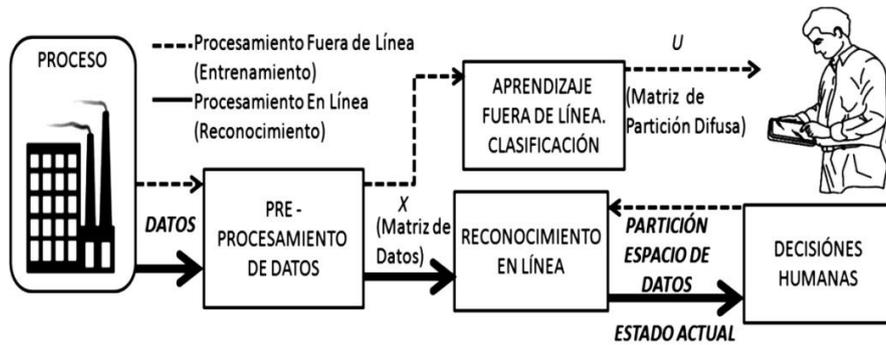


Fig. 1: Esquema de monitoreo basado en clasificadores (Isaza et al., 2007)

En el esquema están presentes una etapa de aprendizaje fuera de línea y una etapa de reconocimiento que se ejecuta en línea. En la etapa de entrenamiento se utilizan los datos históricos del proceso para entrenar (modelar los estados funcionales a través de los grupos) un clasificador difuso. Posteriormente, el clasificador es utilizado en línea (reconocimiento) para procesar cada nueva muestra tomada del proceso. El resultado es ofrecer al operador información en tiempo real de los grados de pertenencia de la muestra a cada clase o, mediante un ciclo de decisión final asignar la clase a la que pertenece la muestra. Estos esquemas, al ser construidos con base en clasificadores con agrupamiento difuso, tienen las mismas características que se describieron en la sección anterior. Entre estas características está la de ser un análisis que no incluye la evolución temporal.

Con el objetivo de mejorar los esquemas de supervisión y monitoreo de procesos una de las estrategias es la de incorporar algoritmos de estimación o predicción (Isermann, 1984). Estos algoritmos permiten presentarle al operador los posibles panoramas o situaciones futuras (en el instante $t+1$) que se pueden presentar de acuerdo con la tendencia actual del proceso.

ESTIMACIÓN DE ESTADOS A PARTIR DE GRADOS DE PERTENENCIA

En la propuesta de estimación de estados, dos elementos fundamentales se destacan inicialmente: la información requerida y la construcción del algoritmo de estimación. Los grados de pertenencia obtenidos al aplicar algún algoritmo de agrupamiento a la base de datos histórica del proceso son la información básica para la propuesta que se presenta en este artículo. Esta información está normalmente contenida en la Matriz de Grados de Pertenencia Difusa (U). En relación con el algoritmo de estimación, se toma como base el trabajo sobre Probabilidades de Transición Difusa propuesto por Du y Yeung (2004). Detalles relevantes de este trabajo se describen a continuación.

Probabilidades de Transición Difusa

Du y Yeung (2004) basan su propuesta en el cálculo de probabilidades asociadas a eventos discretos establecida por Markov (Ross, 1996). Markov demostró que se puede calcular la probabilidad P_j que se presente el evento discreto j en el tiempo (muestra) $t+1$ si se conoce la probabilidad P_k del evento discreto k en el tiempo t , ec. (5).

$$P_j(S_{t+1}) = \sum_{k=1}^m P_{kj} * P_k(S_t), \quad j = 1, 2, \dots, m \tag{5}$$

Donde la condición en (6) se debe satisfacer, P_{kj} es la probabilidad de transición entre eventos discretos de k a j , m el número de eventos discretos (clases en un clasificador) y S corresponde al término de muestra.

$$1 = \sum_{j=1}^m P_{kj}, \quad k = 1, 2, \dots, m \tag{6}$$

A partir de la ec. (5), Du y Yeung proponen trabajar con probabilidades difusas y hacer la estimación de la probabilidad difusa (FP) para el estado discreto j en el instante de muestreo ($t+1$). Esta probabilidad (ver ec. (7)) se estima a partir de la Probabilidad de Transición Difusa entre los estados j y k (FP_{kj}) y la

probabilidad difusa $FP_k(S_t)$ que el proceso se encuentre en el estado k en el instante actual de muestreo t . Y se debe cumplir la condición especificada en ec. (8).

$$FP_j(S_{t+1}) = \sum_{k=1}^j FP_{kj} * FP_k(S_t), \quad j = 1, 2, \dots, m., \text{ donde en el límite superior de la sumatoria } k = j \tag{7}$$

$$1 = \sum_{j=1}^m FP_{kj}, \quad k = 1, 2, \dots, m \tag{8}$$

Las probabilidades difusas son obtenidas mediante un clasificador (basado en intervalos) desarrollado por los autores (Du y Yeung, 2004). Definir el clasificador y calcular las FP_{kj} constituyen la etapa de entrenamiento fuera de línea del sistema de monitoreo. El cálculo de la estimación (etapa de monitoreo en línea) de la probabilidad difusa $FP_j(S_{t+1})$ requiere de las FP_{jk} encontradas en entrenamiento y de la probabilidad difusa actual $FP_k(S_t)$ calculada por el clasificador. De acuerdo con (7) donde $k=j$ en el límite superior de la sumatoria, la aplicación de la propuesta está restringida a procesos no renovables.

Método propuesto para la estimación de estados

Teniendo en cuenta que la mayoría de procesos complejos presentan transiciones con retorno entre estados funcionales, el estudio de la propuesta de Du y Yeung (2004) permite plantear una generalización donde se incluyan transiciones de retorno. Así, la ecuación (7) es modificada en la sumatoria para el valor de m ($k=m$) y se propone la ecuación (9).

$$FP_j(S_{t+1}) = \sum_{k=1}^m FP_{kj} * FP_k(S_t), \quad j = 1, 2, \dots, m., \text{ donde en el límite superior de la sumatoria } k = m \tag{9}$$

Ya que se busca estimar los grados de pertenencia que va a tener el proceso en el tiempo $t+1$ para los estados funcionales j ($j=1,2,\dots,m$) presentes en el proceso, la relación bijectiva existente entre posibilidad (π) y probabilidad (p) demostrada por Dubois y Prade (1983), permitiría proponer el reemplazo de las probabilidades por los grados de pertenencia difusa en (9). La propuesta de Dubois y Padre se basó en el Principio de Consistencia establecido por Zadeh (1978). Dubois y Prade (1983) dedujeron una formulación matemática para hacer el cálculo de probabilidades a partir de grados de pertenencia (10) y viceversa (11). Como consecuencia, el comportamiento de los eventos discretos expresado en términos de probabilidades y el comportamiento de los eventos discretos expresado en términos de posibilidades –donde las probabilidades son obtenidas a partir de las posibilidades o viceversa– conserva el mismo patrón de comportamiento y la relación entre los valores correspondientes.

$$\forall_i, i = 1, \dots, m. \quad p_i = \sum_{j=i}^m \frac{1}{j} (\pi_j - \pi_{j+1}) \tag{10}$$

$$\forall_i, i = 1, \dots, m. \quad \pi_i = \sum_{j=1}^m \min(p_i, p_j) \tag{11}$$

De tal manera que a partir de (9), teniendo en cuenta la posibilidad de retorno entre estados, se pueden reemplazar las probabilidades difusas por los grados de pertenencia obtenidos por cualquier método de agrupamiento difuso. Así, el grado de pertenencia del proceso en el instante $t+1$ al estado j , se puede definir como:

$$\mu_j(S_{t+1}) = \sum_{k=1}^m WFT_{kj} * \mu_k(S_t), \quad j = 1, 2, \dots, m \tag{12}$$

La ecuación matricial (13), que equivale al sistema de ecuaciones obtenido a partir de la ec. (12), permitiría estimar los grados de pertenencia difusos del proceso para cada uno de los estados funcionales j [$\mu(S_{2 \dots n})$] –en el instante $(t+1)$ –, conociendo los grados de pertenencia difusos [$\mu(S_{1 \dots (n-1)})$] –en el instante t –, y el vector de Pesos de Transición Difusa, WFT .

$$[\mu(S_{2 \dots n})] = [\mu(S_{1 \dots (n-1)})] * [WFT_{kj}] \tag{13}$$

Con la información histórica del proceso, la ecuación (13) –con más ecuaciones que incógnitas– es solucionada para WFT usando el método de Mínimos Cuadrados con restricción de la solución a valores

mayores o iguales que cero. El vector WFT representa las posibles transiciones ponderadas que pueden ocurrir –en el tiempo t – desde en el estado k a los diferentes estados del proceso $j(j=1,2, \dots k, \dots m)$. Donde

$$[\mu(S_{2...n})] = \begin{bmatrix} \mu_1(S_2) \\ \vdots \\ \mu_1(S_n) \\ \mu_2(S_2) \\ \vdots \\ \mu_2(S_n) \\ \vdots \\ \mu_m(S_2) \\ \vdots \\ \mu_m(S_n) \end{bmatrix}, [WFT_{kj}] = \begin{bmatrix} WFT_{11} \\ \vdots \\ WFT_{1m} \\ WFT_{21} \\ \vdots \\ WFT_{2m} \\ \vdots \\ WFT_{m1} \\ \vdots \\ WFT_{mm} \end{bmatrix} \tag{14}$$

y

$$[\mu(S_{1...(n-1)})] = \begin{bmatrix} \vec{A}_1 & \vec{0} & \dots & \vec{0} & \vec{0} \\ \vec{0} & \vec{A}_1 & & \vec{0} & \vec{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vec{0} & \vec{0} & \dots & \vec{A}_1 & \vec{0} \\ \vec{0} & \vec{0} & \dots & \vec{0} & \vec{A}_1 \end{bmatrix} \begin{bmatrix} \vec{A}_2 & \vec{0} & \dots & \vec{0} & \vec{0} \\ \vec{0} & \vec{A}_2 & & \vec{0} & \vec{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vec{0} & \vec{0} & \dots & \vec{A}_2 & \vec{0} \\ \vec{0} & \vec{0} & \dots & \vec{0} & \vec{A}_2 \end{bmatrix} \dots \begin{bmatrix} \vec{A}_2 & \vec{0} & \dots & \vec{0} & \vec{0} \\ \vec{0} & \vec{A}_2 & & \vec{0} & \vec{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vec{0} & \vec{0} & \dots & \vec{A}_2 & \vec{0} \\ \vec{0} & \vec{0} & \dots & \vec{0} & \vec{A}_2 \end{bmatrix} \tag{15}$$

Donde n es el número de muestras (vectores de datos de entrenamiento), \vec{A} corresponde al vector de grados de pertenencia del instante anterior (que en el caso de datos históricos para el estado j va de $\mu_j(S_t)$ a $\mu_j(S_{n-1})$) y

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}; \vec{A}_i = \begin{bmatrix} \mu_i(S_1) \\ \mu_i(S_2) \\ \vdots \\ \mu_i(S_{n-1}) \end{bmatrix}, \quad i = 1, 2, \dots m. \tag{16}$$

La ec. (12), de donde se obtiene la ec. (15), es desarrollada para cualquier estado funcional j en la ec. (17). En la ec. (17), los valores denominados Pesos de Transición Difusa (WFT) existen porque se tiene en cuenta la transición desde cualquier estado existente al estado j (esto representa la posibilidad de transiciones de retorno). Entonces, el grado de pertenencia al estado j ($\mu_j(S_{t+1})$) –en el instante $t+1$ – es calculado con:

$$\mu_j(S_{t+1}) = WFT_{1j} * \mu_1(S_t) + WFT_{2j} * \mu_2(S_t) + \dots + WFT_{mj} * \mu_m(S_t) \tag{17}$$

Las ecuaciones (13), (14) y (15) tienen como argumentos de entrada grados de pertenencia difusa (obtenidos al usar un método de agrupamiento difuso), y el resultado en la etapa de entrenamiento es la matriz de pesos de transición difusa (WFT).

Para tener en cuenta la dinámica del proceso, además de los valores de los grados de pertenencia que se obtienen a partir del histórico de datos, se propone incluir el cambio de los grados de pertenencia $\Delta\mu$. Incluir $\Delta\mu$ permite tener en cuenta, en forma directa, la información de las variaciones –tendencia (Stewart, 2012)– en los grados de pertenencia a medida que evoluciona el proceso. En forma general, para cualquier estado j , el cambio en los grados de pertenencia entre dos instantes de tiempo está definido por:

$$\Delta\mu_j = \mu_j(S_t) - \mu_j(S_{t-1}) \tag{18}$$

Estos cambios en los grados de pertenencia, evaluados para el histórico de datos del proceso, son el argumento de entrada a las ecuaciones (19), (20) y (21), donde se reformulan las ecuaciones (13), (14) y (15) con el objetivo de poder estimar los cambios en los grados de pertenencia difusa. En la etapa de entrenamiento, las ecuaciones (19), (20) y (21) son evaluadas para encontrar el vector Delta de Pesos de Transición Difusa (ΔWFT).

$$[\Delta\mu(S_{2...n-1})] = [\Delta\mu(S_{1...(n-2)})] * [\Delta WFT_{kj}] \tag{19}$$

Donde

$$[\Delta\mu(S_{2...n-1})] = \begin{bmatrix} \Delta\mu_1(S_2) \\ \vdots \\ \Delta\mu_1(S_{n-1}) \\ \Delta\mu_2(S_2) \\ \vdots \\ \Delta\mu_2(S_{n-1}) \\ \vdots \\ \Delta\mu_m(S_2) \\ \vdots \\ \Delta\mu_m(S_{n-1}) \end{bmatrix}, [\Delta WFT_{kj}] = \begin{bmatrix} \Delta WFT_{11} \\ \vdots \\ \Delta WFT_{1m} \\ \Delta WFT_{21} \\ \vdots \\ \Delta WFT_{2m} \\ \vdots \\ \Delta WFT_{m1} \\ \vdots \\ \Delta WFT_{mm} \end{bmatrix} \tag{20}$$

y

$$[\Delta\mu(S_{1...(n-2)})] = \begin{bmatrix} \vec{0} & \vec{0} & \dots & \vec{0} & \vec{0} \\ \vec{B}_1 & \vec{0} & \dots & \vec{0} & \vec{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vec{0} & \vec{0} & \dots & \vec{B}_1 & \vec{0} \\ \vec{0} & \vec{0} & \dots & \vec{0} & \vec{B}_1 \end{bmatrix} \begin{bmatrix} \vec{B}_2 & \vec{0} & \dots & \vec{0} & \vec{0} \\ \vec{0} & \vec{B}_2 & \dots & \vec{0} & \vec{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vec{0} & \vec{0} & \dots & \vec{B}_2 & \vec{0} \\ \vec{0} & \vec{0} & \dots & \vec{0} & \vec{B}_2 \end{bmatrix} \dots \begin{bmatrix} \vec{B}_m & \vec{0} & \dots & \vec{0} & \vec{0} \\ \vec{0} & \vec{B}_m & \dots & \vec{0} & \vec{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vec{0} & \vec{0} & \dots & \vec{B}_m & \vec{0} \\ \vec{0} & \vec{0} & \dots & \vec{0} & \vec{B}_m \end{bmatrix} \tag{21}$$

Donde B corresponde al vector de cambio de grados de pertenencia del instante anterior y

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}; \vec{B}_i = \begin{bmatrix} \Delta\mu_i(S_1) \\ \Delta\mu_i(S_2) \\ \vdots \\ \Delta\mu_i(S_{n-1}) \end{bmatrix}, \quad i = 1, 2, \dots, m \tag{22}$$

El cálculo de WFT y ΔWFT –en la etapa de entrenamiento– va a permitir la estimación de los grados de pertenencia para la muestra en S_{t+1} durante la etapa de monitoreo.

Utilizando la ecuación (13), se estiman los grados de pertenencia $\mu_j(S_{t+1})$ a todos los m estados funcionales –en el instante $t+1$ –. Para esto se utiliza la matriz WFT y los grados de pertenencia difusa $\mu_j(S_t)$ actuales que son entregados por el clasificador.

El cálculo del cambio entre grados de pertenencia ($\Delta\mu_j(S_t)$, ec. (18)) es realizado para los grados de pertenencia entregados por el clasificador, y es utilizado en el cálculo de la estimación de cambio $\Delta\mu_j(S_{t+1})$ de acuerdo con la ec. (19) que se expresa en términos de $\Delta\mu_j(S_t)$ y ΔWFT así: $\Delta\mu_j(S_{t+1}) = \Delta\mu_j(S_t) * \Delta WFT$.

Entonces, el cálculo de la denominada Estimación Inicial de Grados de Pertenencia ($\mu_j'(S_{t+1})$, ec. (23)) incluye la estimación de grados de pertenencia $\mu_j(S_{t+1})$ y la estimación de cambio $\Delta\mu_j(S_{t+1})$.

$$\mu_j'(S_{t+1}) = \mu_j(S_{t+1}) + \Delta\mu_j(S_{t+1}) \tag{23}$$

Para fortalecer la tendencia de los grados de pertenencia en los cambios de estado, se propone incluir una medida de la cantidad de información con la que se clasifica a un individuo en una clase en el cálculo de la estimación final de los grados de pertenencia $\mu_j''(S_{t+1})$. Cuando se presenta un cambio de estado en un proceso, la relación de los grados de pertenencia para las diferentes muestras cambia. Si el tiempo de muestreo elegido para el proceso es lo suficientemente corto para tener muestras intermedias durante el cambio de estado, se evidencia que en la transición la pertenencia de las muestras a las diferentes clases no está bien definida. En la transición los grados de pertenencia de cada muestra $\mu_j(S_t)$ a las clases ($j=1, \dots, m$) pueden ser muy similares entre sí. Si se tiene en cuenta que el índice de información propuesto por Isaza [23] es una medida de qué tanta fortaleza tiene el grado de pertenencia mayor μ_M (el conjunto difuso se asemeja a un conjunto tipo *singleton*) a los demás μ_j , permite identificar que se está presentando una transición en el proceso. El índice de información $I_D(\mu)$ es evaluado de acuerdo con (24) entregando valores entre '0' y '1'; '0' corresponde a la situación en la que todos los grados de pertenencia son iguales, y '1' en que uno de los grados de pertenencia es el máximo y los demás son cero. El vector $\lambda_i(\mu)$ contiene las diferencias entre el grado de pertenencia máximo μ_M y los demás grados de pertenencia del mismo individuo o muestra (ver ec. (25)). Donde $\mu_M = \max[\mu_i]$ y $C = m - 1$.

$$I_D(\mu) = \sum_i \lambda_i \cdot e^{\lambda_i} / C \cdot \mu_M \cdot e^{\mu_M}, \quad i = 1, 2, \dots, m \tag{24}$$

$$\lambda_i(\mu) = \{\mu_M - \mu_i\}_{i=M}, \quad i = 1, 2, \dots, m \tag{25}$$

El índice I_D (ec. (24)), basado en la entropía difusa, presenta valores bajos para los grados de pertenencia registrados durante las transiciones entre estados del proceso. Esto porque, durante las transiciones, no hay un grado de pertenencia que tenga más ‘fuerza’ con respecto a los demás; es decir, no hay un grado de pertenencia que sea significativamente mayor que todos los otros. Por el contrario, antes o después de la transición, cuando el sistema definitivamente se encuentra en un estado funcional (éste se ha definido con certeza), el índice tiene un valor mayor.

En las muestras próximas a la transición, se requiere que la tendencia en la estimación de los grados de pertenencia ($\Delta\mu(S_{t+1})$, ec. (26)) sea más fuerte para que se puedan evidenciar los posibles cambios. Entonces, el cálculo final de estimación ($\mu_j''(S_{t+1})$, ec. (27)) incluye el inverso de I_D calculado en t como el factor de proporcionalidad para la tendencia $\Delta\mu(S_{t+1})$.

$$\Delta\mu(S_{t+1}) = \mu_j'(S_{t+1}) - \mu_j(S_t) \tag{26}$$

$$\mu_j''(S_{t+1}) = \mu_j(S_t) + \frac{1}{I_D(\mu)} \Delta\mu(S_{t+1}) \tag{27}$$

Los valores de la estimación de los grados de pertenencia $\mu_j''(S_{t+1})$ son la información que determina el estado funcional estimado. El valor más alto de $\mu_j''(S_{t+1})$ corresponde al estado funcional que tiene mayor posibilidad de ocurrencia. Es importante tener en cuenta que también se tiene la información de la posibilidad de transición a los otros estados, lo que es una herramienta muy útil para tomar decisiones en un proceso.

LA PROPUESTA DE ESTIMACIÓN DE GRADOS DE PERTENENCIA EN UN SISTEMA DE MONITOREO

En la Fig. 2 se presenta el esquema de monitoreo que incluye el nuevo algoritmo de estimación propuesto.

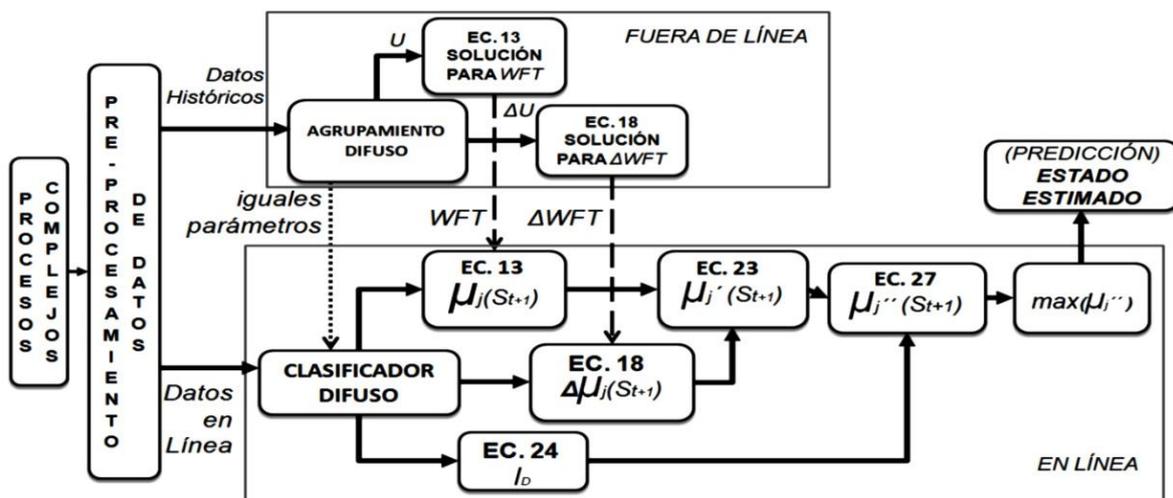


Fig. 2: Esquema de monitoreo incluyendo la propuesta de estimación

El esquema está compuesto por una etapa de entrenamiento (fuera de línea) y otra de monitoreo (en línea). Fuera de línea -usando los datos históricos del proceso- el clasificador difuso es entrenado y, con la clasificación difusa correspondiente, se construye el estimador (matrices de transición WFT y ΔWFT). Cada muestra tomada en línea (tiempo t) es el argumento de entrada al clasificador para obtener los grados de pertenencia difusa. Ellos son la información de entrada a los estimadores. Así, el resultado de la estimación final es la información que se presenta al operador del proceso.

Con el objetivo de mostrar detalles en la forma cómo opera la metodología propuesta, se presenta un ejemplo sencillo. Se consideran las variables de la Tabla 1 como datos de entrenamiento (muestras 1 a 6).

Entrenamiento fuera de línea

Las muestras (histórico de datos), los correspondientes grados de pertenencia (matriz U) encontrados con el clasificador entrenado (parte inicial en la etapa de entrenamiento) y los cambios en los grados de pertenencia (matriz ΔU , ec. (18)) son presentados en la Tabla 1.

Tabla 1: Un ejemplo sencillo

Muestras	Variables		Grados de pertenencia difusa, Matriz U			Cambios en Grados de pertenencia difusa, Matriz ΔU		
	x ₁	x ₂	μ ₁	μ ₂	μ ₃	Δμ ₁	Δμ ₂	Δμ ₃
1	1	1	0.571	0.302	0.259	--	--	--
2	1	2	0.646	0.423	0.396	0.075	0.121	0.137
3	2	2.5	0.573	0.743	0.510	-0.073	0.320	0.114
4	2	3	0.416	0.743	0.468	-0.157	0.000	-0.042
5	3	3	0.298	0.423	0.642	-0.118	-0.320	0.174
6	3	2	0.455	0.423	0.684	0.157	0.000	0.042
1P	1.2	1.5	0.713	0.392	0.333	--	--	--
2P	1.8	2.2	0.567	0.678	0.518	-0.146	0.286	0.185

Usando la información contenida en U y ΔU se construyen las ecuaciones (13) y (19) que se solucionan para WFT y ΔWFT respectivamente (ver ec. (28), (29) y (30)).

$$\begin{bmatrix} 0.646 \\ 0.573 \\ 0.416 \\ 0.298 \\ 0.455 \\ 0.423 \\ 0.743 \\ 0.743 \\ 0.423 \\ 0.423 \\ 0.396 \\ 0.510 \\ 0.468 \\ 0.642 \\ 0.684 \end{bmatrix} = \begin{bmatrix} 0.571 & 0 & 0 & 0.302 & 0 & 0 & 0.259 & 0 & 0 \\ 0.646 & 0 & 0 & 0.423 & 0 & 0 & 0.396 & 0 & 0 \\ 0.573 & 0 & 0 & 0.743 & 0 & 0 & 0.510 & 0 & 0 \\ 0.416 & 0 & 0 & 0.743 & 0 & 0 & 0.468 & 0 & 0 \\ 0.298 & 0 & 0 & 0.423 & 0 & 0 & 0.642 & 0 & 0 \\ 0 & 0.571 & 0 & 0 & 0.302 & 0 & 0 & 0.259 & 0 \\ 0 & 0.646 & 0 & 0 & 0.423 & 0 & 0 & 0.396 & 0 \\ 0 & 0.573 & 0 & 0 & 0.743 & 0 & 0 & 0.510 & 0 \\ 0 & 0.416 & 0 & 0 & 0.743 & 0 & 0 & 0.468 & 0 \\ 0 & 0.298 & 0 & 0 & 0.423 & 0 & 0 & 0.642 & 0 \\ 0 & 0 & 0.571 & 0 & 0 & 0.302 & 0 & 0 & 0.259 \\ 0 & 0 & 0.646 & 0 & 0 & 0.423 & 0 & 0 & 0.396 \\ 0 & 0 & 0.573 & 0 & 0 & 0.743 & 0 & 0 & 0.510 \\ 0 & 0 & 0.416 & 0 & 0 & 0.743 & 0 & 0 & 0.468 \\ 0 & 0 & 0.298 & 0 & 0 & 0.423 & 0 & 0 & 0.642 \end{bmatrix} \begin{bmatrix} WFT_{11} \\ WFT_{12} \\ WFT_{13} \\ WFT_{21} \\ WFT_{22} \\ WFT_{23} \\ WFT_{31} \\ WFT_{32} \\ WFT_{33} \end{bmatrix} \tag{28}$$

$$\begin{bmatrix} -0.073 \\ -0.157 \\ -0.118 \\ 0.157 \\ 0.320 \\ 0.000 \\ -0.320 \\ 0.000 \\ 0.114 \\ -0.042 \\ 0.174 \\ 0.042 \end{bmatrix} = \begin{bmatrix} 0.075 & 0 & 0 & 0.121 & 0 & 0 & 0.137 & 0 & 0 \\ -0.073 & 0 & 0 & 0.320 & 0 & 0 & 0.114 & 0 & 0 \\ -0.157 & 0 & 0 & 0.000 & 0 & 0 & -0.042 & 0 & 0 \\ -0.118 & 0 & 0 & -0.320 & 0 & 0 & 0.174 & 0 & 0 \\ 0 & 0.075 & 0 & 0 & 0.121 & 0 & 0 & 0.137 & 0 \\ 0 & -0.073 & 0 & 0 & 0.320 & 0 & 0 & 0.114 & 0 \\ 0 & -0.157 & 0 & 0 & 0.000 & 0 & 0 & -0.042 & 0 \\ 0 & -0.118 & 0 & 0 & -0.320 & 0 & 0 & 0.174 & 0 \\ 0 & 0 & 0.075 & 0 & 0 & 0.121 & 0 & 0 & 0.137 \\ 0 & 0 & -0.073 & 0 & 0 & 0.320 & 0 & 0 & 0.114 \\ 0 & 0 & -0.157 & 0 & 0 & 0.000 & 0 & 0 & -0.042 \\ 0 & 0 & -0.118 & 0 & 0 & -0.320 & 0 & 0 & 0.174 \\ 0 & 0 & -0.118 & 0 & 0 & -0.320 & 0 & 0 & 0.174 \end{bmatrix} \begin{bmatrix} \Delta WFT_{11} \\ \Delta WFT_{12} \\ \Delta WFT_{13} \\ \Delta WFT_{21} \\ \Delta WFT_{22} \\ \Delta WFT_{23} \\ \Delta WFT_{31} \\ \Delta WFT_{32} \\ \Delta WFT_{33} \end{bmatrix} \tag{29}$$

$$\begin{bmatrix} WFT_{11} \\ WFT_{12} \\ WFT_{13} \\ WFT_{21} \\ WFT_{22} \\ WFT_{23} \\ WFT_{31} \\ WFT_{32} \\ WFT_{33} \end{bmatrix} = 0.0610 \begin{bmatrix} 0.8102 \\ 0.8043 \\ 0.1740 \\ 0 \\ 0.0610 \\ 0.0069 \\ 0.1449 \\ 0.2737 \\ -0.9739 \end{bmatrix} ; \begin{bmatrix} \Delta WFT_{11} \\ \Delta WFT_{12} \\ \Delta WFT_{13} \\ \Delta WFT_{21} \\ \Delta WFT_{22} \\ \Delta WFT_{23} \\ \Delta WFT_{31} \\ \Delta WFT_{32} \\ \Delta WFT_{33} \end{bmatrix} = \begin{bmatrix} 0.1427 \\ 1.7945 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.0956 \\ 1.2307 \\ 0.1695 \end{bmatrix} \tag{30}$$

El clasificador difuso entrenado y los vectores WFT (Pesos de Transición Difusa) y ΔWFT (Delta de Pesos de Transición Difusa) son los elementos que van a permitir realizar el monitoreo en línea del proceso.

Monitoreo en línea

En el método propuesto, una estimación puede ser obtenida cuando han sido clasificadas la muestra actual (S_t) y la muestra anterior (S_{t-1}). Entonces, para la estimación de los grados de pertenencia para la muestra S₃ de prueba, se utilizan las dos muestras de prueba (1P y 2P) ubicadas al final de la Tabla 1 para ejemplificar el procedimiento.

Como dato en línea se toma del proceso la muestra S_{2P} ($x_{1P}=1.8$, $x_{2P}=2.2$) con la clasificación correspondiente ($\mu_{1P}=0.567$, $\mu_{2P}=0.678$ y $\mu_{3P}=0.518$), y son calculadas las diferencias en los grados de pertenencia (18) para las muestras S_{2P} y S_{1P} ($\Delta\mu_{1P}=-0.146$, $\Delta\mu_{2P}=0.286$ y $\Delta\mu_{3P}=0.185$). Conocidos WFT , ΔWFT , μ y $\Delta\mu$ (para S_{2P}) se construyen y resuelven las ecuaciones (13) y (19) como se observa en (31) y (32) respectivamente.

$$\begin{bmatrix} \mu_1(S_{2P}) \\ \mu_2(S_{2P}) \\ \mu_3(S_{2P}) \end{bmatrix} = \begin{bmatrix} 0.567 & 0 & 0 & 0.678 & 0 & 0 & 0.518 & 0 & 0 \\ 0 & 0.567 & 0 & 0 & 0.678 & 0 & 0 & 0.518 & 0 \\ 0 & 0 & 0.567 & 0 & 0 & 0.678 & 0 & 0 & 0.518 \end{bmatrix} \begin{bmatrix} WFT_{11} \\ WFT_{12} \\ WFT_{13} \\ WFT_{21} \\ WFT_{22} \\ WFT_{23} \\ WFT_{31} \\ WFT_{32} \\ WFT_{33} \end{bmatrix} = \begin{bmatrix} 0.5344 \\ 0.6392 \\ 0.6078 \end{bmatrix} \quad (31)$$

$$\begin{bmatrix} \Delta\mu_1(S_{2P}) \\ \Delta\mu_2(S_{2P}) \\ \Delta\mu_3(S_{2P}) \end{bmatrix} = \begin{bmatrix} -0.146 & 0 & 0 & 0.286 & 0 & 0 & 0.185 & 0 & 0 \\ 0 & -0.146 & 0 & 0 & 0.286 & 0 & 0 & 0.185 & 0 \\ 0 & 0 & -0.146 & 0 & 0 & 0.286 & 0 & 0 & 0.185 \end{bmatrix} \begin{bmatrix} \Delta WFT_{11} \\ \Delta WFT_{12} \\ \Delta WFT_{13} \\ \Delta WFT_{21} \\ \Delta WFT_{22} \\ \Delta WFT_{23} \\ \Delta WFT_{31} \\ \Delta WFT_{32} \\ \Delta WFT_{33} \end{bmatrix} = \begin{bmatrix} -0.0031 \\ -0.0343 \\ 0.0314 \end{bmatrix} \quad (32)$$

En (33) se observa el resultado de la estimación inicial de acuerdo con (23).

$$\begin{bmatrix} \mu'_1(S_{2P}) \\ \mu'_2(S_{2P}) \\ \mu'_3(S_{2P}) \end{bmatrix} = \begin{bmatrix} 0.5344 \\ 0.6392 \\ 0.6078 \end{bmatrix} + \begin{bmatrix} -0.0031 \\ -0.0343 \\ 0.0314 \end{bmatrix} = \begin{bmatrix} 0.5313 \\ 0.6049 \\ 0.6392 \end{bmatrix} \quad (33)$$

El valor de I_D (ec. (24) y ec. (25)) es calculado en (34) con los grados de pertenencia de la muestra S_{2P} .

$$\mu_M = \mu_{2P} = 0.678, \begin{bmatrix} \lambda_1 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0.111 \\ 0.160 \end{bmatrix}, I_D = \frac{0.111 * e^{0.111} + 0.160 * e^{0.160}}{2 * 0.678 * e^{0.678}} = 0.1167 \quad (34)$$

El valor de cambio $\Lambda\mu(S_{2P})$ (26) evaluado como la diferencia entre la estimación inicial y los grados de pertenencia de S_{2P} es calculado en (35).

$$\begin{bmatrix} \Lambda\mu_1(S_{2P}) \\ \Lambda\mu_2(S_{2P}) \\ \Lambda\mu_3(S_{2P}) \end{bmatrix} = \begin{bmatrix} 0.5313 \\ 0.6049 \\ 0.6392 \end{bmatrix} - \begin{bmatrix} 0.567 \\ 0.678 \\ 0.518 \end{bmatrix} = \begin{bmatrix} -0.0357 \\ -0.0731 \\ 0.1212 \end{bmatrix} \quad (35)$$

La estimación final de los grados de pertenencia $\mu_j''(S_{2P})$ (27) es evaluada en (36).

$$\begin{bmatrix} \mu''_1(S_{2P}) \\ \mu''_2(S_{2P}) \\ \mu''_3(S_{2P}) \end{bmatrix} = \begin{bmatrix} 0.567 \\ 0.678 \\ 0.518 \end{bmatrix} + \frac{1}{0.1167} \begin{bmatrix} -0.0357 \\ -0.0731 \\ 0.1212 \end{bmatrix} = \begin{bmatrix} 0.2611 \\ 0.0516 \\ 1.5566 \end{bmatrix} \quad (36)$$

El estado final estimado para la muestra S_{3P} es determinado por el máximo valor de pertenencia estimado para S_{3P} , que en (36) corresponde al grado de pertenencia asociado a la clase $j=3$.

CASOS DE ESTUDIO

En el artículo se presentan dos casos estudio: un sistema de monitoreo térmico en un conductor eléctrico - con el que se expone de manera detallada la aplicación de la metodología propuesta- y un sistema de monitoreo en una caldera (Kempowsky et al., 2006). Estos dos casos permiten hacer un análisis del desempeño del estimador en el esquema de monitoreo.

Monitoreo térmico en línea de transporte de energía: Descripción del proceso

Una fibra óptica en la cual el índice de refracción del núcleo se modifica de forma periódica en la dirección de propagación de la luz es una red de Bragg en fibra óptica (FBG). Esta característica hace que la FBG sea un dispositivo selectivo que refleja una longitud de onda específica conocida como longitud de onda de Bragg. En las FBG existe una relación o respuesta lineal de corrimiento espectral con los cambios de temperatura. Además, las FBG presentan reducido tamaño y peso, inmunidad electromagnética y la posibilidad de multiplexación y de sensado remoto. Estas características facilitan la configuración de un sensado quasi-distribuido de temperatura en líneas de transmisión eléctrica (Udd y Spillman, 2011). El tramo de línea de transmisión de energía (sistema de prueba) con los herrajes y aisladores requeridos fue instalado en el interior del laboratorio de Alta Tensión de la Facultad de Ingeniería de la Universidad de Antioquia. El conductor ASCR 1/0 de 8 metros fue conectado a un transformador que se alimentó con una fuente eléctrica variable, para permitir el control del voltaje aplicado. Un valor ajustable de corriente recorre el circuito formado por el arrollamiento secundario del transformador y el conductor. Sobre el conductor se instalaron 3 redes de Bragg grabadas sobre la misma fibra óptica a 2, 4 y 6 metros medidos desde uno de los extremos. Las longitudes de onda λ_B (a 25°C) correspondientes a los sensores 1, 2 y 3 respectivamente fueron: 1554.11 nm, 1556.14 nm y 1558.18 nm. En uno de los extremos de la fibra fue instalado un interrogador de Redes de Bragg basado en un láser sintonizable. En la Fig. 3 se muestra el sistema de monitoreo. Los registros de temperatura fueron obtenidos a través de dispositivos y equipos con tecnología óptica basada en Redes de Bragg.

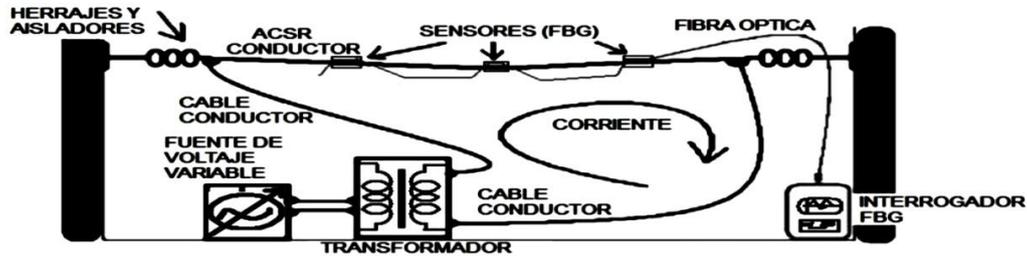


Fig. 3: Esquema del sistema de monitoreo térmico con tecnología óptica

El conductor fue sometido a calentamiento por el aumento controlado de corriente y por fuentes externas de temperatura (llama con soplete de butano). Los perfiles térmicos obtenidos que fueron usados para la etapa de entrenamiento y prueba del sistema de monitoreo se presentan en las Fig. 4 y 5 respectivamente. El eje 'x' corresponde al muestreo de la señal (1 muestra/seg). Las figuras en el eje 'y' muestran los valores de variación de longitud de onda (lectura del interrogador de fibra) para cada sensor. La temperatura real se puede calcular en función del valor de la variación de longitud de onda (WL), ver ec. (37). Las temperaturas oscilaron entre 30°(temperatura ambiente) y 110° Celsius.

$$T(^{\circ}C) = 102.629 * (WL) + 30 \tag{37}$$

Las bases de datos de entrenamiento y de prueba contienen 9339 y 5699 muestras respectivamente. Cada muestra tiene 3 variables ($S1=$ Variación en Longitud de onda sensor 1, $S2=$ Variación en Longitud de onda sensor 2, y $S3=$ Variación en Longitud de onda sensor 3), una por cada FBG. Las 7 clases por identificar, definidas en la Tabla 2, corresponden al comportamiento de la temperatura en el conductor por efecto de la circulación de corriente eléctrica (comportamiento normal del sistema) y fuentes externas de temperatura (no deseadas).

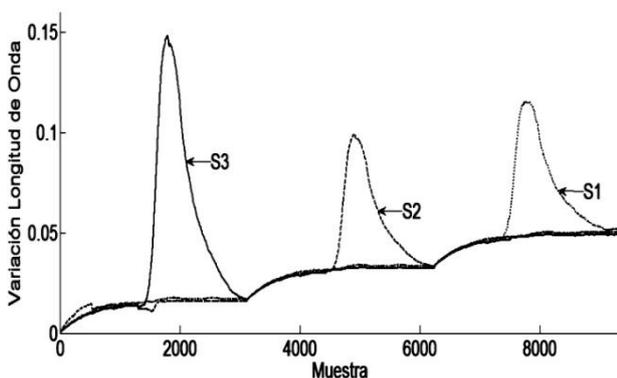


Fig. 4: Temperatura FBG (Datos Entrenamiento)

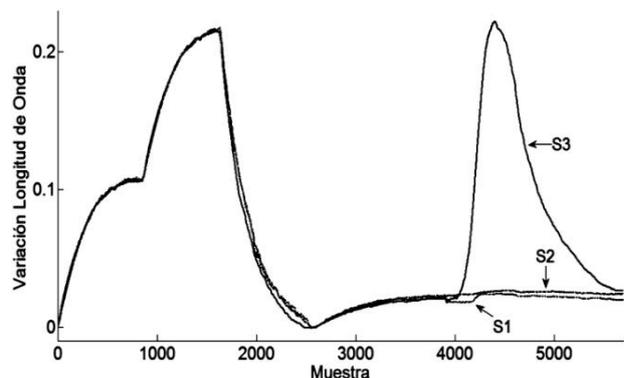


Fig. 5: Temperatura FBG (Datos Prueba)

Pre tratamiento de las señales

Antes de aplicar el método de agrupamiento, se realizó un pre tratamiento a las señales seleccionadas de temperatura. La primera parte del pre tratamiento de las señales corresponde a: a) el filtraje de la señal para eliminar ruidos de alta frecuencia, b) el submuestreo debido a la gran cantidad de datos (se registra un dato cada 10 muestras) y, finalmente, c) la eliminación de la inercia térmica en los tres sensores por la circulación de la misma corriente (el valor absoluto de la diferencia entre señales fue la acción tomada: $|S1-S2|$ denominado $S12$, $|S1-S3|$ denominado $S13$ y $|S2-S3|$ denominado $S23$). La segunda parte del pre tratamiento tiene el objetivo de integrar información que permita diferenciar los tiempos de presencia de falla y recuperación después de fallo. El valor de la pendiente calculado en una ventana de 10 muestras para cada señal (base original de entrenamiento) generó tres nuevas variables ($IS1$, $IS2$ e $IS3$). Las señales resultantes del pre tratamiento (normalizadas de acuerdo con (38)) se presentan en las Fig. 6 y 7.

$$x_{in} = (x_i - x_{min}) / (x_{max} - x_{min}) \tag{38}$$

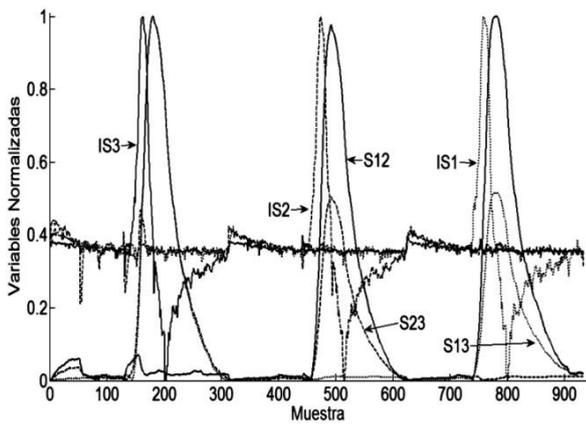


Fig. 6: Base de entrenamiento con pre tratamiento.

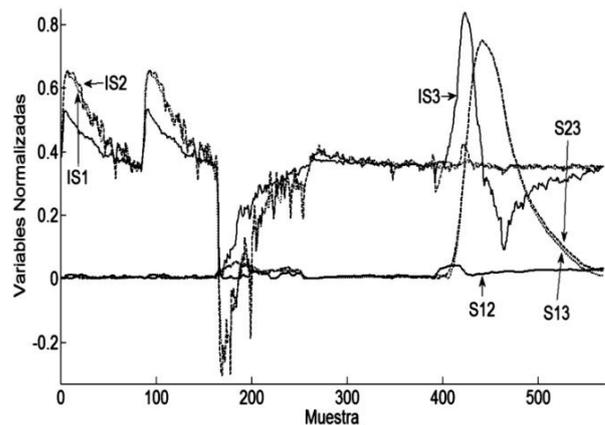


Fig. 7: Base de prueba con pre tratamiento.

Entrenamiento del clasificador difuso: Entrenamiento fuera de línea

El clasificador es entrenado con el algoritmo LAMDA en un esquema supervisado (reconocimiento), i.e. el número de clases es establecido a priori. En el proceso los estados funcionales normal y de falla están presentes. Dentro del estado funcional de falla, la presencia y el aislamiento de la falla deben ser diferenciados. La asociación de las 7 clases a estados funcionales se presenta en la Tabla 2.

Tabla 2. Línea de Energía: Clases Vs Estados Funcionales

Clase	Estado Funcional
C1	Operación Normal: Variación uniforme de temperatura en los 3 sensores
C2	Falla: Incremento de temperatura, fuente externa activa cerca al sensor 3
C3	Recuperación: Descenso de temperatura, fuente externa inactiva cerca al sensor 3
C4	Falla: Incremento de Temperatura, fuente externa activa cerca al sensor 2
C5	Recuperación: Descenso de temperatura, fuente externa inactiva cerca al sensor 2
C6	Falla: Incremento de temperatura, fuente externa activa cerca al sensor 1
C7	Recuperación: Descenso de temperatura, fuente externa inactiva cerca al sensor 1

Los resultados obtenidos con el clasificador en la etapa de entrenamiento se presentan en la Fig. 8a, la línea continua representa la evolución de las clases preestablecidas (clases de referencia) y la línea segmentada la evolución de las clases reconocidas por el clasificador.

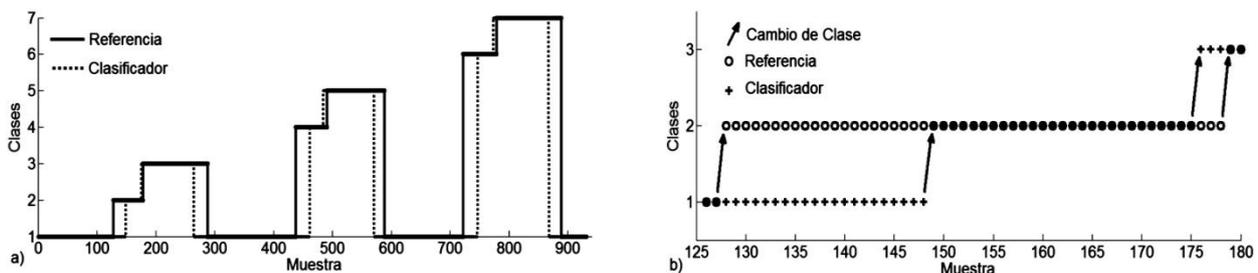


Fig. 8: Clasificación obtenida con LAMDA. a) Clasificación para base de datos completa, b) Ampliación entre muestras 125 y 180.

Todas las clases son bien reconocidas, pero en los intervalos de transición hay diferencias en la identificación de los cambios de estado. En la Fig. 8b la primera transición de C1 a C2 según el clasificador está retrasada con respecto al cambio en las clases de referencia; pero en la transición de C2 a C3 el clasificador identifica el cambio de clase con anterioridad. La observación y análisis del comportamiento natural de las señales tomadas del proceso permitieron establecer que las variaciones registradas no son significativas hasta después de transcurrir varios muestreos a partir del momento en que se produce una perturbación. Debido a la naturaleza del método de agrupamiento –en su relación dimensional– no es posible que el clasificador reconozca un cambio de clase si no hay cambio considerable en la ubicación espacial de la muestra; es decir, se requiere un cambio significativo en los valores de las variables medidas.

Cálculo del estimador para grados de pertenencia y para cambios de grados de pertenencia: Entrenamiento

Con la información de los grados de pertenencia de los datos históricos (resultado de aplicar un método de agrupamiento difuso, en este caso LAMDA), se resuelven las ecuaciones (13) y (19) para WFT y ΔWFT . Una solución tanto para WFT como para ΔWFT se muestra en (39) y (40) respectivamente.

$$[WFT] = \begin{bmatrix} 0.9949 & 0 & 0 & 0.0087 & 0 & 0.0008 & 0 \\ 0 & 0.94 & 0 & 0 & 0 & 0 & 0 \\ 0.0028 & 0.0470 & 1 & 0.0047 & 0.0024 & 0.0062 & 0.0026 \\ 0 & 0 & 0 & 0.9219 & 0 & 0 & 0 \\ 0.0022 & 0.0080 & 0 & 0.0051 & 0.9978 & 0.0079 & 0.0001 \\ 0 & 0 & 0 & 0 & 0 & 0.9283 & 0 \\ 0.0014 & 0.0039 & 0 & 0.0091 & 0 & 0.0561 & 0.996 \end{bmatrix} \quad (39)$$

$$[\Delta WFT] = \begin{bmatrix} 0.6049 & 0 & 0.0701 & 0 & 0 & 0 & 0.0009 \\ 0 & 0.5434 & 0.2325 & 0.0132 & 0.0025 & 0.0049 & 0.0767 \\ 0 & 0 & 0.0723 & 0 & 0 & 0 & 0 \\ 0.3290 & 0.1041 & 0.1745 & 0.6731 & 0.3880 & 0.2693 & 0.311 \\ 0 & 0 & 0 & 0 & 0.0316 & 0 & 0 \\ 0.1245 & 0.0853 & 0.0661 & 0.25400 & 0.4016 & 0.5946 & 0.5020 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (40)$$

Monitoreo del proceso en línea

La etapa de monitoreo tiene como base las matrices WFT (39) y ΔWFT (40) (calculadas una sola vez en entrenamiento) y el resultado de la clasificación difusa $\mu_j(S_{t-1})$ y $\mu_j(S_t)$. Estos grados de pertenencia difusos corresponden a los grados registrados en la muestra anterior y en la muestra actual respectivamente. Con la información anterior y los pasos descritos en la metodología, se procede hasta evaluar la ecuación (26) para obtener la estimación final de los grados de pertenencia difusos $\mu_j^*(S_{t+1})$. El algoritmo de estimación en el esquema de monitoreo es validado con la base de datos de entrenamiento y con la base de datos de prueba.

Prueba de monitoreo con base de datos de entrenamiento y prueba: Monitoreo en línea

Los estados funcionales (clases) para la base de entrenamiento y de prueba, asignados de acuerdo con los máximos grados de pertenencia estimados, son presentados en las Fig. 9a y 10a respectivamente. En las Fig. 9 y 10 la línea continua representa la evolución de los estados de referencia; la línea segmentada, la evolución de los estados según el clasificador difuso; y la línea segmentada con puntos, la evolución de las estimaciones.

Análisis de resultados

Se observa en la Fig. 9a –donde se prueba el estimador con los datos de entrenamiento– que el clasificador difuso entrenado (línea segmentada) es capaz de identificar los estados presentes, pero en las transiciones presenta retrasos con respecto a la clase de referencia. También se observa que el predictor de estados (línea segmentada con puntos) es capaz de identificar los estados presentes y que en todas las transiciones se adelanta al clasificador como mínimo en 3 muestras (ver detalles en Fig. 9b). En la Fig. 10a, donde se presentan los resultados para toda la base de datos de prueba, el estimador cumple satisfactoriamente la tarea de predicción adelantándose en todos los casos a los estados identificados con el clasificador (ver detalles en Fig. 10b). El desempeño del predictor es adecuado, ya que las variaciones en las señales del proceso –debido a las perturbaciones– producen variaciones graduales en los grados de pertenencia obtenidos por el clasificador, y estas tendencias son aprendidas por el estimador. Una vez se inicia el fallo en la muestra 128 -fallo que afecta muy debilmente y en forma gradual el proceso- el clasificador lo detecta solo hasta la muestra 149, pero la predicción estima acertadamente el cambio de estado desde la muestra 146.

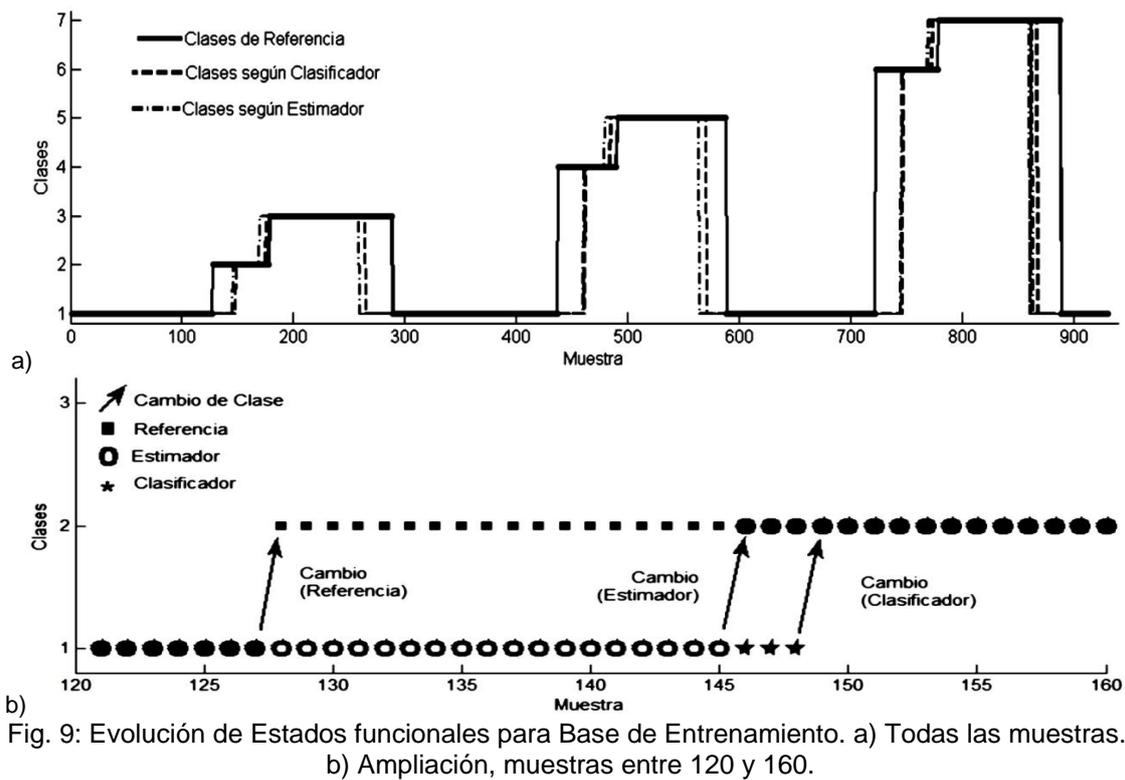


Fig. 9: Evolución de Estados funcionales para Base de Entrenamiento. a) Todas las muestras. b) Ampliación, muestras entre 120 y 160.

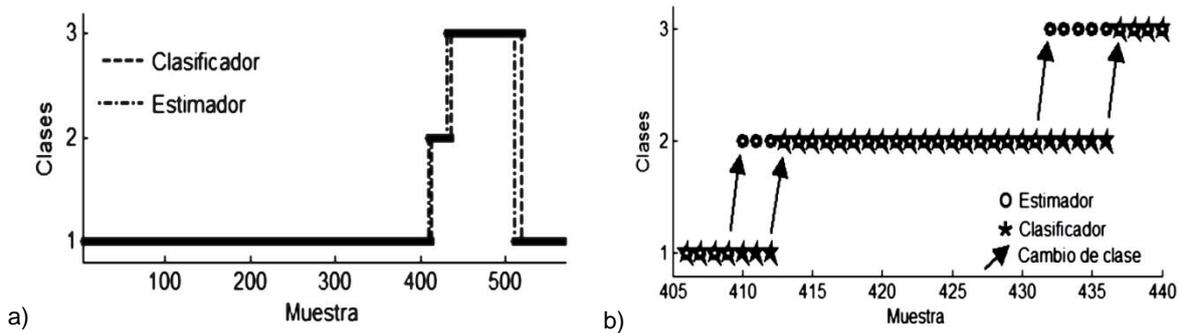


Fig. 10: Evolución de Estados Base de Prueba. a) Todas las muestras. b) Muestras 406 a 450.

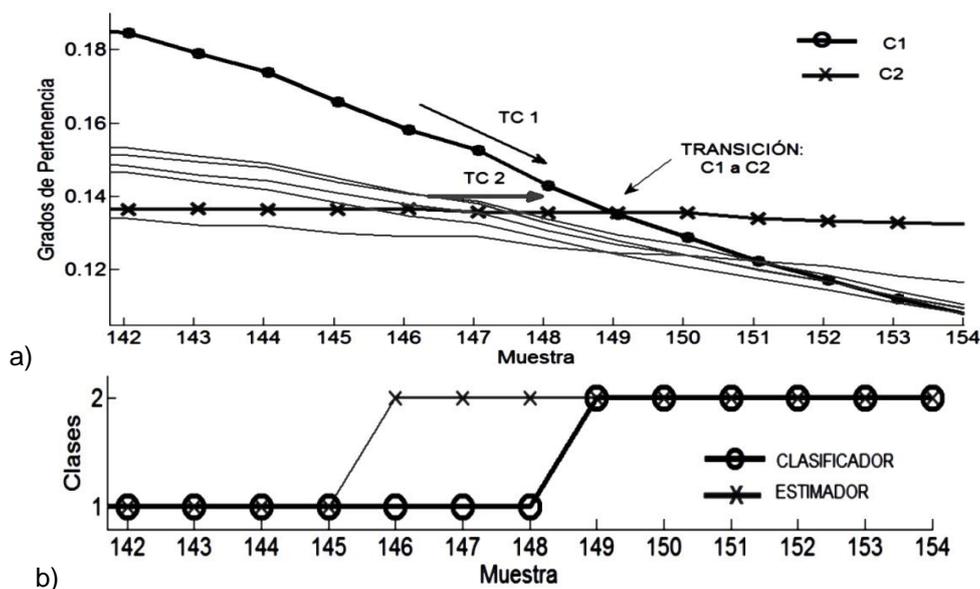


Fig. 11: Resultados estimación para entrenamiento (muestras 142 a 154). a) Grados de pertenencia difusos obtenidos por el clasificador, b) Clases según Clasificador y Estimador

En la Fig. 11a se presentan los grados de pertenencia obtenidos con el clasificador difuso para las muestras 142 a 154, y en la Fig. 11b, las clases obtenidas tanto por el clasificador como por el estimador. En la Fig. 11b el clasificador difuso registra el cambio de la clase C1 a la clase C2 a partir de la muestra 149 (ver línea continua con círculos negros) mientras que el estimador predice el cambio de clase a partir de la muestra 146 (ver línea continua gris con 'x' grises). El estimador –con base en la tendencia TC1 (de disminución) y de la tendencia TC2 (constante)– predice valores de grados de pertenencia en la clase C1 más bajos que los valores de la clase C2, por lo cual se obtiene una estimación acertada con 4 muestras de anticipación (ver Fig. 11a).

Caso de estudio: Subsistema de Caldera

El segundo caso de estudio corresponde al subsistema de caldera de un generador a vapor (Kempowsky, 2006). El generador de vapor de prueba fue diseñado como una versión a escala de un generador de vapor real para una planta nuclear. La operación del proceso es la siguiente: el flujo de agua de alimentación es generado por una bomba que impulsa el agua a la caldera. Para mantener constante el nivel de agua en la caldera, un controlador On–Off opera en la vía de la bomba. El valor de la potencia de calentamiento de la caldera dependerá de la presión del acumulador de vapor. Cuando la presión del acumulador cae por debajo de un valor mínimo la resistencia de calentamiento que da la máxima potencia de calefacción es activada, y cuando alcanza la presión máxima la resistencia de calentamiento es desactivada para mantener la presión en el punto de referencia. El histórico de datos contiene 937 muestras de entrenamiento, 1800 muestras de prueba, y 5 descriptores correspondientes a las variables físicas: el flujo de agua de alimentación, potencia de calefacción, presión de caldera, nivel de caldera y flujo de vapor de salida. Con el objetivo de homogenizar la influencia de las dimensiones propias de cada variable, los datos son normalizados con respecto al valor máximo y mínimo de cada variable.

Los grados de pertenencia difusos (matriz U) para cada una de las muestras de la base de datos histórica fueron obtenidos aplicando el algoritmo LAMDA. Cinco clases fueron identificadas y cada muestra fue clasificada de acuerdo a su máximo grado de pertenencia. En la Tabla 3 las 5 clases fueron asociadas a estados funcionales después de una validación previa por el experto en el proceso (100% de desempeño del agrupamiento).

Tabla 3: Subsistema de Caldera: Clases vs Estados Funcionales.

Estado Funcional	Operación Normal	Regulación: Presión	Regulación: Nivel	Regulación: Nivel y Presión
Clase	C1	C2 - C3	C4	C5

Con base en la matriz U , las matrices WFT y ΔWFT son obtenidas solucionando las ecuaciones (13) y (19). Y, aplicando los pasos del monitoreo en línea, se prueba el estimador para la base de entrenamiento y para la base de prueba (ver Fig. 12 y 13 respectivamente). Se puede observar en la Fig. 12a, correspondiente a la base de entrenamiento, que se predicen correctamente las clases. Además, es importante resaltar que la estimación evita en algunos casos, la oscilación que presenta el clasificador entre las clases C2 y C3 (ver detalle en la Fig. 12b). Las clases C2 y C3 corresponden a un mismo estado de regulación de presión de acuerdo con la tabla 3. La clase C3 se diferencia de la clase C2, en que los valores de sus variables están cercanos a los límites que determinan el siguiente estado funcional.

En la Fig. 13a, que corresponde a la predicción para la base de prueba, se puede observar que el estimador conserva la característica deseada de eliminar oscilaciones entre las clases C2 y C3 –la clase C2 y C3 corresponden a un mismo estado– (ver detalle en Fig. 13b). También es importante resaltar que se predice acertadamente el cambio de la clase C2 a C5; transición que no estuvo presente en la base de datos de entrenamiento con la cual se obtuvo el predictor (ver detalle en Fig. 13c).

Análisis de resultados

Este caso de estudio presenta una diferencia con respecto al tiempo de respuesta del sistema de temperatura. En el proceso de la caldera, los cambios son más rápidos. En el caso del monitoreo de temperatura, una vez que se presenta la perturbación sobre el sistema, las variables del sistema varían lentamente, y es posible encontrar una tendencia de comportamiento antes de presentarse un cambio de clase. Para el caso de la caldera los cambios son drásticos (cambios de clase en las muestras 8 y 9, siguientes a la presencia de la perturbación) como se puede observar en la Fig. 14.

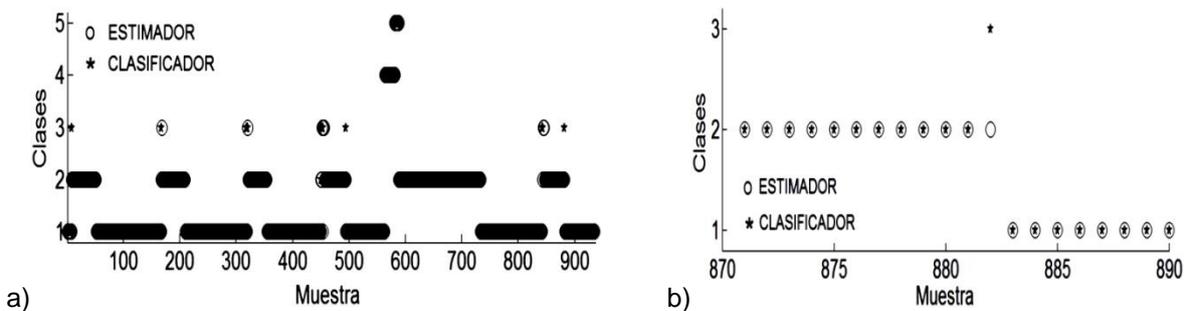


Fig. 12: Evolución de las clases (estados funcionales) según clasificador y estimador: a) Base de entrenamiento completa, b) ampliación entre muestras 870 a 890

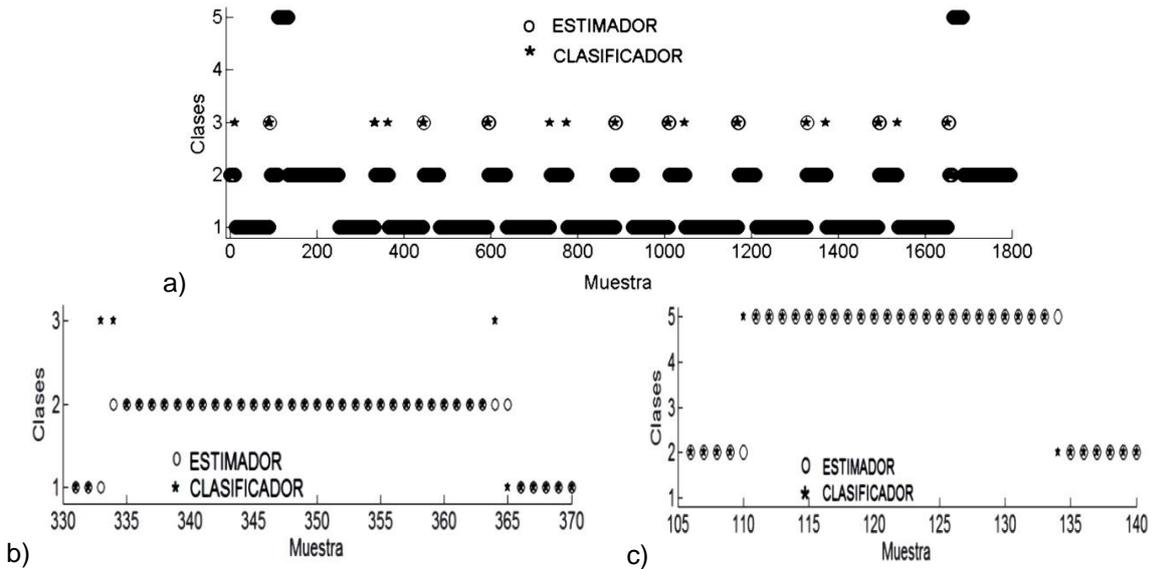


Fig. 13: Evolución de las Clases (Estados) según clasificador y estimador para prueba. a) Todas las muestras, b) Ampliación entre muestras 330 y 370, c) Ampliación entre muestras 105 y 140

En la Fig. 14a –donde se hace una ampliación sobre la grafica de grados de pertenencia difusa obtenidos para el histórico de datos (entre muestras 4 y 12) – se puede observar que inmediatamente se presenta la perturbación sobre el sistema (muestra 7 para este caso) los grados de pertenencia cambian fuertemente (grados de la clase C1 y grados de la clase C3 en las muestras 7 y 8) generando un cambio de clase. Este rápido cambio en los grados de pertenencia está asociado al rápido cambio de los valores de las variables del proceso debido a alguna perturbación. Esta dinámica no tiene una tendencia de cambio; por lo cual, el predictor puede como máximo hacer una estimación igual a la clase obtenida por el clasificador eliminando transiciones oscilatorias (C3).

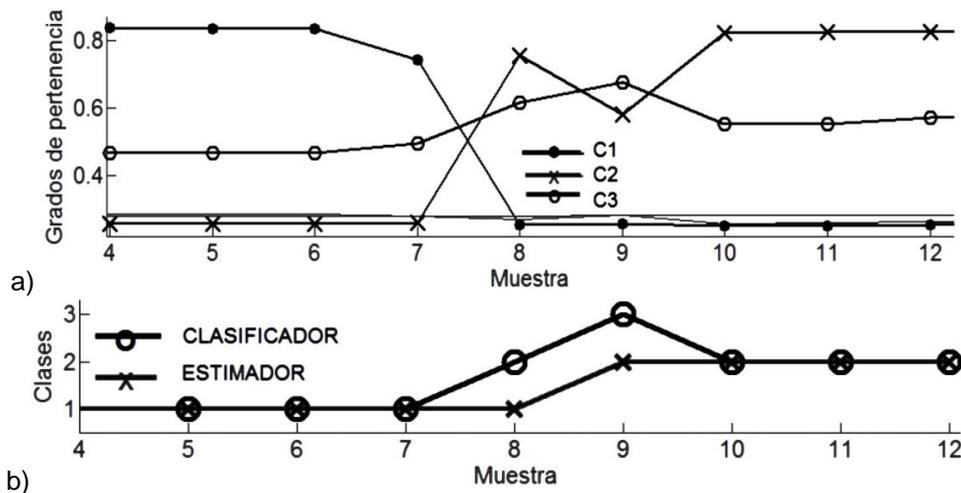


Fig. 14: Resultados de estimación para base de prueba (muestras 4 a 12). a) Grados de pertenencia difusos obtenidos por el clasificador, b) Clases según clasificador y estimador

En la Fig. 14b, las clases obtenidas por el clasificador (línea continua con círculos) y las clases obtenidas por el predictor (línea continua con 'x') son comparadas. La perturbación que se presenta después de registrada la muestra 7 produce el cambio drástico de los grados de pertenencia como se registra en la muestra 8. Por lo anterior, la predicción para la muestra 8 sigue la tendencia de los valores para la muestra 7 y permanece en la clase C1. Posteriormente, con el registro de la muestra 8, el estimador predice la clase C2 para la muestra 9, y de nuevo la estimación de la muestra 10 con base en el registro de la muestra 9 es la clase C2. Estas estimaciones para las muestras 8, 9 y 10 han evitado la oscilación entre las clases C2 y C3 antes de que el proceso evolucione de la clase C1 a la clase C2.

CONCLUSIONES

En este artículo se presentó un método para predecir estados funcionales con base en información extraída del proceso por medio de clasificadores difusos.

La relación existente entre valores de probabilidad y valores de posibilidad permitió utilizar una estructura de estimación propia del campo de las probabilidades como son las cadenas de Markov, para la estimación de grados de pertenencia difusa.

La propuesta desarrollada para la estimación de estados permite utilizar cualquier método de agrupamiento difuso que suministre la información base (grados de pertenencia difusos) para la ejecución de la metodología.

Se integró la metodología de estimación de grados de pertenencia difusa a un esquema general de monitoreo de procesos basado en clasificadores.

La metodología propuesta fue probada con la información obtenida de dos procesos industriales reales, evaluándose como satisfactorio el desempeño de la predicción en el esquema de monitoreo.

Cuando un sistema tiene una respuesta lenta ante una perturbación, el predictor es capaz de estimar adecuadamente el estado siguiente. Si la respuesta del sistema es muy rápida, de tal manera que una vez se presenta una perturbación haya un cambio de estado, el predictor es capaz de seguir este comportamiento en forma conservativa (como mínimo manteniendo el estado anterior al registro de la falla).

AGRADECIMIENTOS

Este trabajo fue apoyado por el acuerdo de cooperación internacional COLCIENCIAS (COL) - ECOSNord (FRA), y el Comité para el Desarrollo de la Investigación -CODI- de la Universidad de Antioquia (COL) como entes financiadores.

REFERENCIAS

Aguado, D. y C. Rosen, *Multivariate statistical monitoring of continuous wastewater treatment plants* Engineering Applications of Artificial Intelligence, 21 (7), 1080-1091 (2008).

Aguilar-Martín, J. y R. Lopez De Mantaras, *The Process of Classification and Learning the Meaning of Linguistic Descriptors of Concepts*, Approximate reasoning in decision analysis, North H., 165-175 (1982).

Aguilar-Martín, J., *Inteligencia Artificial para La Supervisión De Procesos Industriales*, 1ª edición, ULA, Mérida, Venezuela (2007).

Bezdek, J.C., *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Publishing Corporation, New York, USA. (1981).

Botero, H. y H. Álvarez, Una Revisión de los métodos más frecuentes para la estimación del estado en procesos químicos, *Revista Dyna*: 76(158), 135-146 (2009).

Botía, J. y otros cuatro autores, *Automaton Based on Fuzzy Clustering Methods for Monitoring Industrial Processes*, J. Eng. Apl. Arti. Intell., [approved for publication] (2012).

Chang, C.T. y C.Y. Chen, *Fault diagnosis with automata generated languages*. Computers and Chemical Engineering: 35(2), 329-341 (2011).

Colomer Llinàs, J., J. Meléndez y J. Ayza, *Sistemas de supervisión*, Cuadernos CEA-IFAC, Cetisa-Boixareu, Barcelona, España (2000).

- Cuadrado Vega A.A., *Supervisión de procesos complejos mediante técnicas de Data Mining con Incorporación de Conocimiento Previo*, Tesis Doctoral, Departamento de Ingeniería Eléctrica, Electrónica, de Computadores y Sistemas, Universidad de Oviedo, Gijón, España (2002).
- Daidone, A., F.Di Giandomenico, A. Bondavalli y S. Chiaradonna, *Hidden Markov Models as a Support for Diagnosis: Formalization of the Problem and Synthesis of the Solution*. Proceedings of 25th IEEE Symposium on Reliable Distributed Systems (SRDS), 245–256, Leeds, UK, October (2006).
- De Silva, S., M.D. Junior, V.L. Junior y M.J. Brennan, *Structural Damage Detection by Fuzzy Clustering*, Mechanical Systems and Signal Processing, 22, 1636-1649 (2008).
- Du, R., y K. Yeung, *Fuzzy transition probability: a new method for monitoring progressive faults. Part 1. The theory*, Engineering Applications of Artificial Intelligence, 17, 457-467, (2004).
- Du, R., y K. Yeung, *Fuzzy transition probability: a new method for monitoring progressive faults. Part 2. Applications examples*, Engineering Applications of Artificial Intelligence, 19, 145-155 (2006).
- Dubois, D. y H. Prade, *Unfair Coins and Necessity Measures: Towards a Possibilistic Interpretation of Histograms*, Fuzzy Sets and Systems, 10, 15–20 (1983).
- Gustafson, D.E., y W.C. Kessel, *Fuzzy Clustering with a Fuzzy Covariance Matrix*, Proceedings of IEEE Conference on Decision and Control, 17, 761 -766 (1978).
- Hedjazi, L., J. Aguilar-Martin y M. Le Lann, *Similarity-margin based feature selection for symbolic interval data*, Pattern Recognition Letters, 32(4), 578-585 (2011).
- Hegyí, A., D. Girimonte, R. Babuška y B. De Schutter, *A comparison of filter configurations for freeway traffic state estimation*, IEEE Intelligent Transportation Systems Conference, Toronto, Canada, September 17-20 (2006).
- Infante, S., L. Sánchez y F. Cedeño, *Filtros para Predecir Incertidumbre de Lluvia y Clima*, Revista de Climatología: 12, 33-48 (2012).
- Isaza C. y otros cuatro autores, *Decision Method for Functional States Validation in a Drinking Water Plant*, 10th International Symposium on Computer Applications in Biotechnology, Cancún, Mexico, 4 a 6 junio (2007).
- Isermann, R., *Process fault detection based on modeling and estimation methods-A survey*, Automatica, 20 (4), 387-404 (1984).
- Isermann, R. y P. Ballé, *Trends in the Application of Model-Based Fault Detection and Diagnosis of Technical Processes*, Control Engineering Practice, 5(5), 709-719 (1997).
- Jyoti, K. y S. Singh, *Data Clustering Approach to Industrial Process Monitoring, Fault Detection and Isolation*, International Journal of Computer Applications (0975 – 8887), 17(2), 41-45, (2011).
- Kempowsky, T., A. Subias y J. Aguilar-Martin, *Process Situation Assessment: From a Fuzzy partition to a finite State Machine*, Engineering Applications of Artificial Intelligence, 19, 461–477 (2006).
- Klir, G.J. y B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, Upper Saddle River, New Jersey, USA (1995).
- Nguyen N.V., V. Shin, G. Shevlyakov, *Power System State Estimation with Fusion Method, The 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore ,26-28 Feb (2010)*.
- Omran, M., A. P. Engelbrecht y A. Salman, *An overview of clustering methods*, Intelligent Data Analysis, 11 (6), 583-605 (2007).
- Prakash, J. y otros cuatro autores, *Nonlinear Bayesian state estimation: Review and recent trends*, International Symposium on Advanced Control of Industrial Processes (ADCONIP), Hangzhou, China, 23-26 May (2011)._Ross, S., *Stochastic Processes*, 2ª edición, John Wiley & Sons, USA (1996).

Roth, J.T. y otros cuatro autores, *Quality and Inspection of Machining Operations: Tool Condition Monitoring*, *Transactions, Journal of Manufacturing Science and Engineering*: 132(4), 1-16 (2010).

Stewart, J., *Single Variable Calculus: Vol. 2, Early Transcendentals*, 7th Edition, MacMaster University, CAN (2012).

Teppola, P., S.P. Mujunen y P. Minkkinen, *A combined approach of partial least squares and fuzzy c-means clustering for the monitoring of an activated-sludge waste-water treatment plant*, *Chemometrics and Intelligent Laboratory Systems*, 41 (1), 95-103 (1998a).

Teppola, P., S.P. Mujunen y P. Minkkinen, *Adaptive Fuzzy C-means clustering in process monitoring*, *Chemometrics and Intelligent Laboratory Systems*, 45 (1-2) , 23-38 (1998b).

Tobon-Mejia, D. A., K. Medjaher, N. Zerhouni y G. Tripot, *Hidden Markov Models for Failure Diagnostic and Prognostic*, *Prognostics and System Health Management Conference (PHM'11)*, Shenzhen, China, 24-25 May (2011).

Udd, E., y W.B. Spillman, *Fiber Optic Sensors: An Introduction for Engineers and Scientists*, 2ª edición, John Wiley & Sons, Hoboken, USA (2011).

Van der Merwe, R. y E.A. Wan, *The square-root unscented Kalman filter for state and parameter-estimation*, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Salt Lake City, USA, 7-11 May (2001).

Waissman, J. y otros cuatro autores, *Wastewater Treatment Process Supervision By Means of a Fuzzy Automaton Model*, *Proceedings of the IEEE International Symposium on Intelligent Control (ISIC)*, 163–168, Rio Patras, Greece, 17-19 de Julio (2000).

Waissman, J., C. Ben-Yousself y G. Vazquez, *Fuzzy Automata Identification Based on Knowledge Discovery in Datasets for Supervision of a WWT Process*, 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications (SETIT), Tunisia, 27-31 de Marzo (2005).

Weber, P. y S. Gentil, *Fault detection using parameter estimation applied to a winding machine*. IAR Annual Conference, Duisburg, Allemagne, Nov. 20-21 (1997).

Zadeh, L.A., *Fuzzy Sets as a Basis of Theory of Possibility*, *Fuzzy Sets and Systems*, 1, 3-28 (1978).

Zadeh, L.A., *Fuzzy probabilities*, *Information Processing and Management*, 20(3), 363-372 (1984).

Zapata, G., J. Cardillo y E. Chacon, *Aportes Metodológicos para el Diseño de Sistemas de Supervisión de Procesos Continuos*, *Información Tecnológica*, La Serena, 22(3), 97-114 (2011).

Zhang, J., G. Welch y G. Bishop, *LoDiM: A novel power system state estimation method with dynamic measurement selection*, *IEEE Power and Energy Society General Meeting*, 1-7, NC, USA, 24-29 July (2011).