

## **Diseño de un Clasificador Difuso para el Establecimiento de los Estados Funcionales de un Sistema de Producción de Aire Medicinal**

**Jhon A. Ramírez<sup>(1)</sup>, Henry O. Sarmiento<sup>(1)</sup>, Jesús M. López-Lezama<sup>(2)</sup>**

(1) Grupo de Investigación en Instrumentación, Control Automático y Robótica ICARO, Facultad de Ingeniería, Politécnico Colombiano Jaime Isaza Cadavid, Carrera 48 No. 7-151, Medellín-Colombia. (e-mail: [jhon\\_ramirez91081@elpoli.edu.co](mailto:jhon_ramirez91081@elpoli.edu.co), [hosarmiento@elpoli.edu.co](mailto:hosarmiento@elpoli.edu.co))

(2) Grupo de Investigación GIMEL, Departamento de Ingeniería Eléctrica, Facultad de Ingeniería, Universidad de Antioquia, Calle 67 No. 53-108, Medellín - Colombia (e-mail: [jmaria.lopez@udea.edu.co](mailto:jmaria.lopez@udea.edu.co)).

*Recibido Abr. 4, 2017; Aceptado Jun. 12, 2017; Versión final Ago. 10, 2017, Publicado Dic. 2017*

---

### **Resumen**

Este artículo presenta el diseño de un clasificador difuso para establecer los estados funcionales de un sistema de producción de aire medicinal. La implementación de la lógica difusa en el clasificador permite que un objeto pueda pertenecer a múltiples clases, previamente definidas en la sintonización del clasificador. Para el diseño y sintonización del clasificador difuso se utilizó el histórico de datos del proceso y se hizo un pre-procesamiento de la información eliminando información redundante. Las clases obtenidas fueron sometidas a un análisis gráfico multivariable. La clasificación obtenida en el entrenamiento permitió identificar todos los estados funcionales útiles para el monitoreo del proceso. Todos los datos de prueba fueron clasificados acertadamente. El establecimiento de estados funcionales a partir de la clasificación difusa se presenta como una alternativa de la inteligencia artificial que permitirá en otra fase la programación de acciones en la ingeniería de mantenimiento a partir del diagnóstico en los procesos.

*Palabras clave: clasificador difuso; estados funcionales; producción de aire medicinal*

## **Design of a Fuzzy Classifier for Establishing the Functional States of a Medical Air Production System**

### **Abstract**

This article presents the design of a fuzzy classifier to establish the functional states of a medical air production system. The implementation of fuzzy logic in the classifier allows an object to belong to multiple classes, previously defined, in the tuning of the classifier. For the design of the classifier, the process historical data was used and a pre-processing of the information was done by eliminating redundant information. The classes obtained were subjected to a multivariable graphical analysis. The classification obtained in the training allowed identifying all functional states useful for the monitoring of the process. All test data were accurately classified. The establishment of functional states based on the fuzzy classification is presented as an alternative of the use of artificial intelligence for the scheduling of actions in the maintenance engineering, from the diagnosis of the processes.

*Keywords: fuzzy classifier; functional states; medical air production*

## INTRODUCCIÓN

Una de las alternativas para mantener la función de cualquier tipo de equipo o proceso es el monitoreo de sus variables críticas. Variables como la temperatura, corriente, voltaje, humedad u otras son monitoreadas constantemente en las plantas de procesos. Los datos en los sistemas de monitoreo son adquiridos por sensores instalados en puntos estratégicos. La instalación se hace fuera de línea para luego ejecutar pruebas en línea llevando el proceso a diferentes estados para diseñar la estrategia de diagnóstico o mejorar su desempeño cuando éste ya se encuentra implementado (Chuei y Chang, 2016). Cuando no se cuenta con sensores, la información puede ser extraída por modelos del proceso (Sarmiento, 2013). Los resultados del monitoreo son utilizados para apoyar la planeación del mantenimiento de las máquinas mediante acciones de tipo preventivas y predictivas, derivadas de los estados funcionales del sistema (Campuzano-Martínez, 2016). Las acciones correctivas no son programadas a partir del análisis de los valores de la medición de las variables críticas; el mantenimiento correctivo está asociado a la reparación de los equipos en el momento en que estos han fallado (Olarte *et al.*, 2010) por lo tanto, si se ha activado una acción correctiva ha sido porque en los resultados del monitoreo no se lograron definir los suficientes estados del proceso que anunciaran la falla. En consecuencia, el equipo operó hasta la pérdida de su función provocando tiempo de paro en el proceso.

La ingeniería ha permitido la optimización de las máquinas en cuanto a tamaño, costo y efectividad. Esta última en relación a la función para la que han sido diseñadas. Para esto, algunas máquinas en la industria cuentan con sistemas sofisticados de control. Sin embargo, el sistema de control se encarga solo de mantener la(s) salida(s) del proceso bajo un objetivo impuesto por el diseñador o el operador de la planta (Aguilar-Martín, 2007), menospreciando el nivel de monitoreo y diagnóstico del proceso (Gentil *et al.*, 2007), el cual permite establecer acciones de mantenimiento eficientes con el objetivo de mantener la funcionalidad del equipo. Por lo tanto, un buen sistema de control debe estar diseñado de manera que las variables controladas y manipuladas se mantengan en el punto de operación deseado y que además verifique que la función se cumple adecuadamente, incluido el diagnóstico de toda la maquinaria del proceso. Lo anterior a su vez debe estar asociado con el alcance que definen las organizaciones en relación a las estadísticas de CMD (Confiabilidad, Mantenibilidad y Disponibilidad) que finalmente se reflejan en disminución de gastos por mantenimiento y un alto nivel de seguridad en la producción (Mora, 2016).

El diagnóstico a partir de la clasificación se plantea como una herramienta para evaluar la aplicación de mantenimiento preventivo antes de la ocurrencia de un fallo (Corrales y Vázquez, 2013). En la clasificación de estados en grandes volúmenes de información los datos del proceso son clasificados y agrupados de acuerdo a la semejanza entre ellos. La lógica difusa combinada con la minería de datos generan clasificadores difusos. Esta combinación hace posible la clasificación de patrones, los cuales pertenecerán parcialmente a una o varias clases de acuerdo a un grado de pertenencia permitiendo de esta manera manejar el concepto de incertidumbre en la discriminación de los objetos a clasificar (Sarmiento, 2013). Las clases generadas en la agrupación de patrones adquieren propiedades cualitativas. Las propiedades son generalmente construidas por el operador experto de acuerdo a la clasificación. Las clases con su respectiva descripción cualitativa se transforman en estados funcionales del proceso. En el diagnóstico basado en clasificadores el operador define tantos estados como clases se obtengan y que tengan un significado válido en la dinámica del proceso. De esta manera se definen estados de falla y estados intermedios que pueden brindar conocimiento del proceso y que se obtienen del historial de datos (Sarmiento *et al.*, 2013). En ese sentido, Hernández *et al.*, (2014) presentan una estrategia para el monitoreo inteligente de una planta de tratamiento de agua a partir de la aplicación de la lógica difusa, la clasificación no supervisada y las redes neuronales artificiales. Mediante la estrategia propuesta se obtienen los estados funcionales y un modelo no lineal del proceso, los resultados indican mejoras en la calidad del agua para las siguientes etapas en la planta de producción de agua potable. En (Uribe y Isaza, 2012) los autores utilizaron la clasificación difusa y el conocimiento del experto para guiar la selección de variables en procesos industriales. La selección guiada permite reducir el conjunto de variables necesarias para representar la estructura intrínseca de los datos históricos del proceso necesitando solo de algunas variables para definir los estados funcionales del sistema. Sarmiento *et al.*, (2013) proponen la estimación de los estados funcionales de sistemas complejos con base en clasificación difusa. La estimación de estados funcionales es combinada con valores de probabilidad y posibilidad permitiendo obtener un sistema predictor capaz de estimar el siguiente estado del proceso.

En este trabajo se utilizó la lógica difusa aplicada a la clasificación de patrones para la estimación de los estados funcionales de un sistema de producción de aire medicinal. La clasificación difusa es implementada sobre el histórico de mediciones de las variables críticas del sistema, denominadas Variables Para el Control de la Producción de Aire Medicinal (VPCPAM). El historial de datos corresponde a un periodo de adquisición aproximado de 1.5 años, desde el 2015, con muestras cada 2 minutos y durante todos los días de la semana.

La clasificación resultante utilizada para establecer los estados funcionales del proceso a partir del descubrimiento de patrones en grandes volúmenes de información, se presentan como una alternativa en la ingeniería de mantenimiento para apoyar la programación de las intervenciones sobre las máquinas o equipos del proceso de Producción de Aire Medicinal. Se presta relevancia al hecho de que el clasificador es implementado sobre un proceso de suministro de aire medicinal a pacientes de entidades hospitalarias, demostrando que la inteligencia artificial se puede utilizar sobre procesos complejos relacionados con la seguridad de las personas.

## SISTEMA DE PRODUCCIÓN DE AIRE MEDICINAL

La producción de aire medicinal es el proceso mediante el cual el aire captado del ambiente es comprimido. La compresión del aire se produce mediante un sistema que modifica la presión y flujo a partir de los requerimientos de consumo. El sistema tiene como objetivo final entregar aire con características de pureza, de manera que pueda ser suministrado al paciente. Las condiciones de consumo del aire medicinal, se determinan a partir de la medición de las VPCPAM. Las VPCPAM utilizadas para evaluar la condición del aire corresponden a la medición de la Temperatura Ambiente (TA), Humedad Relativa (HR), Partes Por Millón de Monóxido de Carbono (PPCO) y la Temperatura de Punto de Rocío (TPR). La TA y HR se miden en el área confinada para la producción de aire medicinal. Las PPCO y TPR se miden directamente en la salida final del aire, es decir, que corresponden a la medición directa del aire suministrado a la red para consumo por parte del paciente. Las VPCPAM deben estar bajo rangos establecidos de acuerdo a las Buenas Prácticas de Manufactura (BPM) para la producción de aire medicinal en sitio. Cualquier medición fuera del rango de operación de las VPCPAM podría ocasionar pérdida de la función en el proceso.

Un sistema de producción de aire medicinal está compuesto por una serie de elementos encargados en conjunto de producir aire bajo normas y BPM. La función principal de este sistema es la generación de aire de forma uniforme y controlada de acuerdo con los estándares de calidad apropiados para el uso médico al que está dirigido. La Fig. 1 muestra el diagrama de entradas/salidas y los subsistemas que participan en la producción de aire medicinal. El sistema cuenta con un compresor de dos cabezotes encargado de comprimir el aire captado del ambiente. El aire comprimido es almacenado en un tanque pulmón. La salida de aire del tanque pulmón es sometida a una etapa de enfriamiento para disminuir el calentamiento del aire ocasionado por efectos termodinámicos. Finalmente, el aire es direccionado a una etapa de triple filtrado y luego al sistema de monitoreo para registrar las VPCPAM en el instante  $k$  de la medición. Si el aire cumple con los estándares de calidad y de BPM se entrega al paciente para su consumo por medio de la red de distribución. De lo contrario, el suministro es interrumpido para realizar el análisis del lote producido y efectuar las intervenciones de mantenimiento que permitan llevar el proceso a su estado normal de funcionamiento.

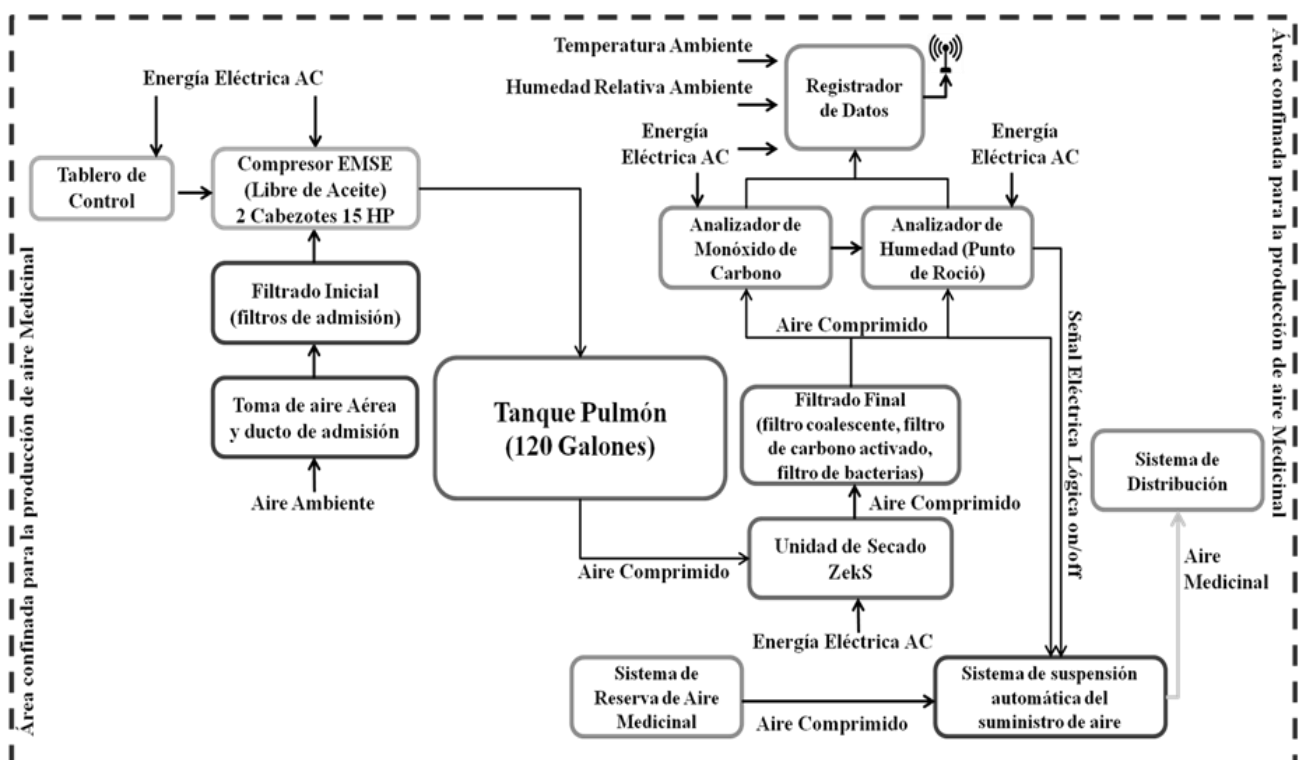


Fig. 1: Diagrama de entradas/salidas y subsistemas de un sistema de producción de aire medicinal.

## CLASIFICACIÓN DE PATRONES

Dentro de las técnicas útiles para realizar la tarea de diagnóstico automático, se encuentran los métodos que identifican estados funcionales por medio de técnicas de clasificación o agrupamiento (Rojas *et al.*, 2008). La clasificación puede verse como una actividad que ocurre en dos etapas. En la primera se lleva a cabo el entrenamiento y aprendizaje. En la segunda se decide si un nuevo objeto es un posible miembro de una clase a partir de propiedades comunes (Lara, 2007). La clasificación permite discriminar estados funcionales y detectar estados normales de funcionamiento, estados de alarma y estados de falla en el diagnóstico del proceso.

### *Entrenamiento y reconocimiento*

En la etapa de entrenamiento o aprendizaje la selección de variables para el clasificador, al igual que la descripción e interpretación de las clases resultantes, pueden ser definidas con el apoyo del experto en el proceso. La selección de parámetros del clasificador se define en un proceso de ajuste de valores con base en el desempeño del clasificador ante los datos de entrenamiento, y fundamentalmente ante los datos de prueba (muestras no utilizadas en entrenamiento) y de validación si se dispone de ellos. El análisis de las señales o variables frente a las clases obtenidas permite definir los estados funcionales del proceso a partir del conocimiento del experto. La eficiencia del clasificador depende del aprendizaje y puede ser evaluada mediante simulación, ingresando muestras conocidas (patrones de prueba) por el experto y que no fueron utilizadas durante el proceso de entrenamiento. La segunda etapa es denominada etapa de reconocimiento y tiene como objetivo reconocer la clase en la cual se ubica cada nueva muestra, y de acuerdo con ello el estado funcional actual del proceso.

El principio del monitoreo de un proceso dinámico basado en clasificadores es el de determinar para cada muestra tomada su pertenencia a una de las clases obtenidas. A estas clases previamente les fue asociado un estado funcional o situación del sistema según el criterio y conocimiento del experto (operario de mantenimiento). Para la implementación (entrenamiento) del clasificador orientado al monitoreo se requieren los datos históricos disponibles representativos del proceso. Posteriormente, ante nuevos datos éste debe “decidir” a qué clases (estados funcionales) pertenecen (Sarmiento, 2013). Los diferentes algoritmos de agrupamiento son clasificados, en lo que se referencia como la forma clásica, en dos principales tipos: algoritmos particionales/jerárquicos y algoritmos de pertenencia exclusiva/no exclusiva. Estos algoritmos de pertenencia exclusiva/no exclusiva, hacia donde se orienta el artículo, son también denominados de agrupamiento concreto/difuso. Los algoritmos de agrupamiento estricto o concreto son aquellos donde un objeto o muestra sólo puede pertenecer a un conjunto o clase (*Hard clustering*), y agrupamiento borroso o difuso (*Fuzzy clustering*) donde los objetos o muestras pueden pertenecer a más de un conjunto con un grado de pertenencia a cada partición. Cualquier método de agrupamiento seleccionado deberá contar con información relevante del proceso. Para garantizar información representativa se debe realizar una optimización de datos de manera que se genere una nueva matriz de información a partir del pre-procesamiento (Prada, 2009).

### *Pre-procesamiento de datos.*

La matriz de datos históricos  $\mathbb{X}$ , dada por la ecuación (1) está conformada por  $n$  filas y  $m$  columnas. Cada fila  $n$  corresponde a un objeto  $x_i$ , el cual contiene los atributos (valores de las variables de los procesos) que diferencian cada fila de la matriz de datos. En general, un objeto es una muestra o patrón de las variables del proceso observadas en un instante  $k$  y almacenadas en el vector  $x_i(k)$ . El vector de datos  $x_i$  debe contener el conjunto de atributos o variables seleccionadas capaces de caracterizar cada objeto de la base de datos. Este vector también es conocido como vector característico (Lückeheide *et al.*, 2007).

$$\mathbb{X} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \dots & \vdots \\ a_{i,1} & a_{i,2} & \dots & a_{i,m} \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ x_n \end{bmatrix}; \rightarrow x_1 = [a_{1,1} \ a_{1,2} \ \dots \ a_{1,m}] \quad (1)$$

Una de las estrategias para pre-procesar el conjunto de datos se basa en el conocimiento y la experticia del operador del proceso. La experiencia adquirida por el experto le da la capacidad para centrarse en un conjunto de datos en especial y eliminar información que resulte redundante. Otra estrategia consiste en la normalización de los datos. La normalización se sugiere para procesos en los cuales las magnitudes de las variables varían considerablemente, lo que puede hacer que aquéllas de mayor magnitud le resten importancia a otras de menor magnitud (Cuadrado, 2002; Lückeheide *et al.*, 2007). Una de las formas de llevar los datos a un rango reducido es mediante la normalización “*Min-Max*”, la cual consiste en transformar los datos a un rango entre  $[0,1]$ . Lo anterior se lleva a cabo aplicando la ecuación (2).

$$\mathbb{X}_{(j,i)} = \frac{\mathbb{X}_{(j,i)} - \text{Mínimo}(i)}{\text{Máximo}(i) - \text{Mínimo}(i)} \quad (2)$$

En esta ecuación,  $\mathbb{X}_{(j,i)}$  es el atributo de la matriz de datos en la columna  $j$  y fila  $i$ ;  $\text{Máximo}(i)$  es el valor máximo del atributo de la fila ( $i$ ), y  $\text{Mínimo}(i)$  es el valor mínimo del atributo de la fila ( $i$ ). La ecuación (2) aplicada a la matriz de datos  $\mathbb{X}$  permite obtener la matriz de datos normalizada que es la utilizada para el entrenamiento del clasificador.

#### Clasificación concreta (hard clustering)

La clasificación concreta forma parte de los métodos que utilizan una función objetivo para discriminar los datos y agruparlos en una clase caracterizada por un punto central denominado centroide (Ghosh y Kumar, 2013). Durante un proceso iterativo cada elemento del arreglo de datos es evaluado minimizando la función objetivo. Existe una gran cantidad de algoritmos de agrupamiento con función objetivo del tipo concreto. Cada uno corresponde a una técnica de optimización diferente entre las que se encuentran: algoritmos iterativos, mapas auto organizados y algoritmos evolutivos, entre otros. En este caso se describe uno de los más representativos denominado *C-means*.

*C-means* (también conocido como *K-means*) es un algoritmo iterativo en el cual un objeto  $x_i$  de la matriz de datos  $\mathbb{X}$  solo puede pertenecer a una clase  $C_j$  limitada por un centroide  $c_j$  como se muestra en la Fig. 2a. Este clasificador concreto tiene como función objetivo minimizar la suma cuadrática de la distancia euclidiana, dada por la ecuación (3), entre los puntos y un prototipo o valor medio del conjunto (Bora y Gupta 2014; Rojas et al., 2008).

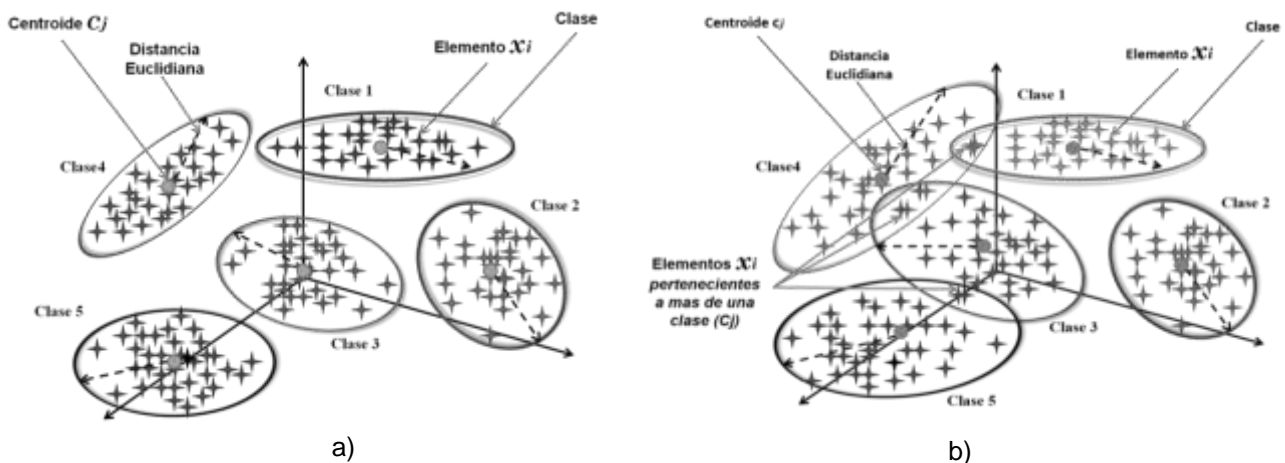


Fig. 2: Representación gráfica de clasificadores: a) Clasificador concreto *C-means*, b) Clasificador difuso.

$$J(U, v) = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (3)$$

Donde:

$J$  = Función Objetivo.

$U$  = Matriz de particiones (de valores binarios).

$v$  = Vector de centro de las clases.

$x_i^j$  = Muestra de los datos a evaluar (*fila* <sub>$i$</sub>  de la matriz de datos  $\mathbb{X}$ ).

$c_j$  = Centroide de la clase.

$\|x_i^j - c_j\|^2$  = Distancia euclidiana medida en un espacio de  $m$  dimensiones entre la muestra  $x_i^j$  y el centro  $c_j$  del cluster  $i$ .

#### Diagrama de flujo del algoritmo *C-means*

El diagrama de flujo para la implementación del algoritmo *C-means* se muestra en la Fig. 3a (Ghosh y Dubey 2013; Ponce, 2010). La condición  $2 \leq C \leq n$  permite definir más de una clase. Sí  $C = 1$  se asume que todos los datos pertenecen a un solo conglomerado y por lo tanto, el algoritmo tendrá convergencia en

una sola iteración. Los parámetros de entrada del algoritmo corresponden al número de clases que se desean encontrar a lo largo de la clasificación, las cuales deben ser definidas para dar inicio al algoritmo (Ghosh y Dubey, 2013; Rojas *et al.*, 2008). Si bien el algoritmo *C-means* es uno de los más representativos, éste presenta algunos inconvenientes en la discriminación de patrones para casos específicos en los que  $n$  patrones diferentes de la matriz de datos  $X = [x_1, x_2, x_3 \dots x_n]$  se ubiquen a una misma distancia del centroide  $c_j$ . En estos casos no existen criterios de desempate establecidos durante la ejecución del algoritmo y su convergencia estaría limitada sobre mínimos locales. Lo anterior demuestra la principal desventaja del clasificador *C-means*; sin embargo, su aplicación fortaleció el estudio de alternativas que permitieran un grado de ambigüedad en la discriminación de patrones.

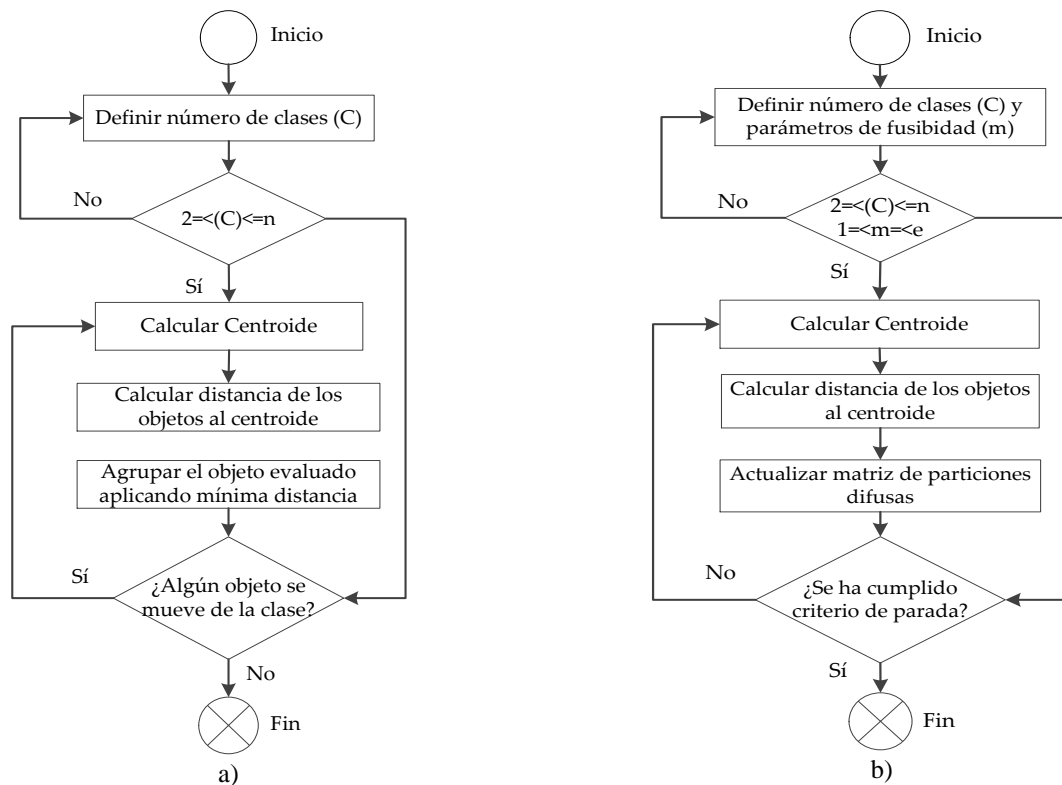


Fig. 3: Diagrama de flujo de algoritmo de clasificación: a) *C-means* y b) *FCM*.

#### Clasificación borrosa o difusa (fuzzy clustering)

La lógica borrosa (también conocida como lógica difusa) es una propuesta que busca incorporar en los sistemas de ingeniería el conocimiento expresado en términos lingüísticos o expresiones verbales. De esta manera la lógica borrosa pretende emular la forma de razonamiento humano (Escalona *et al.*, 2003), en especial el grado de incertidumbre que trae consigo la descripción de percepciones humanas expresadas lingüísticamente por términos que carecen de exactitud y que permiten un grado de pertenencia entre diferentes clases. Mientras que en la lógica convencional un elemento  $A$  solo puede pertenecer mediante un grado de pertenencia  $\mu = [0 \text{ ó } 1]$ ; (falso o verdadero) a un único conjunto  $X$ , en la lógica borrosa el elemento ( $A$ ) puede pertenecer simultáneamente a los conjuntos  $X$  y  $Y$  de una colección de objetos  $U$ , mediante un grado de pertenencia  $\mu_{X(A)} = [0,1]$  y  $\mu_{Y(A)} = [0,1]$ . El grado de pertenencia  $\mu$  se encuentra en todo el intervalo continuo  $[0,1]$  (Jantzen, 2007) y se extiende en general para  $n$  elementos y  $m$  conjuntos mediante funciones de pertenencia. Una función de pertenencia se define como aquella aplicación que asocia a cada elemento de un conjunto borroso el grado con que pertenece al valor lingüístico asociado (Aguilar-Martín, 2007). Lo anterior permite expresar la pertenencia del elemento  $A$  a los conjuntos  $X$  y  $Y$  mediante la definición mostrada en la ecuación (4).

$$X = \{(A, \mu_{X(A)}) | A \in U\}; Y = \{(A, \mu_{Y(A)}) | A \in U\}; (\mu_{X(A)}) = [0,1]; (\mu_{Y(A)}) = [0,1] \quad (4)$$

La clasificación borrosa se presenta como una alternativa para la discriminación de patrones que presentan cierto grado de similitud y que pertenecen simultáneamente a más de una clase. En esta clasificación se utiliza el concepto de partición difusa definida como la generalización de la partición concreta en la cual se permite un grado de pertenencia  $\mu = [0 \text{ ó } 1]$  pero ahora en el intervalo de valores reales  $\mu = [0, 1]$  donde el mayor grado de pertenencia es 1 y el menor es 0, (Ojeda, 2010; Sarmiento, 2013). La clasificación borrosa

permite que  $n$  patrones diferentes de la matriz de datos  $\mathbb{X} = [x_1, x_2, x_3 \dots x_n]$  pertenezcan parcialmente a una clase  $C_j$  o a varias clases  $C = [C_1, C_2, C_3 \dots C_n]$  limitada por su respectivo centroides  $c_j$  como se muestra en la Fig. 2b. Lo anterior hace posible extender el espacio de búsqueda sujeto a pocas restricciones que permite mayor flexibilidad en la clasificación (Lara, 2007). Entre los algoritmos utilizados para la clasificación borrosa y que hacen uso de una función objetivo se encuentra el algoritmo Fuzzy *C-means* (FCM). A diferencia del algoritmo *C-means* en el cual se hacen particiones duras, en el algoritmo FCM se trabajan con particiones suaves, restringidas por la función objetivo definida en la ecuación (5).

$$J_m(U, v) = \sum_{i=1}^N \sum_{j=1}^C \mu_{ij}^m \|x_i - c_j\|^2 \quad (5)$$

Donde:

$J_m$  = Función objetivo.

$U$  = Matriz de particiones de valores reales en el intervalo continuo  $[0,1]$ .

$v$  = Vector de centro de las clases;  $m$  = Parámetro que regula la fusividad de la partición.

$\mu_{ij}$  = Pertenencia de los datos  $x_i$  a la clase  $j$ .

$x_i$  = Muestra de los datos a evaluar;  $c_j$  = Centroide de la clase  $j$ .

$\|x_i^j - c_j\|^2$  = Distancia euclidiana medida en un espacio de  $m$  dimensiones entre la muestra  $x_i^j$  y el centro  $c_j$  del cluster  $i$ .

Los centroides de cada clase se calculan como se muestra en la ecuación (6).

$$c_j = \frac{\sum_{i=1}^N \mu_{ij}^m * x_i}{\sum_{i=1}^N \mu_{ij}^m} \quad (6)$$

Como el FCM es un algoritmo iterativo, la actualización de la matriz de particiones para cada iteración se calcula mediante la ecuación (7). La convergencia del algoritmo tiene lugar cuando se cumple un número de iteraciones definido en los parámetros de inicialización o mediante la comparación de la matriz de particiones frente a un parámetro de tolerancia  $e$  como se muestra en la ecuación (8).

$$u_{ij} = \frac{1}{\sum_{j=1}^c \left( \frac{\|x_i - c_j\|}{\|x_i - c_j\|} \right)^{\frac{2}{m-1}}} \quad (7)$$

$$\|U^{k+1} - U^k\| < e \quad (8)$$

El parámetro  $m$  es conocido como grado de fusividad y se le pueden asignar valores en el intervalo ( $0 \leq m \leq \infty$ ); sin embargo, se recomiendan valores en el intervalo  $[0, 2]$  (Celemín et al., 2013). El diagrama de flujo para la implementación del algoritmo FCM se ilustra en la Fig. 3b (Ghosh y Dubey, 2013; Ponce, 2010).

## PRUEBAS Y RESULTADOS

Para establecer los estados funcionales del sistema de producción de aire medicinal se sintonizó un clasificador FCM. Los datos iniciales adquiridos del proceso se registraron en una base de datos de 99750 muestras para las 4 variables (matriz  $[99750 \times 4]$ ). Los datos iniciales fueron sometidos a pre-procesamiento eliminando información redundante. La variable HR fue previamente filtrada para eliminar la componente de alta frecuencia presente en su medición. El pre-procesamiento generó una nueva base de datos de 5609 muestras (matriz  $[5609 \times 4]$ ), disponiendo de esta el 66% de datos para entrenamiento (3679 muestras) y el 34% de datos para prueba (1930 muestras). El objeto (vector característico) de la matriz de datos se muestra en la ecuación (9) y está conformado por la medición en el instante  $k$  de las VPCPAM. Cada variable medida corresponde a un atributo del objeto  $x_i$  como se indica en la ecuación (10).

$$x_{i(k)} = [PPMCO_{(k)} \quad TPR_{(k)} \quad HR_{(k)} \quad TA_{(k)}]; \quad k: \text{ instante de medición}; \quad (9)$$

$$\text{Atributos}_{(k)} = PPMCO_{(k)}; TPR_{(k)}; HR_{(k)}; TA_{(k)} \quad (10)$$

En la ecuación (11) se pueden observar los 5 primeros objetos de la matriz de datos normalizada del sistema de producción de aire medicinal.

$$X = \begin{bmatrix} 0.2404 & 0.3456 & 0.8234 & 0.1663 \\ 0.5490 & 0.3344 & 0.7897 & 0.1685 \\ 0.3769 & 0.3368 & 0.8135 & 0.1641 \\ 0.3086 & 0.3336 & 0.7716 & 0.1663 \\ 0.000 & 0.3368 & 0.8274 & 0.1641 \end{bmatrix} = \begin{bmatrix} \text{objeto normalizado } x_1 \\ \text{objeto normalizado } x_2 \\ \text{objeto normalizado } x_3 \\ \text{objeto normalizado } x_4 \\ \text{objeto normalizado } x_5 \end{bmatrix} \quad (11)$$

El entrenamiento del clasificador fue realizado utilizando el software Matlab®. Los centros fueron elegidos a priori para un total de 6 particiones en las que se encuentran los diversos estados funcionales asociados a cada clase y de acuerdo al conocimiento del experto sobre el sistema de producción de aire medicinal. El esquema utilizado para el diseño del clasificador se muestra en la Fig. 4.

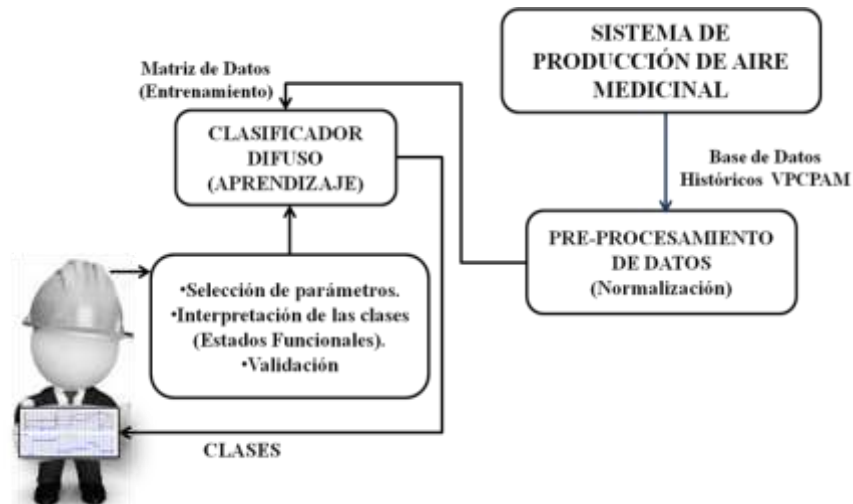


Fig. 4: Esquema utilizado para entrenar el clasificador del sistema de producción de aire medicinal.

**Establecimiento de los estados funcionales**

Para establecer los estados funcionales del sistema de producción de aire medicinal se analizaron los resultados gráficos del clasificador. En las Fig. 5a y 5b se ilustra la respuesta gráfica del histórico de datos y el comportamiento del clasificador sintonizado. El análisis gráfico multivariable y la experiencia de 4 operadores del proceso, incluido el ingeniero jefe de producción, permitieron establecer los estados funcionales reportados en la Tabla 1. Se puede observar que la Fig. 5a muestra una transición entre la clase 1 y la clase 6, la cual se da al inicio de las muestras. Posteriormente, se evidencian varias transiciones entre la clase 1 y la clase 5. Después de la muestra 80 se puede observar una transición de la clase 1 a la clase 5 y posteriormente a la clase 6. Por otro lado, en la Fig5b se presenta una transición de la clase 3 a la clase 6 después de la muestra 1220 y transiciones posteriores entre las clases 6, 4, 2 y 5. No aparecen transiciones a la clase 1 ya que las partes por millón de monóxido de carbono se mantuvieron por debajo del máximo permitido.

Tabla 1: Descripción de los estados funcionales a partir del clasificador FCM diseñado.

Clase	Estado Funcional
C1	Falla: Incremento considerable en las partes por millón de monóxido de carbono (PPMCO)
C2	Falla: Incremento considerable de la temperatura ambiente (TA) y decremento considerable de la humedad relativa HR
C3	Alerta: Incremento de la temperatura ambiente (TA < TA de la clase C2) y decremento de la humedad relativa (HR > HR de la clase 2)
C4	Alerta: incremento de la temperatura de punto de rocío (TPR) e incremento de la temperatura ambiente (TA)
C5	Alerta: Incremento considerable de la humedad relativa (HR)
C6	Operación Normal: Las VPCPAM se encuentran dentro de los rangos normales de operación.



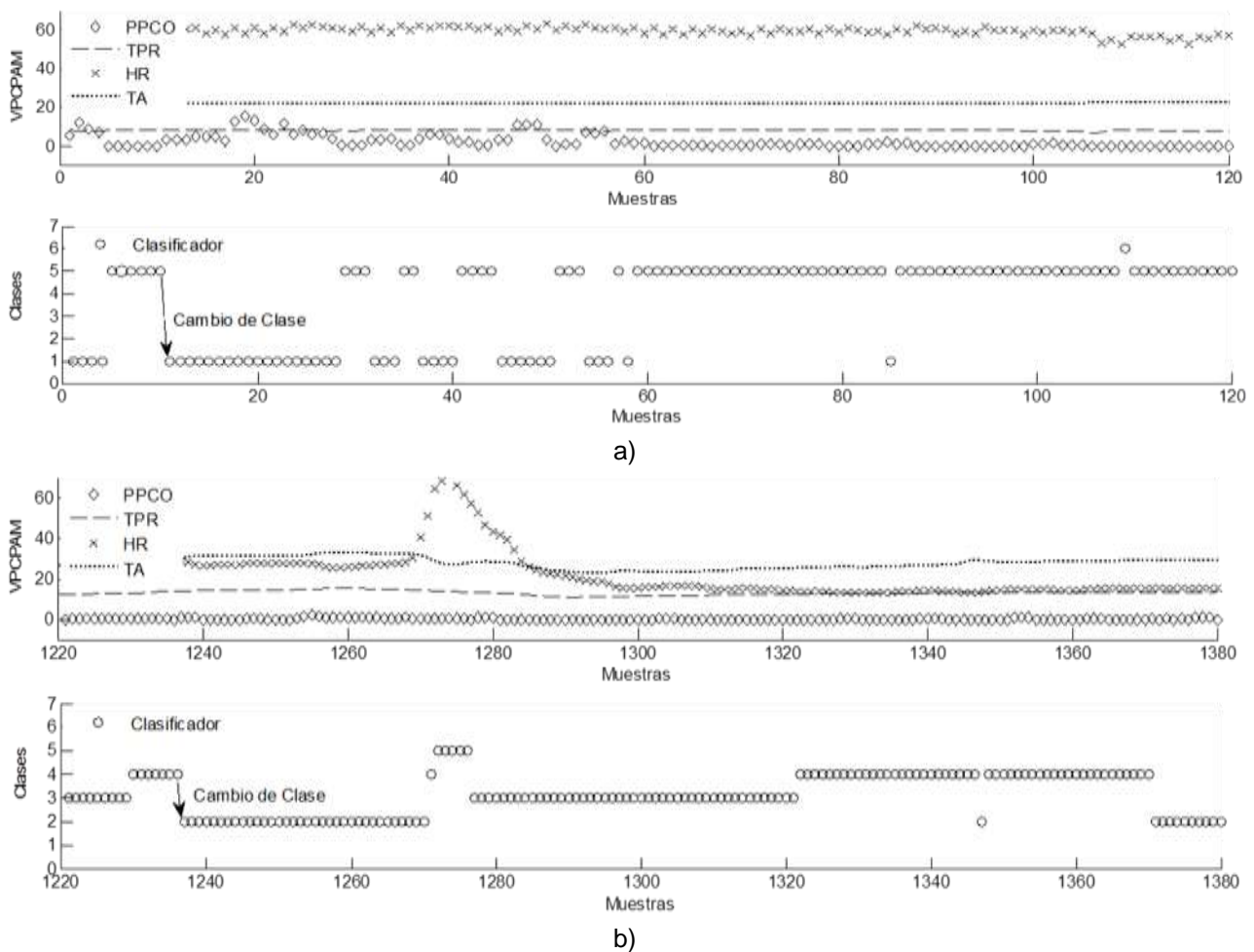


Fig. 5: Respuesta gráfica de las VPCPAM y del clasificador sintonizado (entrenamiento) para el sistema de producción de aire medicinal a) Ampliación entre 0 a 120 muestras. b) Ampliación entre 1220 a 1380 muestras.

*Validación de clases (estados funcionales)*

Se realizaron medidas de validez de las clases para la partición obtenida. Para ello se usaron el Coeficiente de Partición (*Partition Coefficient- PC*) y la Entropía de la Partición (*Partition Entropy – PE*) dados mediante las ecuaciones (12) y (13), respectivamente. Donde *c* es el número de clases, *N* el número de muestras y  $\mu$  el grado de pertenencia (Xie y Beni, 1991). Los resultados obtenidos de estos índices (presentados en la Tabla 2), en conjunto con el reconocimiento y validación de las clases, con base en el conocimiento experto de los involucrados directamente sobre el proceso, dan cuenta del número válido y acertado de clases (C1 a C6).

$$PC(c) = 1/N \sum_{i=1}^c \sum_{j=1}^N \mu_{ij}^2 \tag{12}$$

$$PE(c) = -1/N \sum_{i=1}^c \sum_{j=1}^N \mu_{ij} \log_2(\mu_{ij}) \tag{13}$$

Tabla 2: Medidas de validez de partición difusa.

Índice de Validación	Resultado	Mínimo Posible	Máximo Posible
Coeficiente de Partición	0.6669	Mín=1/c	1
		0.1667	
Entropía de la Partición	0.0007	0	Max= Log <sub>2</sub> (c)
			2.585

Prueba del clasificador

La prueba tiene como objetivo determinar el grado de generalización del clasificador obtenido ante muestras que no fueron utilizadas para el entrenamiento. Durante esta etapa se presentan los datos de prueba al clasificador, se analiza la clasificación generada a partir de la matriz final de centros obtenida en el entrenamiento dada por la ecuación (14) y se discrimina el objeto presentado  $x_i$  del instante  $k$  en una de las seis clases  $C_j = 1, 2, 3, 4, 5, 6$ .

$$v = \begin{bmatrix} 0.2580 & 0.0156 & 0.0395 & 0.0117 & 0.0124 & 0.0107 \\ 0.3562 & 0.8831 & 0.6893 & 0.7126 & 0.3506 & 0.3539 \\ 0.8085 & 0.1539 & 0.3073 & 0.2599 & 0.7944 & 0.6230 \\ 0.1779 & 0.7923 & 0.4236 & 0.4867 & 0.1879 & 0.2006 \end{bmatrix} \quad (14)$$

En la Fig. 6 se ilustra la respuesta gráfica de prueba del clasificador. La Fig. 6 muestra reconocimiento de todas las clases de acuerdo al estado cualitativo asociado. Mediante la observación y análisis detallado de las variables dinámicas del proceso se verificó la acertada clasificación de todos los estados funcionales presentes, es decir un desempeño de clasificador del 100% en las 1930 muestras.

Como otra información relevante, desde el punto de vista de la importancia que reviste el trabajar con métodos difusos, se señala y analiza el comportamiento e información que se deriva de los grados de pertenencia obtenidos en entrenamiento y prueba; específicamente en los casos de transición de estados funcionales. Las transiciones o la evolución del sistema alrededor de las mismas, es información valiosa al operador por cuanto a partir de estos valores difusos se pueden tomar decisiones sobre el proceso, decisiones que podrían evitar evolucionar hacia estados de falla, por ejemplo, en situaciones de fallas incipientes que van aumentando con el tiempo (i.e. deterioro del dispositivo, equipos, y el mismo proceso, entre otras).

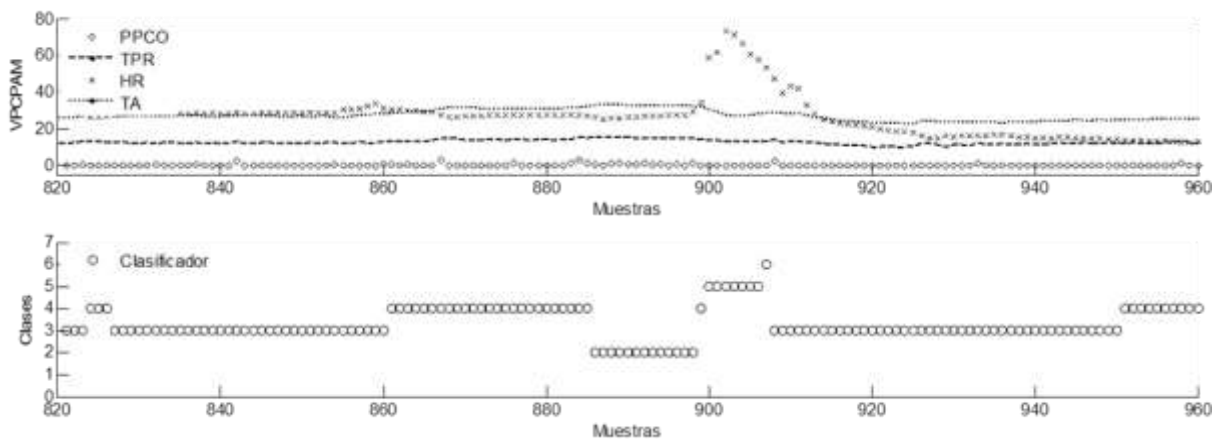


Fig. 6: Respuesta gráfica de prueba del clasificador, ampliación entre 820 a 960 muestras.

En la Fig. 7 se puede observar el comportamiento de los grados de pertenencia obtenidos con el clasificador para las muestras 850 a 870 donde se presenta el cambio de estado funcional C4 a C3 específicamente en la muestra 861. Si el operador es informado continuamente del valor de los grados de pertenencia, este podría observar que las dos funciones de pertenencia con valores mayores corresponden a estados asociados a C3 y C4 (ver Tabla 3), y visualizar que a partir de la muestra 858 se empieza a configurar una tendencia de cambio.

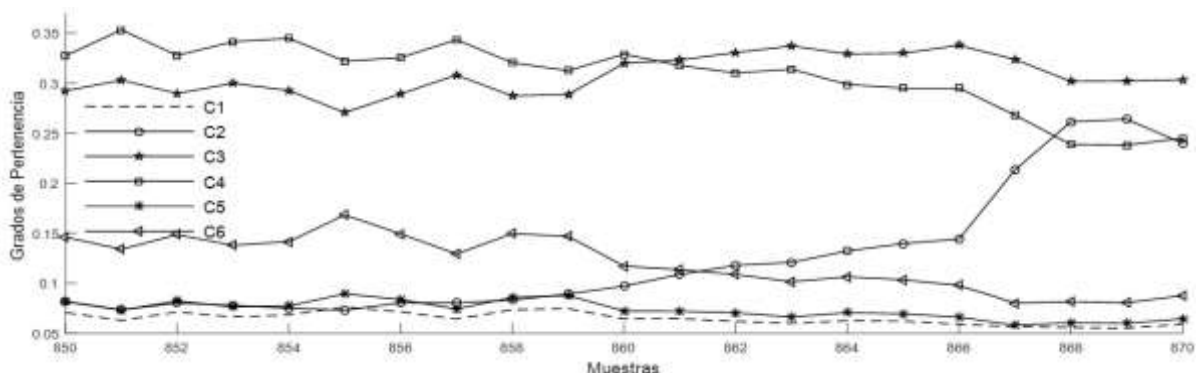


Fig. 7: Evolución de grados de pertenencia muestras 850 a 870. Valores más altos para clases C3 y C4.

Tabla 3: Transición de estados funcionales (C4 a C3), evolución de grados de pertenencia

Muestra	$\mu_{C1}$	$\mu_{C2}$	$\mu_{C3}$	$\mu_{C4}$	$\mu_{C5}$	$\mu_{C6}$	Clase
857	0.064	0.080	0.308	0.343	0.074	0.129	C4
858	0.074	0.083	0.287	0.320	0.086	0.150	C4
859	0.075	0.089	0.289	0.313	0.088	0.147	C4
860	0.064	0.097	0.320	0.329	0.072	0.117	C4
861	0.065	0.109	0.323	0.318	0.072	0.114	C3
862	0.062	0.118	0.331	0.310	0.070	0.109	C3

El operario puede identificar una tendencia de cambio desde el estado de alerta (C4), hacia otra alerta igualmente no deseable para el proceso (C3). Esta es la información que le sería útil para tomar decisiones de acuerdo a las necesidades del proceso, con el objetivo de llevar el proceso a estados normales o deseables. De igual forma, analizando los grados de pertenencia resultado de procesar los datos de prueba (ver Fig. 8) en el clasificador entrenado fue posible identificar entre las muestras 1350 y 1370 la tendencia al cambio del estado de alarma C4 al estado de falla C5 (transición en muestra 1359). Esta información permitiría al operador realizar acciones de corrección sobre el proceso para evitar el estado de falla.

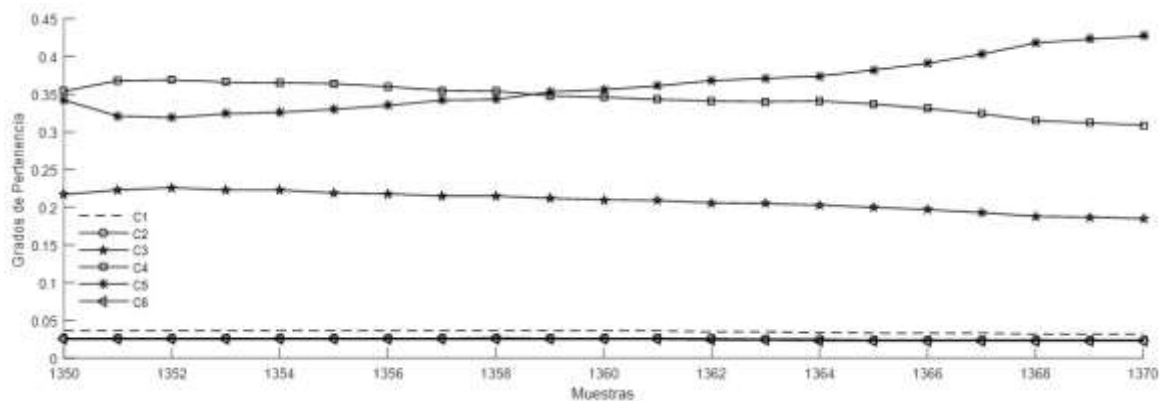


Fig. 8: Grados de Pertenencia a partir de datos de prueba (muestras 1350 a 1370). Transición de Alarma C4 a Estado de Falla C5.

También, resultado del monitoreo basado en clasificadores difusos, y como apoyo visual al operador del proceso se genera el grafo de conexiones entre estados funcionales a partir de los resultados del entrenamiento. Este grafo acompañado de la visualización en línea de la clasificación y grados de pertenencia correspondientes constituirá la información útil en la tarea de supervisión para la toma de decisiones que se requiera. A partir de los resultados de clasificación obtenidos en entrenamiento, el grafo experimental de la Fig. 9 fue construido con base en los estados funcionales y su evolución en el proceso, utilizando la denominada matriz de transiciones (Kempowsky, 2006), ver Tabla 4.

El grafo de conexiones (ver Fig.9), que puede ser actualizado con los resultados de prueba y en etapa de reconocimiento en línea, permite ubicar en cada instante el estado del proceso y proyectar, en forma ágil para el operador el camino posible que pueda tomar.

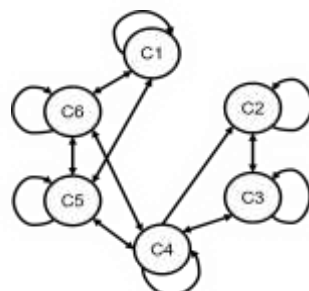


Fig. 9: Grafo experimental del proceso a partir de datos de entrenamiento.

Tabla 4: Matriz de transición de estados

	C1	C2	C3	C4	C5	C6
C1	149	0	0	0	34	2
C2	0	475	5	0	0	0
C3	0	4	549	27	0	0
C4	0	1	26	387	1	4
C5	32	0	0	1	1309	44
C6	3	0	0	4	43	577

## CONCLUSIONES

En este artículo se presentó el diseño de un clasificador FCM para establecer los estados funcionales de un sistema de producción de aire medicinal. El clasificador fue diseñado integrando la lógica difusa y el historial de datos del proceso. La etapa de prueba del clasificador entrenado permitió determinar el grado de generalización del clasificador ante objetos no utilizados en la etapa de entrenamiento. Los resultados en la etapa de prueba mostraron reconocimiento de todas las clases de acuerdo al estado cualitativo asociado.

La información de los grados de pertenencia obtenidos con el clasificador difuso resultante, le permitirá al operador o encargado del proceso definir el tipo de acciones a realizar con el fin de evitar llegar a estados funcionales no deseables o llevar al sistema a realizar transiciones para salir de ellos.

Trabajar el monitoreo de procesos con base en clasificadores del tipo difuso, hace posible la extracción de información de los datos históricos del proceso tal como: grados de pertenencia, clases asociadas a estados funcionales, visualización de tendencias de los estados del proceso, y grafos de estados del proceso con visualización en la etapa de reconocimiento en cada muestreo. Todo lo anterior resulta como información útil en la toma de decisiones que debe afrontar un operador del proceso en su trabajo continuo.

La clasificación obtenida asociada a los estados del proceso de producción de aire medicinal se considera como una herramienta de apoyo para el departamento de ingeniería de mantenimiento de las organizaciones. La clasificación permite mostrar resultados claros, confiables y de fácil interpretación para la posterior planeación y programación del mantenimiento. Durante el desarrollo del trabajo se tuvieron en cuenta los conocimientos del grupo de ingeniería involucrado con el proceso de producción de aire. El conocimiento de los expertos del proceso fue fundamental para la asignación de clases y la definición de los estados funcionales del proceso.

## REFERENCIAS

Aguilar-Martín, J., *Inteligencia Artificial para la Supervisión de Procesos Industriales*, 1ª Ed., ULA, Mérida, Venezuela (2007)

Bora, D. y A. Gupta, A comparative study between fuzzy clustering algorithm and hard clustering algorithm, *International Journal of Computer Trends and Technology*, 2(10), 108-113 (2014)

Campuzano-Martínez, I.R., Diagnóstico de generadores eléctricos de potencia con técnicas de monitoreo en línea y fuera de línea, *Información Tecnológica*, 27(2), 11-20 (2016)

Celemín, C., H. Martínez y M. Melgarejo, Sintonización de clasificadores difusos por medio de algoritmos genéticos inicializados en FCM, *Ingeniería y Competitividad*, 1(5), 9-20 (2013)

Chuei, W.C. y T. Chang, Timed-automata based method for synthesizing diagnostic tests in batch processes, *Computers & Chemical Engineering*, 4(84), 12-27 (2016)

Corrales, L. y A.R. Vázquez, Clasificación de fallas con redes neuronales para grupos electrógenos, *Ingeniería Energética*, 2(34), 137-150 (2013)

Cuadrado, A.A., Supervisión de procesos complejos mediante técnicas de data mining con incorporación de conocimiento previo, Tesis de Doctorado, Departamento de Ingeniería Eléctrica, Electrónica, de Computadores y Sistemas, Universidad de Oviedo, España (2002)

- Escalona, L.W., C.J. Jiménez, A.J. Ferrera y F.J. Arteaga, Diseño de un sistema para la detección de fallas en planta compresora con mantenimiento centrado en confiabilidad usando lógica difusa, Ingeniería UC, 1(10),1-9 (2013)
- Gentil, S., B. Celse, S. Charbonnier, V. Cocquempot, F. Hamelin, S. Lesecq, D. Maquin, J. Montmain, J. Ragot, D. Sauter, Supervision des Procédés Complexes, Hermes Science Publication Lavoisier, Paris, Francia (2007)
- Ghosh, S. y S.K. Dubey, Comparative Analysis of K-means and Fuzzy C-means Algorithms, International Journal of Advanced Computer Science and Applications, 4(4), 35-40 (2013)
- Hernández, H.R., J.L. Camas, A. Medina, M. Perez y V. Le Lann Fault, Diagnostic by Lambda methodology applied to drinking water plant, IEEE Latin America Transactions, 12(6), 985-990 (2014)
- Jantzen, J., Foundation of fuzzy control, John Wiley & Sons Ltd., Chichester, England (2007)
- Kempowsky, T., A. Subias, J. Aguilar-Martin, Process situation assessment: From a fuzzy partition to a finite state machin., Eng. Appl. of AI 19(5), 461-477 (2006)
- Lara, O.J., Nuevas metodologías no invasivas de diagnóstico de defectos incipientes en rodamientos de bola, Tesis Doctoral, Departamento de Ingeniería Mecánica, Universidad Carlos III de Madrid, Madrid, España (2007)
- Lückeheide, S., J. Velásquez y L. Cerda, Segmentación de los contribuyentes que declaran IVA aplicando Herramientas de Clustering, Revista Ingeniería de Sistemas, 21, 87-110 (2007)
- Mora, A., Mantenimiento industrial efectivo, Edición 2016, Editorial COLDI LTDA., Medellín, Colombia (2016)
- Ojeda, B., Aportación a la extracción de conocimiento aplicada a datos mediante agrupamiento y sistemas difusos, Tesis Doctoral, Escuela Técnica Superior de Ingenieros de Telecomunicación, U. Politécnica de Madrid, Escuela Técnica Superior de Ingenieros de Telecomunicaciones, Madrid, España (2010)
- Olarte, W., M. Botero y B. Cañón, Técnicas de mantenimiento predictivo utilizadas en la industria, Scientia et Technica, 45, 223-226 (2010)
- Ponce, P., Inteligencia artificial con aplicaciones a la ingeniería. 1ª Ed., Alfaomega, Ciudad de México, México (2010)
- Prada, M., Técnicas de extracción del conocimiento basadas en data mining visual para la supervisión de procesos industriales. Análisis de la dinámica basado en mapas auto-organizados, Tesis Doctoral, Departamento de Ingeniería Eléctrica y de Sistemas y Automática, Universidad de León, España (2009)
- Rojas, J., J. Chavarro y R. Moreno, Técnicas de lógica difusa aplicadas a la minería de datos, Scientia et Technica, 3(40), 1-6 (2008)
- Sarmiento, H.O., C. Isaza y T. Kempowsky, Estimación de estados funcionales en procesos complejos con base en agrupamiento difuso, Información Tecnológica, 2(24), 79-98 (2013)
- Sarmiento, H.O., Metodología para el establecimiento y ponderación automática de conexiones entre estados funcionales de un proceso como herramienta para el diagnóstico y la predicción de fallos, Tesis Doctoral, Departamento de Ingeniería, Universidad de Antioquia, Medellín, Colombia (2013)
- Uribe, C. y C. Isaza, Selección de variables guiada por conocimiento del experto para el monitoreo basados en datos de procesos industriales, Revista Facultad de Ingeniería Universidad de Antioquia, 65, 122-125 (2012)
- Xie, X.L. y G.A., Beni, Validity measure for fuzzy clustering, IEEE Trans. PAMI, 3(8), 841-846 (1991)

