



Detección de fallos de pequeña magnitud en sistemas industriales multimodo.

M. Quiñones Grueiro, A. Prieto-Moreno, O. Llanes-Santiago

Departamento de Automática y Computación, CUJAE, Habana, Cuba

RESUMEN / ABSTRACT

La detección de fallos de pequeña magnitud es uno de las cualidades que deben poseer las herramientas de diagnóstico basadas en datos históricos en procesos industriales multimodo. La existencia de múltiples modos de operación, las transiciones entre estos y el ruido presente en las mediciones de las variables complejiza la tarea de detectar este tipo de fallos lo cual se debe al reto que representa identificarlos datos que corresponden a cada régimen de funcionamiento estable. En este trabajo se propone un nuevo método para la identificación de modos de operación estables. Este se basa en una modificación del algoritmo de agrupamiento de las k-medianas que se ha denominado k-medianas incremental. Sus principales ventajas son la robustez ante la presencia de transiciones, datos fuera de rango (*outliers*) y ruido de tipo no gaussiano en los datos, no es necesario conocer la cantidad de modos de operación previamente para parametrizarlo y no requiere de múltiples corridas. A partir de los modos obtenidos se presenta el esquema de detección a utilizar. Para probar la efectividad de la estrategia propuesta se utilizó como proceso un tanque reactor-calentador continuamente agitado (CSTH por sus siglas en inglés). Los resultados obtenidos a partir de simular un fallo con diferentes magnitudes en cada modo de operación demuestran la efectividad de la propuesta realizada.

Palabras claves: procesos industriales, multimodo, detección de fallos, diagnóstico de fallos, algoritmo de agrupamiento.

The monitoring of small magnitude faults is a specification that should be accomplished by data-driven fault diagnosis techniques in multimode industrial processes. This is a challenging problem because it is difficult to identify the data of each stable operation mode considering the presence of the transitions between them and the noise of the measurements. A method for the identification of stable operation modes is proposed in this paper. This is based in a modification of the clustering algorithm known as k-medians and is named incremental k-medians. Its main advantages are: it is robust against the presence of transitions, outliers and noise in the dataset, it is not required to know the number of modes previously to parameterize it and it is not required to make multiple runs of it. The monitoring scheme is built upon the identified modes. The proposed strategy is tested using a continuous stirred tank heater (CSTH). Results obtained for a fault with different magnitude in each mode demonstrate the effectiveness of the proposal.

Key words: multimode, industrial processes, fault detection, fault diagnosis, clustering.

Monitoring of small magnitude faults in multimode industrial processes.

1. INTRODUCCIÓN.

La industria moderna tiene como necesidades principales la exigencia de altos niveles de productividad, el mantener los estándares de calidad fijados, garantizar la seguridad de los trabajadores y cumplir con las regulaciones medioambientales establecidas por las normas internacionales. La ocurrencia de averías tiene un impacto negativo en los procesos; pues ello puede provocar paradas no planificadas que interrumpen el ciclo productivo y accidentes que afecten la integridad del personal, que causen la degradación de equipos o violen las regulaciones medioambientales. Sea un fallo un comportamiento no deseado que puede llevar a una avería¹, el diagnóstico de fallos se convierte en requisito fundamental para complementar las tareas orientadas a la producción^{1,2}.

El diagnóstico de fallos se puede dividir en cuatro etapas fundamentales: detección, aislamiento, identificación y recuperación y para cada una se han desarrollado un conjunto de técnicas específicas¹. Estas pueden clasificarse en las basadas en modelos

cuantitativos, que agrupan métodos basados en el conocimiento; en modelos analíticos de los sistemas y en modelos estadísticos a partir de datos históricos³⁻⁴.

El desarrollo de la automática ha impulsado la utilización de sistemas de adquisición de datos en múltiples industrias. La gran cantidad de información que se genera con estos sistemas ha impulsado la utilización de las técnicas de diagnóstico de fallos basadas en datos históricos por su aplicabilidad a escala industrial.

Los fallos llamados de pequeña magnitud representan un reto en la etapa de detección de los sistemas de diagnóstico debido a que no provocan modificaciones drásticas en la dinámica de los procesos y a que en muchas ocasiones los sistemas de control enmascaran su presencia.

Si se tienen en cuenta además la característica multimodal de muchos procesos industriales, los puntos fuera de rango (*outliers*) y el ruido presente en las mediciones la tarea se complica aún más. Ello se debe a la complicación que presupone definir los regímenes estables de funcionamiento a monitorizar dada la existencia de transiciones entre cada modo, las cuales pueden producirse puesto que muchos procesos operan de forma continua. La característica multimodal está determinada por cambios en el modo de producción (referencias de los controladores), en las características de la materia prima o en las condiciones ambientales⁶⁻⁹.

La estrategia para la detección de fallos basada en múltiples modelos locales para cada modo en procesos multimodo ha sido adoptada en varios trabajos con buenos resultados⁷⁻⁹. Para su aplicación se debe contar primeramente con los datos referidos a cada modo de operación y luego definir un criterio para decidir en línea en qué modo se encuentra operando el proceso. Por ello una de las problemáticas existentes es cómo separar los datos de cada régimen cuando se encuentran mezclados con transiciones, *outliers* y ruido¹⁰.

En este artículo se presenta como aporte un novedoso algoritmo de agrupamiento que se ha denominado *k*-medias incremental. El mismo permite separar los datos correspondientes a los modos estables; es robusto ante la presencia de transiciones, datos fuera de rango (*outliers*) y ruido en las mediciones; no es necesario establecer la cantidad de modos a estimar previamente y no requiere de múltiples corridas. A partir de la utilización de los modelos obtenidos, al construir el esquema de detección se logran buenos resultados en la detección de fallos de pequeña magnitud simulados en cada modo.

El trabajo está organizado de la siguiente manera: la sección 2 presenta un resumen sobre la utilización de las técnicas de agrupamiento en la identificación de modos de operación y el algoritmo de agrupamiento propuesto. La estrategia de detección basada en múltiples modelos se explica en la sección 3. La sección 4 presenta el problema de prueba utilizado y los casos simulados. Los resultados experimentales y la discusión sobre los mismos aparecen en la sección 5. Finalmente, son presentadas las conclusiones.

2. Técnicas de agrupamiento. El algoritmo *k*-medias incremental.

2.1. Técnicas de agrupamiento para la identificación de modos estables.

Las técnicas de agrupamiento son métodos no supervisados que se utilizan para agrupar instancias de poblaciones de datos en diferentes conjuntos de acuerdo a medidas de similitud dentro de las mismas. Estas pueden clasificarse en función de los principios que siguen los algoritmos en cuatro categorías: basados en particionamiento, en jerarquía, en divisiones y en densidad¹¹.

Los primeros son los más populares y parten de definir la cantidad de grupos (*clusters*) a estimar, lo que está condicionado por el conocimiento a priori que se posea de las muestras. Los algoritmos más populares entre estos son las *K*-medias y sus extensiones¹². Esta familia de algoritmos se ha adaptado de diferentes formas para aplicarlos en la identificación fuera de línea de modos de operación en los procesos industriales⁶.

Cuando se pueden presentar transiciones en los datos de la planta la implementación clásica de estos no es capaz de separar correctamente los modos de operación. Por ello, se han desarrollado nuevas variantes basadas en la utilización de una ventana de tiempo deslizante¹³⁻¹⁵ teniendo en cuenta la correlación temporal que existe entre las variables en cada modo estable de operación. Aun así, en la mayoría de los casos se requieren de múltiples corridas de los algoritmos y establecer criterios para determinar la partición final.

En este trabajo se propone un método para la identificación de modos estables de operación llamado *k*-medias incremental que resulta fácil de parametrizar, y tiene como ventajas que no necesita que se definan por adelantado la cantidad de modos a estimar, no requiere de múltiples corridas ni de establecer criterios para determinar la partición final de los datos y además es robusto ante

la presencia de transiciones, datos fuera de rango (*outliers*) y ruido en los datos. A continuación se presenta su funcionamiento en detalles.

2.2 K-medias incremental.

El algoritmo de las k-medias se basa en el algoritmo de las k-medias^{16, 17}. Este último se basa en minimizar el error cuadrático, en términos de cierta medida de distancia, entre los puntos de cada grupo con su respectivo centroide¹⁸. El centroide se representa en este caso a partir del promedio de todos los puntos. Los puntos se reasignan iterativamente de acuerdo a la cercanía a cada centroide y este se recalcula en cada iteración hasta que el cambio en la ubicación de todos los centroides no varía.

El algoritmo de las k-medias es sensible a la presencia de *outliers* y ruido en los datos¹⁹, lo que es común en la industria. Para mejorar la robustez se han desarrollado variantes como PAM²⁰ (Particionamiento alrededor de los medoides) en las cuales el centro no se representa por el promedio sino por una observación, y una versión en la cual se utiliza la mediana²¹.

Para lograr una mejor adaptabilidad del algoritmo de las k-medias se propone el algoritmo llamado Seguidor-Líder en el cual los grupos se construyen secuencialmente según el grado de pertenencia de una nueva observación a cada grupo^{22, 23}. El método propuesto se basa en utilizar la adaptabilidad del algoritmo del Seguidor Líder con la robustez de emplear una medida de similitud entre grupos basada en la mediana, por ello se ha denominado k-medias incremental.

El primer parámetro a establecer en el algoritmo es la cantidad de observaciones *m* a tener en cuenta en la ventana temporal. Este debe seleccionarse de 3 a 5 veces mayor que la mayor constante de tiempo entre todas las variables del proceso²⁴ para así garantizar que se pueda capturar correctamente la dinámica del proceso. Luego se presenta secuencialmente cada muestra y se evalúa el grado de pertenencia ξ respecto al grupo actual según la siguiente medida de similitud

$$\xi = e^{-(\|x_{med}-x\|)^2} \quad (1)$$

donde ξ se encuentra entre 0 y 1 determinando la probabilidad de pertenencia al grupo, x_{med} es la mediana del grupo y x es la muestra a asignar. Si una muestra tiene un grado de pertenencia al grupo actual tal que $\xi > \xi_{lim}$ entonces se reinicializa la ventana temporal desde esta muestra hasta *m* observaciones más. El parámetro ξ_{lim} permite definir cuán definidos deseamos que queden los grupos resultantes.

Al reinicializar una ventana debe tenerse en cuenta que el proceso puede encontrarse en estado de transición, que la muestra actual es un *outlier* o el ruido ha provocado una variación en el valor de la observación. Para excluir las observaciones correspondientes a estos fenómenos de los grupos resultantes, puesto que se desea determinar los estados estables del proceso, se establece el parámetro de deslizamiento de la ventana m_d . Este define la cantidad de muestras consecutivas, a continuación de la ventana actual, que deben pertenecer al grupo actual para considerar que el mismo representa un estado estable. Si las transiciones en el proceso ocurren de forma rápida entonces su valor puede definirse entre el 10 % y 20 % del tamaño de la ventana.

Si al evaluar las siguientes m_d muestras alguna no pertenece al grupo actual entonces la ventana se desplaza en una muestra y la muestra desechada (primera de la ventana) se considera un *outlier* o una transición en caso de que sean varias consecutivas. Por último se determina la similitud entre los grupos resultantes utilizando las medianas de cada uno y se unifican aquellos que cumplan que $\xi < \xi_{lim}$. Esto se realiza dado el caso en que un *outlier* o el ruido en las mediciones pudieron haber provocado que un modo estable se separara en dos grupos. El pseudocódigo del método se resume en el algoritmo 1.

Finalmente cabe destacar que el algoritmo solo requiere una corrida, no necesita que se establezcan la cantidad de modos a estimar previamente y por último es capaz de identificar los modos de operación estables aun cuando en los datos se mezclen transiciones, *outliers* o estén contaminados por ruido.

3. Estrategia de detección basada en múltiples modelos.

La estrategia que se utiliza en este trabajo para construir el esquema de monitorización parte de emplear los datos de los modos estables identificados. Primeramente se determina para cada modelo local la media de los datos μ_k y la matriz de covarianza Σ_k que los caracteriza. Luego se transforman las variables originales de cada modo de operación por separado utilizando Análisis de Componentes Principales (PCA). Esta técnica para la reducción de dimensión se basa en transformar las variables según la ecuación (2) de forma que en el nuevo espacio queden las componentes principales que representen la mayor variabilidad del conjunto original¹. Para este trabajo la cantidad de componentes retenidas (a) se tomaron de forma que representaran el 80 % de la varianza total.

$$T = XP(2)$$

Donde P es la matriz de transformación que contiene los a autovectores asociados con los primeros a autovalores de la matriz de covarianza de los datos, X es la muestra de datos originales y T es el espacio reducido de dimensión a .

ALGORITMO 1. K-medianas incremental.

```

{m: Número de muestras que definen un modo estable,  $m_d$ : Número de muestras consecutivas que determinan la existencia de un
nuevo modo estable,  $\xi_{lim}$ : Probabilidad que define la pertenencia a un grupo, n: cantidad de muestras, C: grupo, x: muestra, i: número
de grupo, j: número de muestra, W: ventana actual, X: datos}
{Inicializar W con m muestras, i=1, j=1}
repeat
 $x_d \leftarrow$  Calcular Mediana( $C_i$ )
 $\xi_j \leftarrow$  Calcular Similaridad( $W, x_j$ )
if ( $\xi_j > \xi_{lim}$ ) then
     $i = i + 1$ ;
    Reiniciar W;
     $j = j + m$ ;
    While  $result == false$  and  $j < n$ 
        Desplazar W en una muestra;
         $j = j + 1$ ;
         $result \leftarrow$  Calcular Similaridad Muestras Continuas( $W, X, j, m_d$ );
    end while
else
 $C_i \leftarrow$  Añadir( $x_j$ );
    Desplazar W en una muestra;
     $j = j + 1$ ;
endif
until  $j \geq n$ 
    
```

A continuación, a partir de utilizar un número reducido del conjunto total de componentes principales, se construyen los límites de control basados en los estadísticos T^2 y SPE (*Squared Prediction Error*), también designado como Q, para cada modo de operación. El cálculo de estos límites se realizó como lo explica Chiang et al.¹ considerando desconocidas las verdaderas medias y covarianzas de la población de las variables y utilizando $\alpha = 0.05$. Esta etapa se resume en la Fig.1.

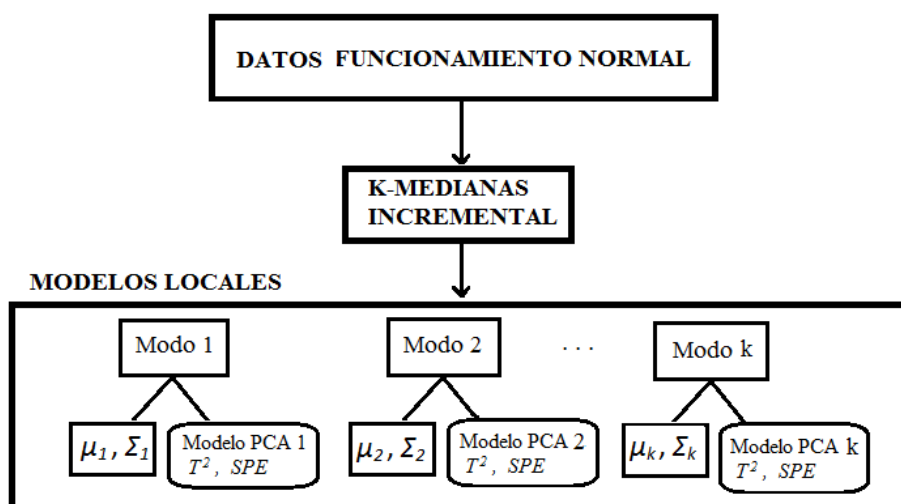


Figura 1: Etapa fuera de línea de la estrategia propuesta.

Una vez establecidos los límites de control para cada modo de operación la detección en línea se realiza de la siguiente forma. Primeramente se supone un modo de operación. Si alguno de los estadísticos viola los límites de control entonces se evalúa la observación actual respecto a los otros modos. Si el valor de alguno de los estadísticos se mantiene fuera de los límites de control durante cinco muestras consecutivas entonces se confirma la presencia de un fallo en el comportamiento del proceso.

4. Problema de prueba y casos simulados.

El problema de prueba utilizado es el tanque calentador continuamente agitado (*Continuous Stirred Tank Heater* CSTH por sus siglas en inglés) desarrollado por Thonhill et al²⁵ y empleado en numerosos artículos para probar alternativas de técnicas de diagnóstico de fallos²⁶⁻²⁸. La simulación de esta planta, caracterizada por su alta no linealidad, mezcla un modelo matemático con datos experimentales obtenidos de la planta real. El ruido de tipo no gaussiano presente en las variables se introduce a partir de mediciones reales.

Un esquema simple de la planta se muestra en la Fig. 2. En el tanque calentador se mezclan bien agua caliente y fría de forma que la temperatura del tanque se presume equivalente a la del flujo de salida. La mezcla es calentada dentro del tanque a través de un serpentín empleando vapor y drenada a través de una larga tubería. Las variables de entrada son la posición de las válvulas de agua fría, agua caliente y vapor. Las variables de salida son la temperatura de la mezcla, el nivel de agua y el flujo de agua fría. Todas las mediciones se obtienen en un rango de 4-20 mA y por tanto la escala de los gráficos mostrados se encuentra en ese rango.

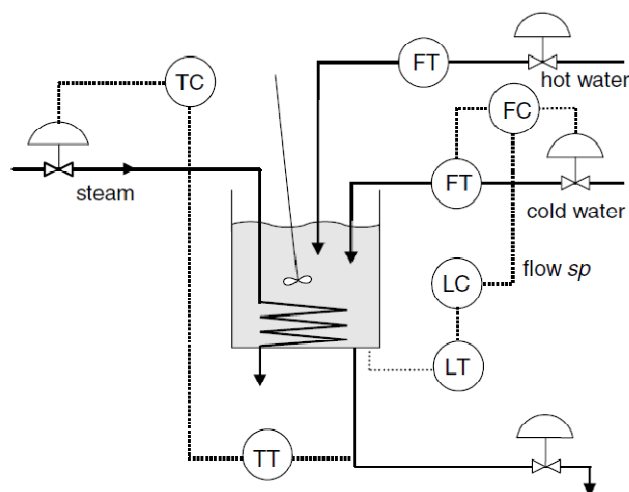


Figura 2: Diagrama de flujo del tanque calentador continuamente agitado.

El cambio en los modos de operación está dado a partir de variaciones en la apertura de la válvula de agua caliente y en las referencias del nivel y la temperatura del agua en el tanque. En la Tabla 1 se muestran los cinco modos de operación establecidos para la simulación. El tiempo de muestreo es de 1s y la simulación de cada modo dura 400 muestras. El cambio entre modos se realiza de forma continua y las variables monitorizadas son el flujo de agua fría, el nivel y la temperatura, pues estas son las que se miden en línea.

Variable (mA)	Modo 1	Modo 2	Modo 3	Modo 4	Modo 5
Nivel	12	12	12	13	13
Temperatura	10.5	10.5	11	10.5	10.5
Apertura de la válvula de agua caliente	0	5.5	5.5	5.5	5

Tabla 1: Modos de operación simulados.

Para determinar la capacidad de detectar fallos de pequeña magnitud se simuló un fallo de tipo desviación (bias) en la medición del sensor que mide el nivel del tanque. La desviación se estableció de forma que se produce una disminución en el nivel. Se establecieron tres amplitudes diferentes para el fallo: Mediano (-0.5), pequeño (-0.3) y muy pequeño (-0.1). Este se introdujo pasadas 100 muestras en cada modo de operación y se encuentra presente durante otras 100 muestras. La calidad en la detección del fallo se mide en función de tres parámetros: el índice de falsas alarmas (FAR), el índice de detección de fallo (FDR) y la demora en la detección del fallo (DD). El primero define la cantidad de muestras capturadas en funcionamiento normal que se clasificaron como fallo y el segundo determina la cantidad de muestras capturadas durante la ocurrencia del fallo que no fueron clasificadas como tal.

5.RESULTADOS Y DISCUSIÓN.

Primeramente se presenta la identificación fuera de línea de los modos de operación y a continuación los resultados obtenidos en la detección para el fallo de pequeña magnitud.

5.1 Identificación de los diferentes modos de operación.

En la Fig.3 se muestran los resultados de la simulación de los cinco modos de operación establecidos de forma continua. Como puede observarse las transiciones se encuentran entre los modos simulados. En dependencia del cambio de modo de operación las transiciones tienen diferentes características y duración. Todo ello complica la separación de los modos de operación de forma que es prácticamente imposible lograrlo empleando algoritmos de agrupamiento clásicos como las *k-medias* o GMM (*Gaussian Mixture Models* por sus siglas en inglés).

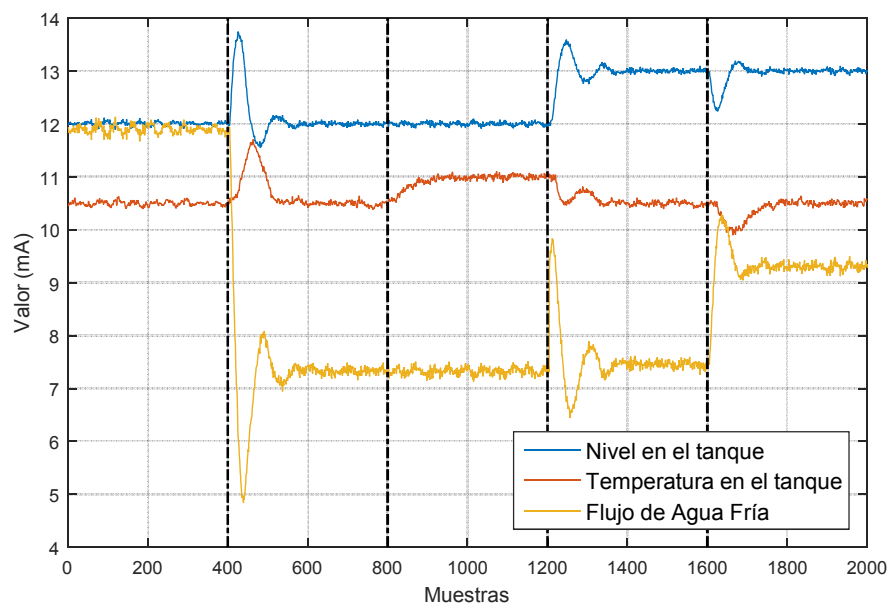


Figura 3: Respuesta temporal de las variables monitorizadas al atravesar los cinco modos de operación.

Al utilizar el algoritmo propuesto (k-medias incremental) se seleccionó primeramente un ancho para la ventana de tiempo m y la ventana de deslizamiento w_d iguales a 40 y 10 muestras respectivamente a partir de analizar los tiempos de respuesta de las variables del proceso mostrados en Thornhill et al²⁵. Con el objetivo de lograr la mejor separabilidad se estableció un umbral de 0.95 para la probabilidad de pertenencia de las muestras a cada grupo encontrado.

Los resultados obtenidos se muestran en la Fig.4. Como se muestra se detectan cinco modos de operación y las transiciones son separadas en todos los casos con excepción del cambio entre el modo 2 y 3. Lo que está dado porque la transición se realiza de forma muy lenta y en un tiempo relativamente corto. A pesar de ello, la influencia en la detección no resulta ser significativa.

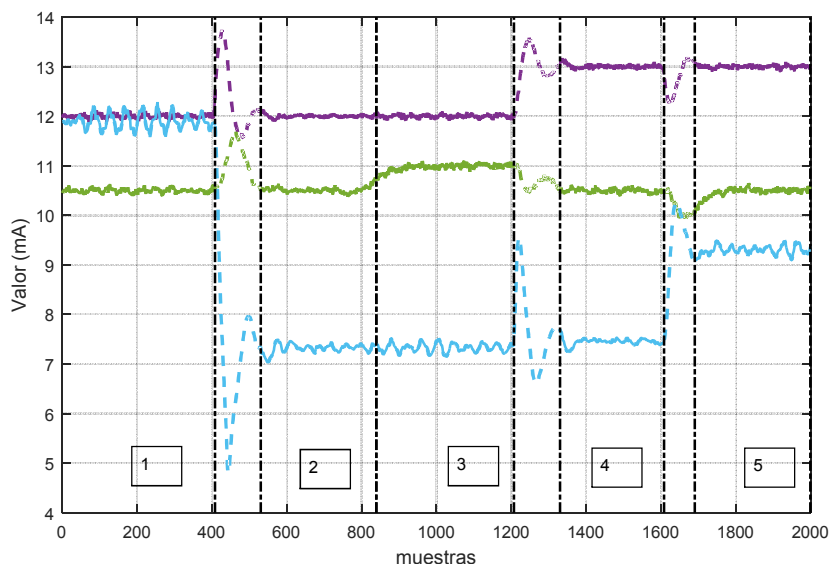


Figura 4: Separación de los cinco modos estables al utilizar el algoritmo k-medianas incremental.

5.2 Resultados en la detección del fallo con diferentes magnitudes..

Una vez identificados los modos de operación según el algoritmo k-medianas incremental se procede a transformar los datos de cada modo por separado según PCA y construir los límites de control para cada modo. Se va a realizar la detección de fallos según la estrategia establecida en la sección 3. Para todos los modelos PCA se utilizó como criterio para la selección de las componentes el 80% de la información.

El resumen de los resultados obtenidos se muestra en la Tabla 2. Los índices FAR y FDR se muestran en porcentaje, mientras que la demora en la detección (DD) en cantidad de muestras.

Modo	Fallo Mediano			Fallo Pequeño			Fallo Muy Pequeño		
	FAR	FDR	DD	FAR	FDR	DD	FAR	FDR	DD
1	5	94	6	5	93	7	5	61	17
2	3	93	7	3	92	8	3	80	11
3	1	93	7	1	91	9	1	64	22
4	0	93	7	0	93	7	0	70	13
5	6	94	6	6	93	7	6	78	11

Tabla 2: Resultados obtenidos para los fallos de diferente magnitud.

Como puede observarse para el fallo mediano y pequeño la tasa de detección se mantiene por encima del 90 % con una tasa de falsas alarmas entre 0 % y 6 %. Ello confirma que la calidad de los modelos PCA que se definieron para cada estado de operación es buena.

El método de detección de modos de operación permite capturar la dinámica de cada modo con suficiente precisión aun cuando se encuentran transiciones mezcladas entre estos y ruido de tipo no gaussiano capturado de las mediciones reales. La tasa de falsas alarmas se debe principalmente a que el ruido presente en el conjunto de datos de validación introduce variaciones en la dinámica que no fueron capturadas durante la construcción de los modelos.

Para los casos donde el fallo es de muy pequeña magnitud se observa un FDR entre 61 % y 80 %. En la Fig.5 se observa por ejemplo que existe muy poca variación en la dinámica del proceso cuando este fallo aparece en el modo 1, que es el caso con peores resultados. Al ser la magnitud del fallo en este escenario extremadamente pequeña se torna muy difícil detectarlo. Por ello, aunque el FDR no supera el 80 % y la demora en la detección (DD) es mayor que en los otros casos, se consideran buenos resultados.

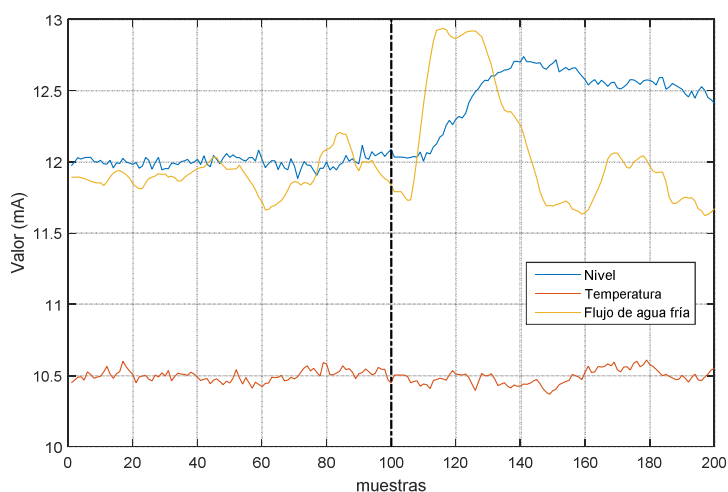


Figura 5: Respuesta temporal al fallo de mediana magnitud que ocurre partir de la muestra 100.

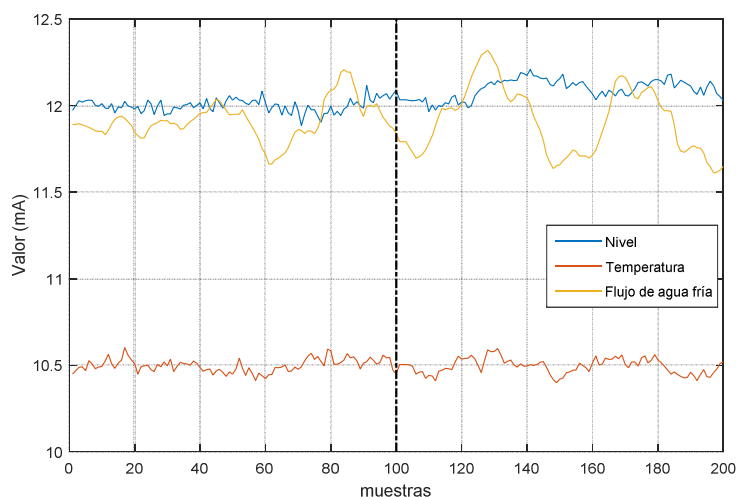


Figura 6: Respuesta temporal al fallo de muy pequeña magnitud que ocurre partir de la muestra 100.

6. CONCLUSIONES.

En este artículo se presentó un método para la detección de modos en sistemas industriales con múltiples modos de operación. Con su utilización es posible detectar fuera de línea los modos estables aun cuando se presenten transiciones en los datos de diferente característica y duración y además es robusto ante la presencia de *outliers*. El método llamado k-medianas incremental se basa en una modificación del algoritmo de las k-medianas empleando una ventana temporal deslizable. Además tiene como ventajas que no se requiere conocer por adelantado la cantidad de modos en que opera el proceso, tampoco se necesita realizar varias corridas del algoritmo para llegar al resultado final y la capacidad de lidiar con ruido de tipo no gaussiano presente en las mediciones. Como limitante principal puede mencionarse que se debe conocer aproximadamente los tiempos de respuesta de las variables en función de definir el ancho de la ventana de tiempo.

La estrategia de detección basada en múltiples modelos demostró ser efectiva para detectar fallos de pequeña magnitud, teniendo en cuenta que se utilicen modelos de caractericen con alta precisión la dinámica de cada modo de operación. Los índices de FAR fueron inferiores al 7 % en todos los experimentos realizados lo que demuestra la alta confiabilidad de la estrategia propuesta. Puede destacarse que aun para fallos de muy baja magnitud el FDR se mantuvo por encima del 50 % y la demora en la detección no superó las 22 muestras. Futuras investigaciones pueden enfocarse en desarrollar una estrategia para la detección de los fallos durante las transiciones entre modos de operación.

8. REFERENCIAS.

- [1] L. H. CHIANG, E. RUSELL and R. D. BRAATZ, *Fault Detection and Diagnosis in Industrial Systems*. Springer-Verlang. London, England, 2001.
- [2] R. ISERMANN and P. BALL, "Trends in the application of model based fault detection and diagnosis of technical processes" en Proc. of the 19th IFAC World Congress, (Piscataway, New Jersey), pp. 1–12, IEEE Press, 1996.
- [3] V. VENKATASUBRAMANIAN, R. RENGASWAMY, and K. YIN, "A review of process fault detection and diagnosis Part I: Quantitative model-based methods," en *Chemical Engineering*, vol. 27, pp. 293–311, 2003.
- [4] V. VENKATASUBRAMANIAN, R. RENGASWAMY, and S. N. KAVURI, "A review of process fault detection and diagnosis Part II: Qualitative models and search strategies," en *Chemical Engineering*, vol. 27, pp. 313–326, 2003.
- [5] V. VENKATASUBRAMANIAN, R. RENGASWAMY and S. N. KAVURI, "A review of process fault detection and diagnosis Part III: Process history based methods," en *Chemical Engineering*, vol. 27, 2003.
- [6] D.-H. HWANG and C. HAN, "Real-time monitoring for a process with multiple operating modes," en *Control Engineering Practice*, vol. 7, pp. 891–902, 1999.
- [7] M. MAESTRI, A. FARALL, P. GROISMAN, M. CASSANELLO, and G. HOROWITZ, "A robust clustering method for detection of abnormal situations in a process with multiple steady-state operation modes," en *Analysis*, vol. 34, pp. 223–231, 2010.
- [8] M. M. RASHID and J. YU, "Hidden Markov Model Based Adaptive Independent Component Analysis Approach for Complex Chemical Process Monitoring and Fault Detection," en *Industrial & Engineering Chemistry Research*, 2012.
- [9] X. WANG, X. WANG, Z. WANG, and F. QIAN, "A novel method for detecting processes with multi-state modes," en *Control Engineering Practice*, vol. 21, no. 12, pp. 1788–1794, 2013.
- [10] H. MA, Y. HU and H. SHI, "A novel local neighborhood standardization strategy and its application in fault detection of multimode processes," en *Chemometrics and Intelligent Laboratory Systems*, vol. 118, pp. 287–300, 2012.

- [11] J. HAN, M. KAMBER, and J. PEI, *Data Mining: Concepts and Techniques*. Ed. Elsevier. Wyman Street, Waltham, USA, 2012.
- [12] A. K. JAIN, “Data clustering: 50 years beyond K-means,” en *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.
- [13] F. WANG, S. TAN, J. PEN and Y. CHANG, “Process monitoring based on mode identification for multi-mode process with transitions” en *Chemometrics and Intelligent Laboratory Systems*, vol. 110, pp. 144–155, 2012.
- [14] Z. ZHU, Z. SONG and A. PALAZOGLU, “Transition Process Modeling and Monitoring Based on Dynamic Ensemble Clustering and Multiclass Support Vector Data Description” en *Industrial & Engineering Chemistry Research*, pp. 13969–13983, 2011.
- [15] Z. ZHU, Z. SONG and A. PALAZOGLU, “Process pattern construction and multi-mode monitoring” en *Journal of Process Control*, Vol. 22, pp. 247–262, 2012.
- [16] E. FORGY, “Cluster analysis of multivariate data: efficiency vs. interpretability of classifications” en *Biometrics*, vol. 21, pp. 768 – 780, 1965.
- [17] J. MACQUEEN, “Some methods for classification and analysis of multivariate observations” en *Proceedings of the Fifth Berkeley Symposium*, Vol. 1, pp. 281 – 297, 1967.
- [18] M. J. ZAKI and M. J. WAGNER, *Data mining and analysis*. New York: Cambridge University Press, 2014.
- [19] XU, R. and WUNSCH, D. C. *Clustering*. Ed. IEEE PRESS, 2009.
- [20] KAUFMAN, L. and ROUSSEEUW, P.; *Finding groups in data: An introduction to cluster analysis*. New York, NY: John Wiley & Sons, 1990.
- [21] V. ESTIVILL - CASTRO and J. YANG, “A fast and robust general purpose clustering algorithm” en *Proceedings 6th Pacific Rim International Conference on Artificial Intelligence PRICAI 2000*, R. Mizoguchi and J. Slaney, Eds., *Lecture Notes in Artificial Intelligence 1886*, New York, NY: Springer - Verlag, pp. 208 – 218, 2000.
- [22] R. DUDA, P. HART and D. STORK, *Pattern classification*, 2nd edition. New York: John Wiley & Sons, 2001.
- [23] B. MOORE, “ART1 and pattern clustering” en *Proceedings of the 1988 Connectionist Models Summer School*, pp. 174 – 185, 1989.
- [24] L. VAZQUEZ S., A. TRUJILLO C., Y. LLOSAS A. and D. DÍAZ M., “Método para detección de estados estacionarios: aplicación a unidades de generación eléctrica” en *RIELAC*, Vol. XXXV, No. 2, pp. 45-61, 2014.
- [25] N. F. THORNHILL, S. C. PATWARDHAN and S. L. SHAH, “A continuous stirred tank heater simulation model with applications” en *Journal of Process Control*, Vol. 18, pp. 347-360, 2008.
- [27] C. NING, M. CHEN and D. ZHOU, “Hidden Markov Model-Based Statistics Pattern Analysis for Multimode Process Monitoring: An Index-Switching Scheme” en *Industrial & Engineering Chemistry Research*, Vol. 53, pp. 11084-11095, 2014.
- [28] X. XU, L. XIE and S. WANG, “Multimode process monitoring with PCA mixture model” en *Computers and Electrical Engineering*, Vol. 40 No. 7, pp. 2101-2112, 2014.

AUTORES

Marcos Quiñones Grueiro. Ingeniero Automático. Ingeniero. CUJAE, La Habana, Cuba. marcosq@electrica.cujae.edu.cu

Alberto Prieto-Moreno. Ingeniero Automático. Doctor en Ciencias Técnicas. CUJAE, La Habana, Cuba.
albrieto@electrica.cujae.edu.cu

Orestes Llanes Santiago. Ingeniero Electricista. Doctor en Ciencias Técnicas. CUJAE, La Habana, Cuba.
orestes@electrica.cujae.edu.cu