

A proposal for indirect monitoring of the anaerobic digestion process using wavelet transform

Jarvio-Hernández Yerdey E.* Moguel-Castañeda Jazael G. ** Puebla H. ** Hernandez-Martinez E.*

**Facultad de Ciencias Químicas, Universidad Veracruzana, Xalapa Veracruz, México.
(e-mail: elisehernandez@uv.mx)*

***Departamento de Energía, Universidad Autónoma Metropolitana-Azcapotzalco, Av. San Pablo 180, México,
D.F. 02200. (e-mail: jzael_moguel@hotmail.com; hpuebla@uam.azc.mx)*

Abstract: The monitoring of key variables in physical, chemical, and biological processes represents important information for analyzing and/or evaluating performance and stability of processes. However, it is not always possible to obtain key variables measurements (e. g. reagent and product concentrations) on real time, therefore, indirect monitoring, using easily measured variables, is a useful alternative. In this paper a study of wavelet transform application for pH time series analysis as an alternative for indirect monitoring of anaerobic digestion is presented. Two case studies, tequila vinasse and cheese whey, were analyzed, determining that wavelet transform can be used as a tool for qualitative monitoring of the anaerobic digestion process operated at different operating conditions.

Keywords: Time series, Wavelet transform, Monitoring, Data mining tools, Bioenergy production

1. INTRODUCCIÓN

La digestión anaerobia (DA) es un proceso ampliamente usado para el tratamiento de efluentes con altas concentraciones de materia orgánica, ya que además de reducir la carga orgánica es posible generar productos de valor agregado como ácidos grasos volátiles (AGV) y biogás. A pesar de las ventajas que ofrece la DA, su operación no es una tarea fácil, principalmente debido consta de cuatro etapas consecutivas conocidas como, hidrólisis, acidogénesis, acetogénesis y metanogénesis, donde cada etapa está ligada a un consorcio bacteriano con necesidades específicas. Lo anterior hace necesario mantener las condiciones de operación y variables del proceso dentro de rangos definidos por la naturaleza biológica del sistema para garantizar un desempeño satisfactorio de las cuatro etapas (Micolucci et al., 2016; Chu y Wang 2017). Entonces, el seguimiento dinámico de las variables clave del proceso como la demanda química de oxígeno (DQO), AGV y biogás es de vital importancia para asegurar la calidad de los productos y seguridad del proceso. Sin embargo, para llevar a cabo esta tarea se requiere de técnicas analíticas y de equipos especializados de alto costo que dificultan el seguimiento del proceso con rapidez. Esto provoca el retraso en la toma de acciones en la operación del proceso.

Por otro lado, se han propuesto metodologías para el monitoreo de procesos basadas en la comprensión de datos y la información estadística, por ejemplo, el análisis de componentes principales (ACP) para la supervisión de procesos; la aplicación de filtros de Kalman para la estimación de variables de estado (Vaishnavi y Muthumari, 2018) o el uso de modelos ARIMA y sistemas de regresión para la predicción de dinámicas de procesos no lineales

(Pradhan et al., 2016; Molina et al., 2018). No obstante, estas metodologías no son capaces de detectar los cambios relacionados con los diferentes fenómenos que interactúan en los procesos (reacciones y/o fenómenos de transporte), así como las escalas de tiempo características en las que se presentan. Otros métodos proponen la aplicación de observadores de estado que permiten reconstruir en línea las variables difíciles de medir como concentraciones de componentes basándose en un modelo de proceso y en algunas mediciones disponibles (Dochain, 2003; Bogaerts y Wouwer, 2003). El buen desempeño y aproximación de esta metodología depende del conocimiento y robustez del modelo matemático del sistema, aunado a esto, suelen ser sensibles ante incertidumbres.

Por otra parte, las metodologías basadas en el análisis multiescala permiten extraer información intrínseca e inferir cualitativamente el comportamiento de variables clave. Estas metodologías se han aplicado a imágenes y series de tiempo de variables que afectan al proceso y son de fácil medición como la temperatura y pH, donde la información obtenida puede ser usada para el monitoreo indirecto de procesos. Como alternativas para el monitoreo de la DA, se han aplicado técnicas multiescala de series de tiempo de fácil medición permitiendo dar seguimiento a los productos principales (Méndez-Acosta et al., 2013; García-Solano et al., 2015). Por ejemplo, en Méndez-Acosta et al. (2013) se realizó el análisis multiescala de series de tiempo de pH de la DA de vinazas tequileras, obteniendo correlaciones entre los índices multiescala y las variables clave del proceso (i.e., DQO, AGV y el metano) las cuales se obtuvieron gracias a la identificación de tres escalas de tiempo características asociadas a cada una de las etapas del proceso de digestión (i.e., hidrólisis, acidogénesis-acetogénesis y metanogénesis).

Así mismo García-Solano et al. (2015) realizaron el análisis de series pH en un digestor anaerobio continuo en la etapa de arranque, logrando detectar desestabilizaciones de esta etapa mediante los índices multiescala. Más adelante, Sánchez-García et al. (2018) presentaron el análisis multiescala R/S de un digestor anaerobio de flujo ascendente para la degradación de lactosuero crudo. Encontrando que los índices multiescala obtenidos de series de tiempo de pH, presentan correlaciones que pueden describir de forma cualitativa las variables clave que se adquieren fuera de línea (i.e., DQO, AGV y producción de biogás).

Otra metodología multiescala ampliamente usada es la transformada de Wavelet cuya ventaja sobre otros análisis de series de tiempo es que el tamaño de la ventana no es fijo, sino que varía en función de la frecuencia sin necesidad de manipular excesivamente la señal original, obteniendo información sobre el tiempo y la frecuencia. En el área de procesos se ha aplicado la transformada de wavelet para el estudio de comportamientos dinámicos (Tahmasebpoor et al., 2015; Bai et al., 2019), por ejemplo, Bai et al. 2019 emplearon esta metodología para el estudio de los efectos hidrodinámicos de la permeación de gas en diferentes partes de un reactor de membrana de lecho fluidizado. Los resultados mostraron correlaciones entre los comportamientos del cambio de la energía de wavelet y la evolución del tamaño de las burbujas a lo largo de la altura del lecho. Además, Prado et al. (2019) realizan el análisis de multifractalidad y coherencia de wavelet en series de tiempo de corriente eléctrica del ánodo de un reactor de tipo pila para la generación de combustible. Sus resultados mostraron información sobre las biopelículas adjuntas a los ánodos del reactor relacionadas a su estabilidad.

Lo anterior sugiere que el análisis de información basado en la transformada de wavelet puede ser utilizada para el monitoreo de las variables clave de procesos físicos, químicos y biológicos como la DA. En ese sentido, en este trabajo se propone el estudio de las variables clave de la digestión anaerobia mediante la aplicación de la transformada wavelet a diferentes series de pH. Se usan dos casos de estudio, lactosuero y vinaza tequilera como sustratos, para evaluar el potencial de la transformada wavelet como sistema de monitoreo indirecto, encontrando que es posible identificar correlaciones entre los coeficientes wavelet multiescala (conocidos también como coeficientes de detalle) y la dinámica de DQO y AGV de los procesos de DA analizados.

2. MATERIALES Y MÉTODOS

2.1 Sistema experimental

2.1.1 Digestión anaerobia de vinaza tequilera

En la DA de vinaza tequilera el sustrato proviene de una fábrica de tequila proveniente de La Laja, México y se utiliza

lodo anaerobio como inóculo. Se utilizó un reactor anaerobio secuencial discontinuo (ASBR), con capacidad de 5.1 l y se realizó la medición en línea del pH cada 2 segundos durante 72 horas de operación, obteniendo 116055 datos mediante un sistema de adquisición de datos de National Instruments cRIO9004. La operación se llevó a cabo a condiciones de temperatura de $T=38^{\circ}\text{C}$ y $\text{pH}=8$. En la Fig. 1a se tienen las fluctuaciones dinámicas de pH adquiridas en línea alrededor de la referencia ($\text{pH}=8$), donde las fluctuaciones al inicio son mayores debido al arranque del proceso y posteriormente se pueden ver discontinuidades en diferentes momentos de la serie de tiempo que son debidas a las acciones del control on-off de pH. Las mediciones de DQO se obtuvieron por el método colorimétrico de reflujo cerrado mediante el digestor DBR200 de HACH y el espectrofotómetro DR2800. Para los AGV se utilizó un HPLC Waters 600 utilizando un detector de 2487-Absorbance. El total de AGV se reporta como concentración de ácido acético (Hac). Finalmente, la composición del biogás se determinó por cromatografía de gases utilizando un Perkin Elmer AutoSystem XL con un detector TCD y un Alltech colon Hayesep D 100/120 (Méndez-Acosta et al., 2013)

2.1.1 Digestión anaerobia con lactosuero

En la DA de lactosuero, el sustrato proviene de una lechería comunitaria ubicada en Acajete Veracruz, México y se utiliza lodo anaerobio como inóculo. El proceso de digestión se llevó a cabo en un reactor anaerobio secuencial discontinuo (ASBR), con un volumen de 3.5 l, además se obtuvo una medición de pH cada segundo durante 120 horas de operación, obteniendo 432000 datos mediante un sistema de adquisición de datos de National Instruments cRIO9074 (Fig. 1b).

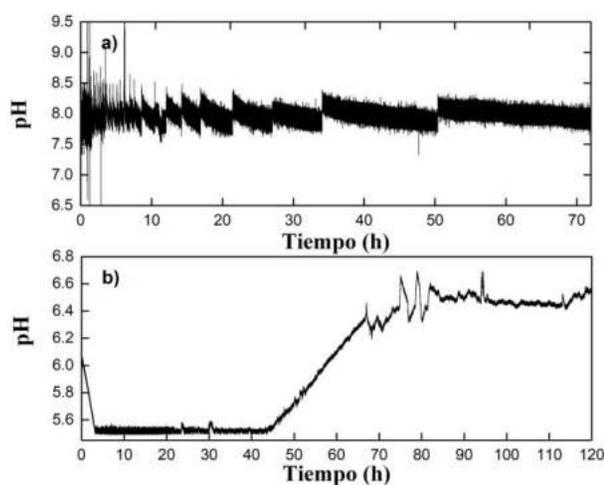


Fig. 1. Series de tiempo de pH, a) DA de vinaza tequilera y b) DA de lactosuero a $T=35^{\circ}\text{C}$ y $\text{pH}=5.5$.

La operación se llevó a cabo en condiciones de temperatura de T=35 y 40 °C y un pH de inicio de 5.5, ya que, el control fue establecido para evitar la acidificación del sistema, es decir, que el pH no fuera menor a 5.5. En cambio, el incremento de pH no fue regulado, por ello es posible observar un aumento 5.5 a 6.5. Las mediciones de DQO se obtuvieron por la técnica estándar micro semiautomatizado. A partir de un cromatógrafo de gases Agilent (modelo 7820 A) equipado con una columna capilar y un detector de ionización de llama se obtuvieron las concentraciones de ácidos orgánicos de cadena corta y se utilizó un factor de conversión para obtener la concentración de VFA (Lara-Musule et al., 2021).

2.2 Análisis de las series tiempo transformada de wavelet

El análisis de la serie de tiempo se realizó mediante la transformada de wavelet discreta (TWD). La TWD emplea ventanas de tiempo extensas para las frecuencias bajas y ventanas de tiempo cortas para frecuencias altas, esto resulta en un análisis de tiempo-frecuencia. La serie de tiempo analizada se descompone mediante la TWD mediante un filtrado sucesivo de dos tipos, pasa altas y pasa bajas y dos muestreos descendentes, cada filtro otorga a su vez una escala de análisis de base 2. El filtro pasa bajas (h_ϕ) es su versión en espejo, mientras que el filtro pasa altas (h_ψ) es una ondícula madre discreta. La salida de los filtros pasa bajas son los coeficientes de aproximación (CA_i), los cuales son utilizados para realizar una reconstrucción de la señal original obteniendo una función aproximada a la original representados. La salida de los filtros pasa altas son los coeficientes wavelet de detalle (CD_i), los cuales indican la proporción con la que la wavelet madre se ajusta o superpone sobre la señal original.

Específicamente, la dinámica de los CD_i varía en el dominio tiempo-frecuencia con tendencias similares a las variables clave del proceso de digestión. Por ello, en este trabajo se propone usar los CD_i como una variable indirecta para el monitoreo de la tendencia de las variables claves de la DA. El número de filtros especifica el nivel de descomposición, siendo la longitud de CA_i y CD_i igual a la mitad de tamaño de la serie de tiempo por cada filtro (Subasi, 2007). Este análisis multinivel de la TWD se puede observar en la Fig. 2.

La aplicación del análisis multiescala de la TWD se basa en el uso de diferentes formas y tamaños llamadas funciones wavelets. Cada transformada se representa de la siguiente forma:

$$\psi_{ab} = (1/\sqrt{a})\Psi(t-b)/a \quad (1)$$

donde a el parámetro de dilatación, b el parámetro de traslación y Ψ la wavelet madre.

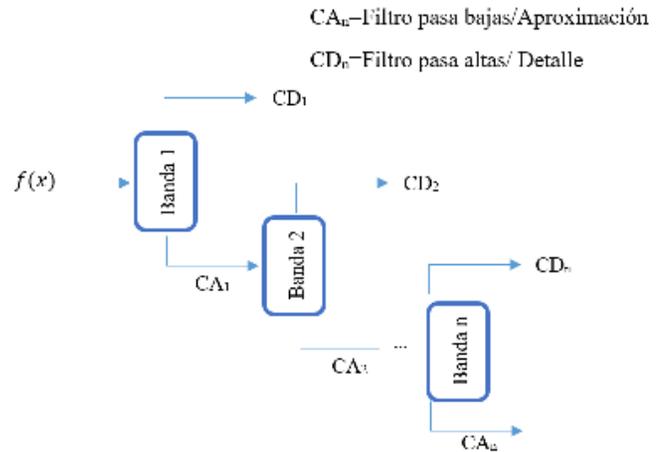


Fig. 2. Diagrama de niveles de descomposición de la TWD

En este estudio las funciones de wavelet madre que se emplearon fueron la de tipo dmey (Meyer discreta) para la vinaza tequilera y wavelet tipo rbior 2.8 (reverse biorthogonal) para las series de lactosuero. Para la selección de la wavelet madre se utilizó como criterio el mejor grado de ajuste dinámico de los coeficientes de detalle con la dinámica experimental de cada caso de estudio y que la escala de análisis correspondiente al nivel de descomposición tuviera relación fenomenológica con el proceso.

Para obtener los CA_i y CD_i se debe seleccionar el nivel de descomposición de la serie de tiempo. Por cada nivel de descomposición se tiene un filtro pasa bajas donde se aplica la función de escalamiento ($\phi_{j,k}$) y un filtro pasa altas donde se aplica la función wavelet descritas ($\psi_{j,k}$) estas funciones se definen como sigue (Mallat, 1989):

$$\phi_{j,k}(n) = 2^{j/2} \phi(2^j n - k) \quad (2)$$

$$\psi_{j,k}(n) = 2^{j/2} \psi(2^j n - k) \quad (3)$$

donde j denota el nivel de la escala y k denota la traslación en ese nivel de la escala. $n=0, 1, 2, \dots, M$; $j=0, 1, 2, \dots, J$; $k=0, 1, 2, \dots, 2^j-1$; $J=\log_2(M)$ y M es la longitud de la serie de tiempo. Los coeficientes de la TWD se obtienen del producto punto de la serie de tiempo original $x(n)$ y las funciones base. Los CA y CD en el i -ésimo nivel se definen como sigue:

$$CA_i = (1/\sqrt{M}) \sum_n x(n) \times \phi_{j,k}(n) \quad (4)$$

$$CD_i = (1/\sqrt{M}) \sum_n x(n) \times \psi_{j,k}(n) \quad (5)$$

Los resultados encontrados se obtuvieron de la comparación de diferentes niveles de descomposición con los datos experimentales de los dos procesos de digestión anaerobia, aplicándose un suavizado en los coeficientes obtenidos para observar de mejor manera la relación existente entre los datos

obtenidos y las dinámicas experimentales. Para el cálculo de los CA y CD se utilizó el toolbox wavelet de la plataforma de Matlab 2017b.

3. RESULTADOS

3.1 Análisis con la transformada de wavelet para la DA de vinaza tequilera

A partir del análisis de la serie de tiempo de pH mediante la TWD se obtuvieron coeficientes wavelet de detalle (CD_i) de diferentes bandas de descomposición. En la Fig 3. se muestra los coeficientes wavelet de detalle obtenidos. La descomposición de la serie va desde el nivel 1 hasta el nivel 8, donde los niveles correspondientes a escalas mayores a 0.02h (i.e. CD₅₋₈) tienen cambios dinámicos que pueden ayudar a la identificación de variaciones en el comportamiento de las variables clave.

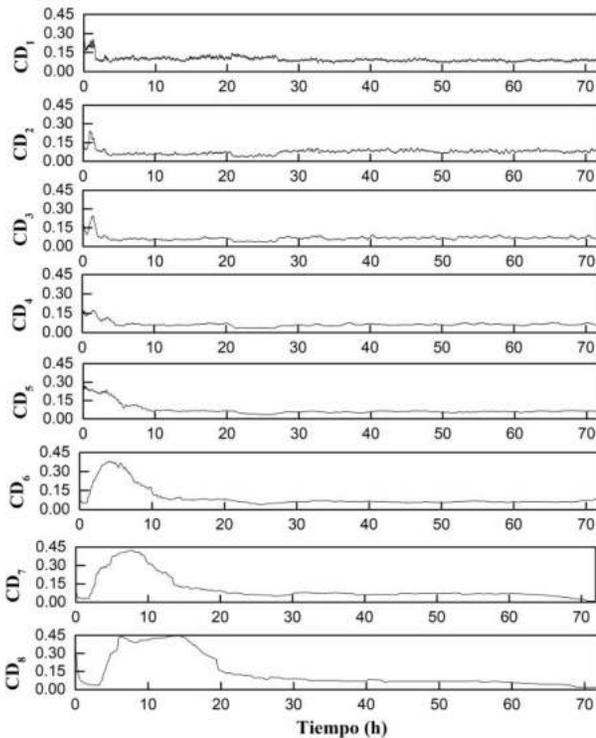


Fig. 3. Niveles de descomposición de la TWD de la serie de tiempo de pH de DA de vinazas tequileras.

Los CD de las bandas 5, 6 y 8 pueden ser correlacionados con la dinámica de las mediciones de DQO, AGV y producción de biogás, respectivamente. Estos resultados mostraron un seguimiento satisfactorio de los datos experimentales (Fig. 4). Debido a la capacidad de la TW como técnica multiescala, cada banda de descomposición puede ser asociada a un valor de escala temporal de modo que:

1. La DQO puede ser relacionada con una escala de análisis de 0.02 h la cual se encuentra en la banda 5.

2. La dinámica del AGV puede ser asociada a la banda 6 la cual está en una escala igual a 0.04 h.
3. La producción de biogás se puede asociar a la escala de 0.18 h encontrada en la banda 8.

Los valores de las escalas obtenidas en este trabajo se encuentran en el intervalo de las zonas de escalamiento reportadas por Mendez-Acosta et al., (2013).

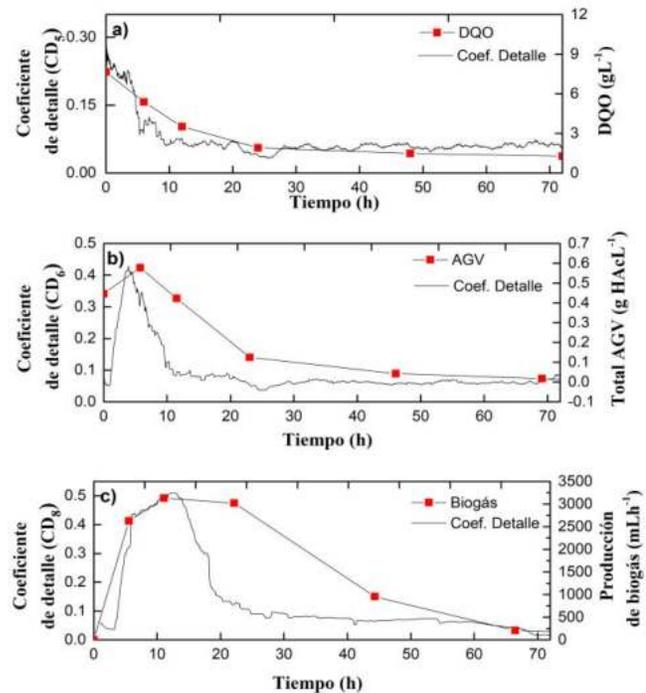


Fig. 4. Comparación de los datos experimentales de la DA de vinazas tequileras con los coeficientes de wavelet

3.2 Análisis con la transformada de wavelet para la DA de lactosuero

En el análisis de las series de tiempo de pH de la DA de lactosuero mediante la TWD también se encontraron tendencias de los CD a diferentes niveles de descomposición correlacionados con los datos experimentales de DQO y AGV. Por otro lado, la producción de biogás no fue correlacionada debido a que las condiciones de operación propuestas no favorecen la etapa metanogénica. En la Fig. 5 se puede observar que el comportamiento de los CD obtenidos con la banda 7 tiene una tendencia similar al de las mediciones fuera de línea de la DQO para dos diferentes temperaturas de operación ($T= 35$ y 40°C). Del mismo modo en la Fig. 6 se muestra la correlación existente de la banda 11 con la concentración total de AGV para ambas condiciones de operación, logrando nuevamente relacionar las bandas o niveles de descomposición a las etapas características presentes en el proceso de DA (i.e. hidrólisis y acidogénesis).

Esto sugiere nuevamente la existencia de escalas características asociadas a fenómenos intrínsecos del proceso. Las escalas características encontradas son de 0.036 h y 0.57 h para las bandas 7 y 11 respectivamente.

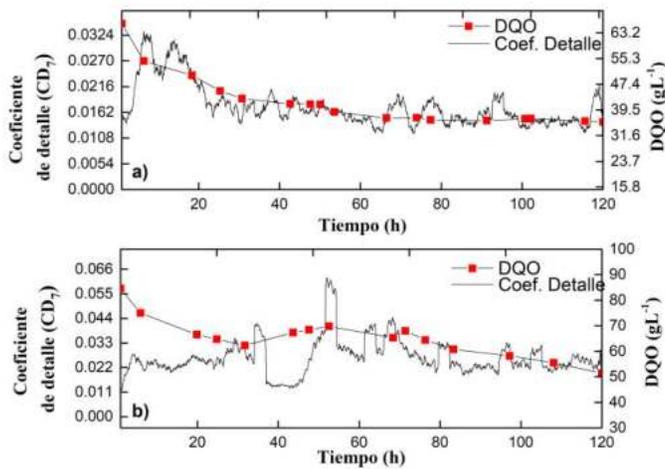


Fig. 5. Comparación de las mediciones de DQO con los coeficientes de wavelet, a) a $T=35^{\circ}\text{C}$ y b) a $T=40^{\circ}\text{C}$

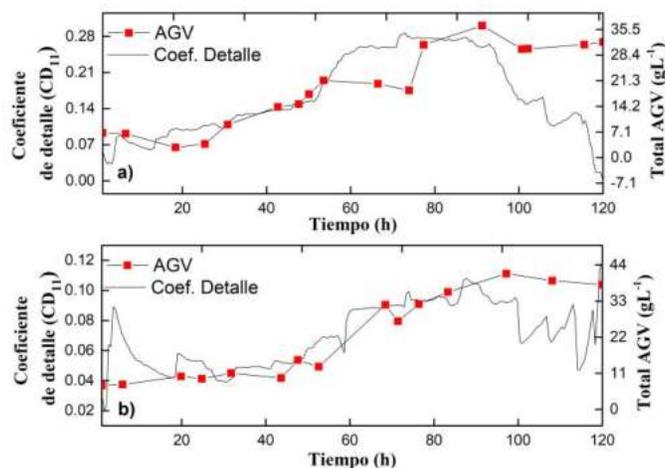


Fig. 6. Comparación de las mediciones de AGV con los coeficientes de wavelet, a) a $T=35^{\circ}\text{C}$ y b) a $T=40^{\circ}\text{C}$

6. CONCLUSIONES

En este trabajo se realizó un estudio de la utilidad de la metodología de la transformada de wavelet para el seguimiento indirecto de variables como DQO, AGV y producción de biogás en procesos de digestión anaerobia. El monitoreo se obtuvo a través del procesamiento de series de tiempo de variables fáciles de medir y que afectan directamente el rendimiento del proceso como lo es el caso del pH. Los resultados mostraron que la descomposición de la serie de tiempo por la transformada de wavelet es una buena alternativa para el seguimiento rápido del comportamiento de algunas variables clave. Si bien, esta

correspondencia solo es de manera cualitativa esta información es valiosa para conocer cómo evoluciona el proceso en el tiempo, aunado a esto una vez calibrada la metodología para un proceso en particular es posible seguir usando el mismo análisis para condiciones de operación diferentes.

En el caso de un seguimiento en línea de las variables clave, deberá considerarse un tiempo de espera necesario para el procesamiento de los datos que dependerá de la disponibilidad de la adquisición de los datos y de las escalas características del proceso, sin embargo, este trabajo no tiene el alcance suficiente y se tiene contemplado para trabajos futuros.

REFERENCIAS

- Bai, C., Xiao, Y., Peng, R., Grace, J. R., and Chen, Y. (2019). Analysis of the Hydrodynamic Effects of Gas Permeation in a Pilot-Scale Fluidized Bed Membrane Reactor. *Applied Sciences*, 9(1), 67.
- Bogaerts P, Wouwer AV. Software sensors for bioprocesses. (2003) Software for bioprocesses. *Isa Transactions*.42(2003), 547-558.
- Chu, C. Y., and Wang, Z. F. (2017). Dairy cow solid waste hydrolysis and hydrogen/methane productions by anaerobic digestion technology. *International Journal of Hydrogen Energy*, 42(52), 30591-30598
- Dochain D. (2003). State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of process control*.13(2003),801-818.
- García-Solano, M., Méndez-Acosta, H. O., Puebla, H., and Hernández-Martínez, E. (2016). Dynamic characterization of an anaerobic digester during the start-up phase by pH time-series analysis. *Chaos, Solitons & Fractals*, 82, 125-130.
- Lara-Musule, A., Álvarez-Sánchez, E., Trejo-Aguilar, G., Acosta-Domínguez, L., Puebla, H., and Hernández-Martínez, E. (2021). Diagnosis and Monitoring of Volatile Fatty Acids Production from Raw Cheese Whey by Multiscale Time-Series Analysis. *Applied Sciences*, 11(13), 5803.
- Mallat, S.G., (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (7), 674–693.
- Méndez-Acosta, H. O., Hernández-Martínez, E., Jáuregui-Jáuregui, J. A., Álvarez-Ramírez, J. and Puebla, H. (2013). Monitoring anaerobic sequential batch reactors via fractal analysis of pH time series. *Biotechnology and bioengineering*, 110(8), 2131-2139.
- Micolucci, F., Gottardo, M., Cavinato, C., Pavan, P. and Bolzonella, D. (2016). Mesophilic and thermophilic anaerobic digestion of the liquid fraction of pressed biowaste for high energy yields recovery. *Waste management*, 48, 227-235.
- Molina, R., Orcajo, G. and Martínez, F. (2018). KBR (Kinetics in Batch Reactors): a MATLAB-based

- application with a friendly Graphical User Interface for chemical kinetic model simulation and parameter estimation. *Education for Chemical Engineers*.
- Pradhan, P., Nayak, B. and Dhal, S. K. (2016). Time Series Data Prediction of Natural Gas Consumption Using Arima Model. *International Journal of Information Technology & Management Information System*, 7(3), 1-7
- Prado, P. F., Mayen, M. A. G., Silva, G. F. and Duarte, I. C. S. (2019). Using Matlab's wavelet toolbox to compare electric signals outputted by microbial fuel cells. *Sensing and Bio-Sensing Research*, 24(2019), 100285.
- Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 32(4), 1084-1093.
- Tahmasebpoor, M., Zarghami, R., Sotudeh-Gharebagh, R., and Mostoufi, N. (2015). Characterization of fluidized beds hydrodynamics by recurrence quantification analysis and wavelet transform. *International journal of multiphase flow*, 69, 31-41.
- Vaishnavi, P. and Muthumari, S. (2018). Application of Unscented Kalman Filter for States Estimation in Evaporator. *International Journal of Pure and Applied Mathematics*, 118(18), 2207-2214.