

# Prediction of Biomethane Production of Cheese Whey by using Artificial Neural Networks

Verónica E. Córdoba, Jorgelina Mussi, Mariano De Paula, Gerardo G. Acosta

**Abstract**— The search for new energy sources has intensified today worldwide. In Argentina, bioenergy continues to be the energy with minor participation in the energy matrix. Therefore, research should focus on searching for new substrates that increase their production without compromising agricultural systems. This work analyzed the cumulative methane production by anaerobic digestion of cheese whey, a food industry waste, using three substrate/inoculum ratios in volatile solids units. An artificial neural network (ANN) was developed to model the process based on the characterization data (input data) and methane production obtained in the laboratory (output data). The training algorithm used for the ANN was backpropagation. This model's validity was analyzed using statistical parameters such as the regression coefficient ( $R^2$ ) and the mean square error (MSE). The results obtained by the ANN were compared with those obtained by conventional kinetic models, such as the First-order and Gompertz models. The experimental biomethane production was in the range  $232 \pm 5$  to  $382 \pm 22$  NmL CH<sub>4</sub>/g VS. The proposed network could predict the experimental data with an  $R^2$  of 0.9992 and a training MSE of 2.7598. The statistics used to compare the goodness of fit between the models ( $R^2$  and RMSE) were higher for the network, demonstrating its ability to model a complex system such as anaerobic digestion.

**Index Terms**— Biomethane, Kinetics, Cheese whey, Modelling, Artificial neural network.

## I. INTRODUCCION

Con el aumento de la demanda energética, y la dependencia mundial de las fuentes de energía no renovables, las reservas de combustibles fósiles se ven disminuidas constantemente. Su quema aumenta la concentración de gases de efecto invernadero en la atmósfera, resultando en un aumento de las temperaturas globales [1]. Para superar estos problemas, numerosas investigaciones se han focalizado en la búsqueda de nuevos combustibles sostenibles [2]. En particular, la bioenergía es esencial para lograr una descarbonización de la economía [1].

En Argentina, la bioenergía se encuentra en pleno proceso de expansión, aunque sigue siendo la de menor participación en la matriz energética [2]. Dentro de los biocombustibles, el biogás es un combustible gaseoso, renovable que se obtiene por digestión anaeróbica (DA) a través del aprovechamiento de residuos agroindustriales, residenciales y sustratos diversos [3]; de esta manera, esta tecnología contribuye, a resolver problemas críticos de la sociedad actual como es la gestión de los residuos, mitigando los impactos negativos sobre los ecosistemas y la reducción de las emisiones globales de gases de efecto invernadero [4].

Una de las particularidades del proceso anaeróbico radica en las características altamente no lineales y complejas del sistema, proceso que depende de la actividad metabólica de diversos microorganismos para degradar sinérgicamente los materiales orgánicos [5], [6]. Los resultados del proceso dependerán de la composición de la biomasa y de las condiciones del proceso [7]. En este sentido, la relación sustrato/inóculo (RSI), ha sido identificado como uno de los factores que afecta no solo a la población microbiana, sino también a las propiedades fisicoquímicas de la fermentación [8], donde una dosis óptima reduce los tiempos del proceso, e incrementa su productividad [9]. La relación óptima no puede ser establecida de antemano, sino que depende fuertemente de factores como la composición del sustrato y la fuente de inóculo [8], [10]. La variabilidad de factores que influyen sobre el proceso, plantea la necesidad de utilizar modelos matemáticos que sean capaces de relacionar las condiciones del proceso como la temperatura, la concentración de inóculo, con el sustrato y la producción de biogás a fin de monitorear, analizar, optimizar y controlar la estabilidad de la DA [11].

Si bien existen diversos modelos que describen el proceso anaeróbico, tales como el modelo ADM1 [12], requiere del conocimiento detallado de parámetros y constantes cinéticas que en algunos casos son difíciles o imposibles de determinar [13]. Otros modelos más sencillos como el modelo de Primer Orden, el de Gompertz o el de Cone, basados en la identificación de la etapa limitante de la velocidad [11], han sido ampliamente utilizados para estudiar la cinética del proceso anaeróbico [14], [15], y establecer el potencial de diversos sustratos, aunque son incapaces de representar el efecto o la influencia de los parámetros fisicoquímicos de los sustratos o del proceso sobre el potencial de producción de biometano. En los últimos años, ha habido un mayor uso de la tecnología de Inteligencia Artificial, particularmente de las técnicas de "Machine Learning" o aprendizaje automático, como los modelos de Redes Neuronales Artificiales (RNA), que

V. E. Córdoba, INTELYMEC, Facultad de Ingeniería, Universidad Nacional del Centro de la Provincia de Buenos Aires UNCPBA. CIFICEN (UNCPBA, CICPBA, CONICET), vcordoba@fio.unicen.edu.ar

J. Mussi: Facultad de Ingeniería, Universidad Nacional del Centro de la Provincia de Buenos Aires UNCPBA. jmussi@fio.unicen.edu.ar

M. De Paula, INTELYMEC, Facultad de Ingeniería, Universidad Nacional del Centro de la Provincia de Buenos Aires UNCPBA. CIFICEN (UNCPBA, CICPBA, CONICET), marianodepaula@gmail.com

G. G. Acosta INTELYMEC, Facultad de Ingeniería, Universidad Nacional del Centro de la Provincia de Buenos Aires UNCPBA. CIFICEN (UNCPBA, CICPBA, CONICET), ggacosta@fio.unicen.edu.ar

permiten modelar y optimizar estos procesos, representando una alternativa a los modelos tradicionalmente se emplean en este dominio. El Perceptrón Multicapa (MLP), constituye una herramienta adecuada que puede ser utilizado para identificar la compleja relación no lineal entre los parámetros de entrada y salida [16]. En particular en los procesos biológicos las RNA pueden ser capaces de predecir su comportamiento sin conocimiento previo de las etapas metabólicas y cinéticas del sistema [17]. Esta herramienta ha sido ampliamente utilizada en sistemas de DA con diversos fines, tales como predecir la producción de biogás de sustratos en ensayos batch [18], o de agua residuales sintéticas [19] y de melazas [20] en reactores mesofílico semicontinuos. También ha sido utilizada para analizar la influencia de la utilización de co-sustratos y de la temperatura sobre la producción de metano [21], o incluso para para establecer la producción de otros gases como H<sub>2</sub>S y NH<sub>3</sub> en el biogás [22]. A pesar de los trabajos anteriormente indicados, según el saber de los autores, no se han encontrado artículos que analicen el efecto de la RSI, sobre la producción de metano en reactores batch utilizando RNA como modelo.

Por lo tanto, el objetivo de este trabajo es evaluar experimentalmente la producción de biometano de un residuo de la industria alimenticia, como es el suero de leche, analizando el efecto de la cantidad de inóculo sobre dicha producción, y modelar el proceso mediante la aplicación de RNA. En este estudio comparativo se presentan los datos obtenidos con las RNA, y se los compara con los obtenidos mediante la utilización de dos modelos de regresión no lineal convencionalmente utilizados para describir el proceso anaeróbico, a fin de evaluar el potencial predictivo de la RNA en este dominio particular.

## II. METODOLOGÍA

### A. Sustratos e Inóculo

Como sustrato se utilizó suero de leche ácido (SL), producido en la elaboración de queso ricota y obtenido de una quesería local ubicado en la ciudad de Olavarría, Buenos Aires, Argentina (36°54'00"S 60°20'00"O). Como inóculo se utilizó lodo de depuradora obtenido de la planta local de tratamiento de aguas residuales, los cuales fueron desgasificados (T=35 ± 1°C) previo a su utilización [7].

### B. Caracterización Físicoquímica

La caracterización de SL y del inóculo se realizó mediante la determinación de sólidos totales (ST, %), y volátiles en base seca (SV, %bs), demanda química de oxígeno (DQO, g/L), pH, nitrógeno amoniacal (NA, g/L) y alcalinidad total (AT, g CaCO<sub>3</sub>/L) a través de métodos APHA [23]. Alcalinidad parcial (AP) e intermedia (AI), asociada a los ácidos grasos volátiles y la capacidad buffer del sistema se determinó según [24].

### C. Diseño Experimental

La producción de biometano a partir de SL se determinó utilizando biorreactores de 1000 mL de capacidad, con 200 mL de SL e inóculo en cantidad suficiente para que la relación entre ambos, denominada relación Sustrato/Inóculo (RSI), en unidades de SV sea 1/1, 3/1 y 6/1. Los reactores se mantuvieron

en condiciones mesofílicas (T= 35±1°C) mediante baño termostático. Cada experiencia fue realizada por duplicado. El experimento se controló diariamente y se detuvo cuando la diferencia diaria de metano producido fue inferior al 0.2%. Las características del inóculo (contenido en SV de 2.6 ± 0.5%bs, pH 7.5, AT 1.7 ± 0.4 g CaCO<sub>3</sub>/L y NA 744±44 mg/L) se encontraron dentro del rango establecido por [7] para asegurar su calidad como tal, con excepción de la AT, que según [7] debe encontrarse por encima de los 3,0 g/L, razón por la cual para llevar a cabo las experiencias, se añadió bicarbonato de sodio a fin de neutralizar las acides del SL y aumentar la alcalinidad de las mezclas a los valores recomendados. El volumen de biogás (mL) se determinó por desplazamiento de volumen de agua y su calidad fue evaluada mediante el porcentaje de metano. Las mediciones se realizaron periódicamente utilizando el instrumento portátil Landgem GA2000 (Landtec, Reino Unido) equipado con celdas infrarrojas para medir metano (error máximo ± 0,5%). En paralelo, se realizó un blanco de inóculo cuya producción se restó de la producción de metano de las muestras de mezcla [7].

### D. Diseño de la RNA

La red neuronal se estructuró mediante la arquitectura de un perceptrón multicapa, entrenado por retropropagación, para ajustar y corregir los valores de los pesos (*w*) y sesgo o *biass* (*b*) de la red. A fin de establecer la mejor arquitectura que se adapte a los datos experimentales y sea capaz de estimar la producción acumulada de biometano, se evaluaron dos tipos de red, con una y dos capas escondidas (Fig. 1.). Para ambas RNA, la capa de entrada estará conformada por 10 neuronas, donde cada una refiere a la composición inicial de las mezclas de SL (sustrato) y lodos (inóculo) en términos de ST, SV, DQO, pH, AT, AP, AI y NH<sub>3</sub>, a la RSI y al tiempo de retención en el reactor (TRH), mientras que la capa de salida tendrá los datos correspondientes a la producción de biometano para cada condición inicial y cada uno de los tiempos analizados. Las variables de entrada seleccionada son variables que han sido establecidas previamente [25], [26] como las más relevantes y de influencia sobre el proceso anaeróbico. Estas variables se refieren a el contenido de humedad (ST), la cantidad de materia orgánica (SV, DQO), y la capacidad buffer del sistema (AT, AI, AP, NA), mientras que la RSI es otro parámetro que influye sobre la evolución del proceso y producción de metano del sistema. La capa de entrada de ambas redes consiste en una matriz de 10x174 datos, mientras que la capa de salida tendrá una dimensión de 1x174 datos.

El programa utilizado para la construcción y el análisis de las RNA fue MATLAB® (MathWork, versión R2017b). Este programa provee varios algoritmos de entrenamiento para redes neuronales. El algoritmo seleccionado fue "trainlm", correspondiente a una función de entrenamiento que actualiza los valores de peso y sesgo de acuerdo con la optimización de Levenberg-Marquardt de retropropagation (LMBP). La principal ventaja de este algoritmo es que proporciona una solución numérica al problema de minimizar una función no lineal, pero además tiene una convergencia estable y es adecuado para problemas de tamaño pequeño y mediano [27].

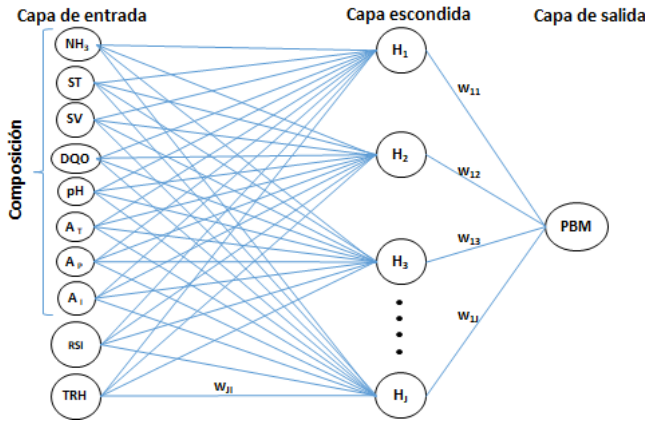


Fig. 1. Arquitectura de la RNA de 1 capa escondida.

Una vez seleccionado el algoritmo, para cada una de las RNA, se varió el número de neuronas de las capas escondidas y las funciones de activación de las capas, a fin de evaluar qué combinación presentaba el mejor ajuste de la red. Los datos totales (174) se dividieron, 70% para el Entrenamiento (1x122) y 30% para la Prueba, lo que incluye el Testeo (1x26) y la Validación (1x26) del modelo. Para determinar el rendimiento de la red, se utilizó el error cuadrático medio (MSE). Matlab provee los MSE de entrenamiento, testeo, validación por separado y del conjunto de datos como un todo, en este sentido, la iteración se detiene cuando el MSE ha alcanzado un mínimo. Otro parámetro a considerar es el coeficiente de correlación  $R^2$ , el cual permite establecer cuán adecuado es el ajuste de los datos experimentales a los datos obtenidos mediante el modelo. Un valor de  $R^2$  igual a 0 significa que hay una relación aleatoria. Cuanto menor sea el valor de la regresión, mayor será la diferencia entre la predicción de la serie temporal y la actual. La mejor arquitectura será aquella que presente el menor MSE, y el máximo  $R^2$ , tanto para el entrenamiento, como para la validación y el testeo.

### E. Modelos Cinéticos de Regresión no Lineal

La manera convencional de analizar el proceso anaeróbico es a través de la aplicación de modelos cinéticos, que proporcionan información útil para su análisis, diseño y operación [28]. Sin embargo, estos modelos sólo se pueden usar con datos experimentales disponibles en condiciones predeterminadas, pero no se pueden utilizar para predecir el rendimiento del proceso en condiciones operativas distintas a las examinadas [29]. Los modelos a utilizar en este trabajo son el modelo cinético de Primer Orden (PO) [30] y el modelo de Gompertz (EG) [31], ambos utilizados ampliamente por varios investigadores para modelar el proceso anaeróbico.

En el modelo de PO la producción de metano ( $B$ , mL/g SV) a lo largo del tiempo ( $t$ , días), se puede describir según (1):

$$B(t) = B_0(1 - e^{-kt}) \quad (1)$$

Donde  $B_0$  es el potencial de producción de biometano máximo (mL  $\text{CH}_4/\text{g SV}$ ) y  $k$  representa la tasa de producción

de metano ( $\text{d}^{-1}$ ). En aquellos casos en los que el proceso muestra un retraso en su producción de metano al inicio del proceso, la misma se puede describir mediante la Ec (2):

$$P(t) = P_0 * \exp\left(-\exp\left(\frac{R}{P_0}\right)(\lambda - t) \exp(1)\right) \quad (2)$$

Donde  $\lambda$  es el tiempo de la fase de retardo (d);  $P_0$  es la producción máxima de metano (mL  $\text{CH}_4/\text{g SV}$ ); y  $R$  es la tasa máxima de producción.

### F. Análisis Estadístico

Los datos experimentales se analizaron estadísticamente mediante análisis de varianza (ANOVA) utilizando Statgraphics Centurion 18 (18.1.12) a fin de establecer diferencias significativas en los parámetros al 95% de confiabilidad. La medida de la bondad del ajuste de los datos experimentales a los modelos estudiados se realizó en base al coeficiente de correlación  $R^2$  y el RMSE, siendo, el mejor modelo, aquel que provee el mayor  $R^2$  y el menor RMSE. En el caso en particular del modelo de RNA, los datos de salida de la red fueron extraídos de Matlab y comparados con los datos experimentales para cada RSI. De acuerdo a [18] RMSE es un indicador de la precisión de los modelos para predecir la respuesta, y se considera como un criterio importante para el ajuste en caso de que el propósito del modelo sea la predicción, mientras que  $R^2$  estima la robustez del modelo [17].

## III. RESULTADOS Y DISCUSIÓN

### A. Caracterización Físicoquímica

La caracterización físicoquímica del SL y las mezclas utilizado en la prueba de DA en función de la RSI se detalla en la Tabla I. Como se puede observar el SL exhibe una elevada concentración de materia orgánica en términos del contenido en SV (5.3%), y DQO (72.290 mg/L), y un valor de pH ácido (4.2). Se observan diferencias significativas entre tratamientos en los valores de DQO y SV ( $p \geq 0.05$ ), aunque en todos los casos, el valor de pH se mantuvo en el rango adecuado sugerido para el crecimiento de bacterias metanogénicas [32].

### B. Producción de Biogás y Metano

El desempeño del proceso anaeróbico se analizó en términos de la remoción de materia orgánica en base al contenido de SV, y a través de las producciones acumuladas de biogás y metano (Fig. 2.). Con el incremento de la RSI de 1 a 6, la producción de biogás y metano disminuyó 59% (522 a 213 mL/g SV) y 81% (455 a 88 mL/g SV), respectivamente. [33] quienes trabajaron con SL en reactores discontinuos reportaron una producción de metano de 424 mL/g DQO (equivalente a 23.4 L  $\text{CH}_4/\text{L SL}$ ). Este valor resultó un 5% inferior al valor obtenido en el tratamiento A (24.5 L  $\text{CH}_4/\text{L SL}$ ), aunque esta diferencia aumentó hasta un 408% para el tratamiento C (88 mL/g SV, equivalente a 4,6 L  $\text{CH}_4/\text{L SL}$ ). La remoción de materia orgánica varió entre 36% y 43%, aunque no se observaron diferencias significativas entre los tratamientos. La menor remoción de SV se observó en el tratamiento C (36%), resultado asociado a una menor concentración de inóculo en el reactor.

La reducción en la productividad del sistema a medida que aumenta la RSI se puede deber a las características del SL, el cual contiene el 55% de los nutrientes de la leche, incluyendo proteínas solubles, lactosa, vitaminas y minerales, además de una alta acidez con valores de pH inferiores a 5. Según [34], la lactosa, se convierte fácilmente en ácidos orgánicos por acción de los microorganismos, pudiendo desestabilizar el proceso. Sin embargo, un incremento de la concentración de microorganismos metanogénicos resultado de una mayor concentración de inóculo, contribuiría a una eficiente conversión de los ácidos (intermediarios en el proceso de DA) en metano, evitando la sobrecarga del proceso [8]. Por lo tanto, el uso de un inóculo adecuado en la proporción correcta

TABLA I:  
COMPOSICIÓN DEL SL, Y LAS MEZCLAS

Parámetro	SL	Tratamiento (RSI)		
		A (1/1)	B (3/1)	C (6/1)
pH	4.3	7.6	7.6	7.5
AT		5.4±0.0a	8.8±0.1b	10.9±0.2c
AI	0.0±0.0	2.2±0.1a	4.0±0.1b	5.0±0.3c
AP		3.2±0.1a	4.8±0.0b	6.0±0.1c
NA	0.13±0.01	0.17±0.01a	0.15±0.0a	0.14±0.01b
ST	5.8±0.0	5.6±0.1a	5.7±0.1a	5.7±0.1a
SV	5.3±0.0	3.5±0.1a	4.0±0.1b	4.4±0.1c
DQO	72.3±0.6	55.8±0.1a	60.6±0.2b	62.6±1.5b

Los valores que se muestran en la tabla son valores medios de las repeticiones ± desviación estándar. Valores con la misma letra en la misma fila no presentan diferencias significativas ( $p > 0.05$ ) de acuerdo con el procedimiento de Fisher de la mínima diferencia significativa (LSD test). Unidades, AT, AP y AI (g/L); NA (g/L); ST y SV (%); DQO (g/L)

permitirá el máximo aprovechamiento de la materia orgánica disponible, tal como se observa en el tratamiento A.

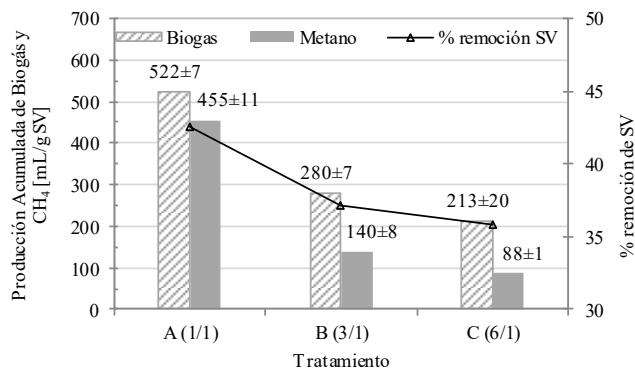


Fig. 2. Producción de biogás, metano y remoción de SV.

### C. Resultados y Selección de la RNA

A fin de establecer la RNA que mejor ajustara a los datos, se analizaron dos arquitecturas, con una y dos capas escondidas (Fig. 1.). Se modificaron en cada una de ellas: a) la función de activación de las capas escondidas, utilizando intercaladamente Logsig (Log-sigmoid) y tansig (tangent sigmoid), y b) el

número de neuronas hasta alcanzar un máximo de 15, similar a lo analizado por [29], [35] para sistemas anaeróbicos. Para la función de activación en la capa de salida se utilizó la función purelin (linear), resultados que se muestran en la Tabla II y III.

La determinación de la RNA más adecuada se basó en obtener un mínimo MSE, y un máximo  $R^2$ . Posteriormente, se corrobora mediante el análisis del histograma de errores, que los mismos se distribuyan alrededor de cero.

Se puede observar que la mejor arquitectura corresponde al sistema de dos capas escondidas (10:15:10:1), un algoritmo de *back-propagation* y funciones de activación tansig /logsig /purelin, cuya arquitectura presenta los menores MSE (2.7598, 5.0656 y 6.1657 para el entrenamiento, validación y testeo respectivamente), y un coeficiente de determinación para el conjunto de datos ( $R^2 = 0.9999$ ) superiores a cualquiera de las arquitecturas de 1 capa escondida. [29], [35] quienes analizaron la producción de biogás de residuos agrícolas y lignocelulósicos, trabajaron con arquitecturas similares en términos de funciones de activación y número de neuronas, aunque la principal diferencia es que estos autores utilizan una red de una sola capa escondida.

En la Fig. 3. Se muestra el histograma de errores, obtenidos como diferencia entre los valores objetivo y los valores predichos por la red. Puede observarse que los de errores se distribuyen cercanos a cero, exhibiendo una distribución cercana a la distribución normal.

El siguiente paso para validar la RNA es analizar los gráficos de regresión (Fig. 4.), que muestra la relación entre las salidas de la red y los objetivos. En un ajuste perfecto, los datos deberían caer a lo largo de una línea de 45°, donde las salidas de la red son iguales a los objetivos, pero la relación rara vez es perfecta en la práctica. Para este trabajo, el ajuste es razonablemente bueno para todo el conjunto de datos, con  $R^2$  superior a 0.99987. Este diagrama también es útil para mostrar si ciertos puntos de datos tienen ajustes deficientes, sin embargo, tal como se observa, todos los puntos caen lo suficientemente cerca de la línea. Por lo tanto, del análisis realizado se puede indicar que el modelo de RNA propuesto presentó un rendimiento de estimación alto y satisfactorio.

### D. Comparación de la Eficiencia de la Predicción de la Red con los Modelos de Ajuste Convencionales.

Los datos experimentales de producción de metano a lo largo del tiempo fueron ajustados a los modelos de regresión no lineales de Primer Orden y Gompertz. Se puede observar de la Fig. 5. que la ecuación cinética de PO no es capaz de explicar el comportamiento del sistema de digestión anaeróbica y producción de biometano en forma completa, mostrando valores de  $R^2$  entre 76% y 90%. En particular, durante los primeros días del proceso se suele producir un retardo o *delay* en la producción de biometano, difícil de explicar por una ecuación de tipo exponencial. Estos problemas generalmente se resuelven mediante el ajuste de los datos a la EG, la cual modela con mayor detalle el comportamiento de los sistemas que contemplan dicha fase ( $R^2 > 97.4\%$ ). En la Fig. 5., se muestra el ajuste obtenido por la RNA a los datos experimentales.

TABLA II  
RESULTADOS OBTENIDOS DURANTE EL ENTRENAMIENTO Y TESTEO DE LA RNA DE UNA CAPA ESCONDIDA

Sistema	N° de neuronas capa escondidas	Función de activación		MSE			Mejor Época	R <sup>2</sup>			
		1°	2°	Entrenamiento	Validación	Testeo		Entrenamiento	Validación	Testeo	Todo
1	5	logsig	purelin	21.1622	36.0273	33.9829	22	0.99944	0.9994	0.99923	0.99941
2	5	tansig	purelin	6.1366	18.4986	25.9704	30	0.99987	0.99963	0.9988	0.99974
3	6	tansig	purelin	13.8154	21.5459	59.6396	15	0.99966	0.99955	0.99911	0.9995
4	6	logsig	purelin	2.8271	3.1311	4.522	71	0.99993	0.99994	0.99992	0.99993
5	8	tansig	purelin	12.8726	9.5479	20.8174	75	0.99972	0.99974	0.99931	0.99968
6	8	logsig	purelin	5.7913	3.6919	7.3408	102	0.99988	0.99982	0.99977	0.99987
7	10	tansig	purelin	5.8753	5.7712	10.225	58	0.99987	0.99979	0.99978	0.99985
8	10	logsig	purelin	7.83	14.1814	30.5562	54	0.99982	0.99967	0.99923	0.99971
9	15	tansig	purelin	5.6594	7.1837	13.5206	69	0.99987	0.99985	0.99969	0.99983
10	15	logsig	purelin	6.4049	5.085	10.8988	64	0.99985	0.99985	0.9998	0.99984

TABLA III:  
RESULTADOS OBTENIDOS DURANTE EL ENTRENAMIENTO Y TESTEO DE LA RNA DE DOS CAPAS ESCONDIDAS.

Sistema	Número de Neuronas		Función de Activación			MSE			Época	R <sup>2</sup>			
	1°	2	1°	2°	3°	Entrenamiento	Validación	Testeo		Entrenamiento	Validación	Testeo	Todo
1	15	10	logsig	tansig	Purelin	3.9601	5.1782	6.8162	106	0.9991	0.9999	0.9997	0.9999
2	15	10	tansig	logsig	Purelin	2.7598	5.0656	6.1657	20	0.9999	0.9998	0.9998	0.9999

MSE: Error cuadrático medio. R<sup>2</sup>: coeficiente de determinación.

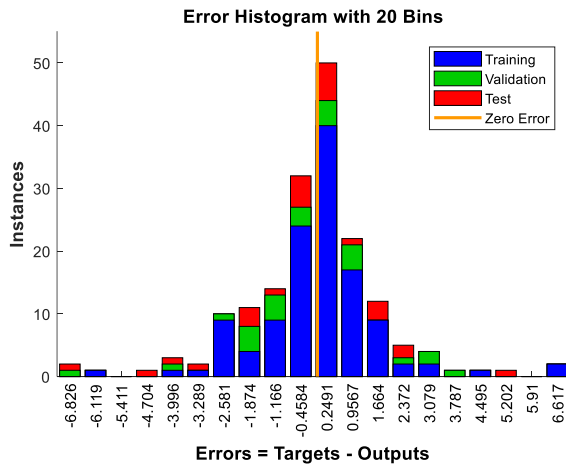


Fig. 3. Histograma de errores. Fuente: MATLAB® software (MathWork, versión R2017b).

Se observa que el ajuste ( $R^2 > 98.55\%$ ) es superior al obtenido por la EG ( $R^2 > 97.4\%$ ), aunque ambos modelos muestran un desempeño adecuado. Sin embargo, los resultados de las RNA son más precisos, lo que se puede corroborar al analizar otro parámetro estadístico como el RMSE, ya que un menor valor implica un modelo capaz de predecir con mayor exactitud los datos. Tomando ambos parámetros en consideración, el modelo de RNA exhibe un mejor desempeño que el modelo de regresión no lineal, con un rendimiento de estimación alto y satisfactorio, lo que asegura un adecuado poder predictivo basado en la capacidad de establecer relaciones entre múltiples entradas y salidas en un entorno altamente complejo, como lo es el proceso anaeróbico.

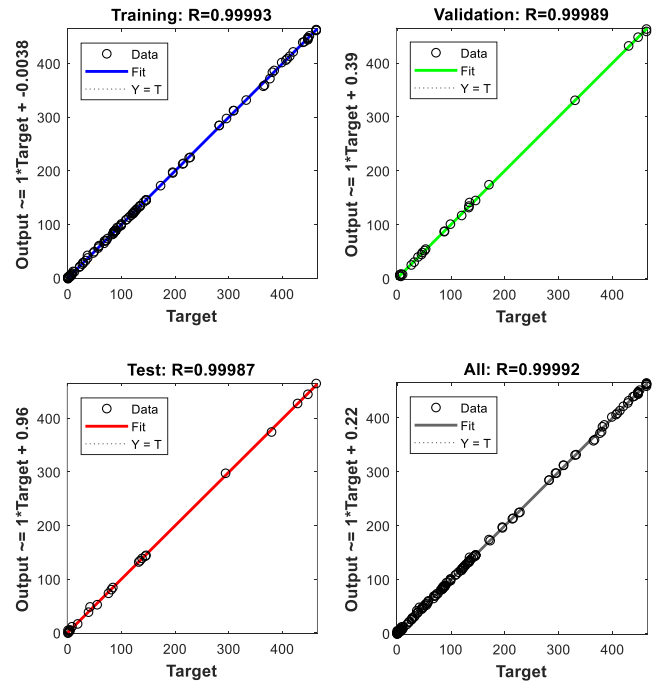


Fig. 4. Diagramas de regresión para el entrenamiento, validación, testeo y datos globales. Fuente: MATLAB® software (MathWork, versión R2017b).

De la comparación de los modelos de regresión no lineal y de la RNA, puede observarse que cada uno es capaz de brindar información sobre parámetros particulares del proceso. El modelo de primer orden proporciona la constante de velocidad de la etapa hidrolítica ( $k$ ), mientras que el modelo de Gompertz

permite determinar el tiempo de latencia  $\lambda$ , en caso de existir y la máxima velocidad de producción de metano, valores que dependen fuertemente de la RSI ( $p < 0.05$ ); sin embargo, estos modelos fallan al intentar relacionar las características fisicoquímicas del sustrato, con su producción de biometano, por lo que resulta imposible extrapolar la información brindada. Por su parte, con el modelo de RNA que relaciona variables de entrada (características del sustrato y RSI) con variables de salida del proceso (PBM), es posible extrapolar el modelo a otros sistemas, además de brindar información que permita su control.

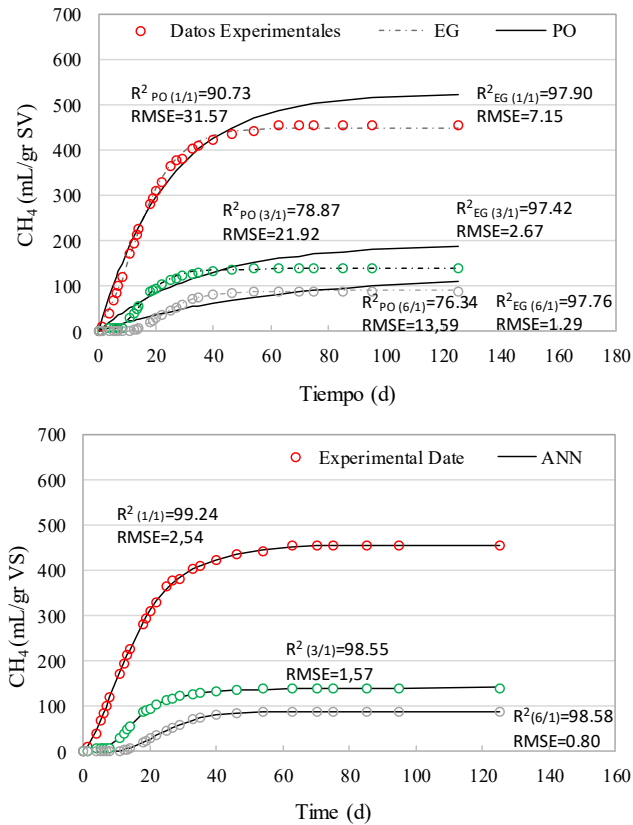


Fig. 5. Superior: Ajuste de datos experimentales al modelo de Primer Orden (PO —) y a la Ecuación de Gompertz (EG ---). Inferior: Ajuste de datos a la RNA (—). En ambos gráficos los símbolos (O, O, O) representan los datos experimentales para las RSI (1/1; 3/1 y 6/1).

#### IV. CONCLUSION

El presente trabajo tuvo como objetivo modelar la producción de biometano de suero de leche bajo tres relaciones sustrato/inóculo diferentes. Los resultados experimentales mostraron una fuerte influencia de la concentración de inóculo sobre la evolución del proceso. El modelado se realizó mediante la utilización de RNA, siendo el algoritmo de entrenamiento de retropropagación de dos capas escondidas (10:15:10:1) y funciones de activación tansig/logsig/purelin el que presentó un mejor ajuste a los valores experimentales. Los resultados obtenidos por la RNA mostraron una elevada correlación, mejor aún que la obtenida a partir de los modelos convencionales analizados (Modelo de Primer Orden y la Ecuación de

Gompertz). La principal ventaja en la utilización de un modelo basado en un Perceptrón Multicapa (MLP) radica en el hecho de que estos modelos son capaces de predecir la producción de metano bajo diferentes condiciones de alimentación del reactor y características del sustrato, sin requerir de conocimientos previos de la cinética del sistema ni de las ecuaciones matemáticas que la gobiernan en comparación con otros modelos más complejos como el ADM1. Estos resultados muestran que los modelos de redes neuronales pueden ser utilizadas no solo para analizar la evolución de la producción de metano, sino que también debido a la versatilidad de la misma, permitiría analizar la evolución de otros parámetros de gran importancia como la alcalinidad o el pH, parámetros que brindan información sobre la estabilidad del proceso. Además, al ser un modelo orientados a datos, representa una ventaja importante para el modelado (y control) de este tipo de procesos biológicos que están ganando importancia en la generación de energía de forma sustentable. Este trabajo demuestra que los resultados de las estimaciones son de mejor calidad, respecto de RMSE y  $R^2$ , que aquellas que proporcionan los modelos de primer orden o Gompertz.

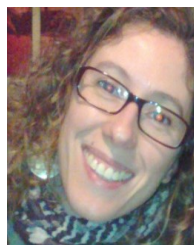
#### AGRADECIMIENTOS

Los autores agradecen a la Cooperativa de Electricidad local de Olavarría por el suministro de los lodos de depuradora utilizados en este estudio y a la empresa La Casiana S.A. por el suministro de suero de queso.

#### REFERENCIAS

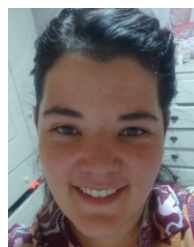
- [1] M. T. Majeed, T. Luni, and T. Tahir, 'Growing green through biomass energy consumption: the role of natural resource and globalization in a world economy', *Environ. Sci. Pollut. Res.*, vol. 29, no. 22, pp. 33657–33673, 2022, doi: 10.1007/s11356-021-18017-w.
- [2] A. G. Nogar, L. D. Rodríguez, C.-V. Bongiorno, and E. M. Santalla, 'Las potencialidades bioenergéticas del Arundo donax L. en Argentina', *Let. Verdes. Rev. Latinoam. Estud. Socioambientales*, no. 30, pp. 84–104, 2021, doi: 10.17141/letrasverdes.30.2021.4607.
- [3] J. Liebetrau, U. Baier, D. Wall, and J. D. Murphy, *Integration of biogas systems into the energy system*, no. 8. IEA Bioenergy Task 37, 2020.
- [4] A. Scott and R. Blanchard, 'The role of anaerobic digestion in reducing dairy farm greenhouse gas emissions', *Sustain.*, vol. 13, no. 5, pp. 1–18, 2021, doi: 10.3390/su13052612.
- [5] I. Andrade Cruz *et al.*, 'Application of machine learning in anaerobic digestion: Perspectives and challenges', *Bioresour. Technol.*, vol. 345, no. 3, pp. 884–902, 2021, p. 126433, 2022, doi: 10.1016/j.biortech.2021.126433.
- [6] R. B. Recio Colmenares and K. J. Gurubel Tún, 'Modelado neuronal de un proceso de digestión aeróbica de aguas residuales', *Bistua Rev. La Fac. Ciencias Basicas*, vol. 16, no. 1, p. 163, 2019, doi: 10.24054/01204211.v1.n1.2018.3204.
- [7] C. Holliger *et al.*, 'Towards a standardization of biomethane potential tests', *Water Sci. Technol.*, vol. 74, no. 11, pp. 1–9, 2016, doi: 10.2166/wst.2016.336.
- [8] X. Ma, T. Jiang, J. Chang, Q. Tang, T. Luo, and Z. Cui, 'Effect of Substrate to Inoculum Ratio on Biogas Production and Microbial Community During Hemi-Solid-State Batch Anaerobic Co-digestion of Rape Straw and Dairy Manure', *Appl. Biochem. Biotechnol.*, vol. 189, no. 3, pp. 884–902, 2019, doi: 10.1007/s12010-019-03035-9.
- [9] J. Kassongo, E. Shahsavari, and A. S. Ball, 'Substrate-to-inoculum ratio drives solid-state anaerobic digestion of unamended grape marc and cheese whey', *PLoS One*, vol. 17, no. 1 January, pp. 1–19, 2022, doi: 10.1371/journal.pone.0262940.
- [10] S. A. Alavi-Borazjani, I. Capela, and L. A. C. Tarelho, 'Over-acidification control strategies for enhanced biogas production from anaerobic digestion: A review', *Biomass and Bioenergy*, vol. 143, no.

- June, p. 105833, 2020, doi: 10.1016/j.biombioe.2020.105833.
- [11] S. Emebu, J. Pecha, and D. Janáčová, 'Review on anaerobic digestion models: Model classification & elaboration of process phenomena', *Renew. Sustain. Energy Rev.*, vol. 160, no. 112288, pp. 1–17, 2022, doi: 10.1016/j.rser.2022.112288.
- [12] D. J. Batstone, J. Keller, I. Angelidaki, S. V. Kalyuzhnyi, S. G. Pavlostathis, and A. Rozzi, 'Anaerobic Digestion Model No 1 (ADM1)', *Water Sci. Technol.*, vol. 45, no. 10, pp. 65–73, 2002.
- [13] Sh.-S. Jeong, C.-W. Suh, J.-L. Lim, S.-H. Lee, and H.-S. Shin, 'Analysis and application of ADM1 for anaerobic methane production', *Bioprocess Biosyst. Engineering*, vol. 27, pp. 81–89, 2005, doi: 10.1007/s00449-004-0370-4.
- [14] D. G. Ibarlucía, E. M. Santalla, and V. E. Córdoba, 'Evaluation of biomethane potential and kinetics modelling of green macroalgae from the South Atlantic Sea: *Codium* sp. (Codiaceae) and *Ulva* sp. (Ulvaceae)', *Environ. Chem.*, vol. 18, no. 7, pp. 311–320, 2021, doi: 10.1071/EN21088.
- [15] D. D. Nguyen, 'Thermophilic anaerobic digestion of model organic wastes: Evaluation of biomethane production and multiple kinetic models analysis', *Bioresour. Technol.*, vol. 280, no. December 2018, pp. 269–276, 2019, doi: 10.1016/j.biortech.2019.02.033.
- [16] W. Y. Chen, 'Artificial Neural Network (ANN) Modelling for Biogas Production in Pre-Commercialized Integrated Anaerobic-Aerobic Bioreactors (IAAB)', *Water (Switzerland)*, vol. 14, no. 9, pp. 0–36, 2022, doi: 10.3390/w14091410.
- [17] T. Beltramo, C. Ranzan, J. Hinrichs, and B. Hitzmann, 'Artificial neural network prediction of the biogas flow rate optimised with an ant colony algorithm', *Biosyst. Eng.*, vol. 143, pp. 68–78, 2016, doi: 10.1016/j.biosystemseng.2016.01.006.
- [18] N. E. Mougari, J. F. Largeau, N. Himrane, M. Hachemi, and M. Tazerout, 'Application of artificial neural network and kinetic modeling for the prediction of biogas and methane production in anaerobic digestion of several organic wastes', *Int. J. Green Energy*, vol. 18, no. 15, pp. 1584–1596, 2021, doi: 10.1080/15435075.2021.1914630.
- [19] F. Tufaner and Y. Demirci, 'Prediction of biogas production rate from anaerobic hybrid reactor by artificial neural network and nonlinear regressions models', *Clean Technol. Environ. Policy*, no. 0123456789, 2020, doi: 10.1007/s10098-020-01816-z.
- [20] K. Yetilmezsoy, F. I. Turkdogan, I. Temizel, and A. Gunay, 'Development of ann-based models to predict biogas and methane productions in anaerobic treatment of molasses wastewater', *Int. J. Green Energy*, vol. 10, no. 9, pp. 885–907, 2013, doi: 10.1080/15435075.2012.727116.
- [21] R. Alrowais, A. Al-otaibi, A. Y. Hatata, M. A. Essa, and M. M. Abdel, 'Comparing the effect of mesophilic and thermophilic anaerobic co-digestion for sustainable biogas production: An experimental and recurrent neural network model study', *J. Clean. Prod.*, vol. 392, no. December 2022, p. 136248, 2023, doi: 10.1016/j.jclepro.2023.136248.
- [22] D. Strik, A. Domnanovich, L. Zani, R. Braun, and P. Holubar, 'Prediction of trace compounds in biogas from anaerobic digestion using the MATLAB Neural Network Toolbox', *Environ. Model. Softw.*, vol. 20, no. 6, pp. 803–810, Jun. 2005, doi: 10.1016/j.envsoft.2004.09.006.
- [23] APHA, *APHA: Standard Methods for the Examination of Water and Wastewater. 20th ed. Washington, DC.*, no. 1. American Public Health Association, 1999.
- [24] S. Jenkins, J. Morgan, and C. Sawyer, 'Measuring anaerobic sludge digestion and growth by a simple alkalimetric titration', *Water Pollut. Control Fed.*, vol. 55, no. 5, pp. 448–453, 1983, [Online]. Available: <http://www.jstor.org/stable/25041903>.
- [25] C. Holliger, 'Towards a standardization of biomethane potential tests', *Water Sci. Technol.*, pp. 1–9, 2016, doi: 10.2166/wst.2016.336.
- [26] I. Angelidaki, 'Defining the biomethane potential (BMP) of solid organic wastes and energy crops: a proposed protocol for batch assays', *Water Sci. Technol.*, vol. 59, no. 5, pp. 927–34, Jan. 2009, doi: 10.2166/wst.2009.040.
- [27] C. Lv, 'Levenberg-marquardt backpropagation training of multilayer neural networks for state estimation of a safety-critical cyber-physical system', *IEEE Trans. Ind. Informatics*, vol. 14, no. 8, pp. 3436–3446, 2018, doi: 10.1109/TII.2017.2777460.
- [28] I. Angelidaki, L. Ellegaard, and B. K. Ahring, 'A Mathematical Model for Dynamic Simulation of Anaerobic Digestion of Complex Substrates: Focusing on Ammonia Inhibition', *Biotechnol. Bioeng.*, vol. 42, pp. 159–166, 1993, doi: 10.1002/bit.260420203.
- [29] F. Almomani, 'Prediction of biogas production from chemically treated co-digested agricultural waste using artificial neural network', *Fuel*, vol. 280, no. July, p. 118573, 2020, doi: 10.1016/j.fuel.2020.118573.
- [30] A. Hublin and B. Zelic, 'Modelling of the whey and cow manure co-digestion process.', *Waste Manag. Res. J. Int. Solid Wastes Public Clean. Assoc. ISWA*, vol. 31, no. 4, pp. 353–60, Apr. 2013, doi: 10.1177/0734242X12455088.
- [31] F. M. Pellerá and E. Gidaracos, 'Effect of substrate to inoculum ratio and inoculum type on the biochemical methane potential of solid agroindustrial waste', *J. Environ. Chem. Eng.*, vol. 4, pp. 3217–3229, 2016, doi: 10.1016/j.jece.2016.05.026.
- [32] M. H. Gerardi, *The microbiology of anaerobic digesters*, 1st ed. John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.
- [33] T. H. Ergüder, U. Tezel, E. Güven, and G. N. Demirel, 'Anaerobic biotransformation and methane generation potential of cheese whey in batch and UASB reactors.', *Waste Manag.*, vol. 21, no. 7, pp. 643–50, Jan. 2001, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11530920>.
- [34] B. Lagoa-Costa, C. Kennes, and M. C. Veiga, 'Cheese whey fermentation into volatile fatty acids in an anaerobic sequencing batch reactor', *Bioresour. Technol.*, vol. 308, no. March, p. 123226, 2020, doi: 10.1016/j.biortech.2020.123226.
- [35] M. Das Ghatak, A. Ghatak, M. Das, and A. Ghatak, 'Artificial neural network model to predict behavior of biogas production curve from mixed lignocellulosic co-substrates', *Fuel*, vol. 232, no. May 2017, pp. 178–189, 2018, doi: 10.1016/j.fuel.2018.05.051.



**Verónica E. Córdoba.** Is a Chemical Engineering, (2006) and Ph.D in Science and Technology, Chemistry Mention (University of San Martín, Bs As, Argentina -2016). Since 2011 she has been working in renewable energies in the INTELYMEC nucleus, developing knowledge about the anaerobic process.

She is currently an Assistant Researcher at CONICET and an Adjunct Professor at the Eng. Faculty at the *Universidad Nacional del Centro de la Provincia de Buenos Aires* (UNCPBA), Argentina.



**Jorgelina Mussi** graduated as Chemical Engineering, Olavarría, Argentina (2019), She is currently a Ph.D student and assistant at the Eng. Faculty at the *Universidad Nacional del Centro de la Provincia de Buenos Aires* (UNCPBA), Argentina.



**Mariano De Paula** Industrial Engineer (Engineering Faculty - UNCPBA, Argentina - 2007), Ph.D. in Engineering, at National Technological University– (2013). He is also a researcher of the Argentinean National Research Council (CONICET) since 2015, working in Engineering Group INTELYMEC, UNCPBA. The topics of interest are the

study of bioinspired techniques, artificial intelligence and Reinforcement Learning and its variants for control and decision-making in complex systems. In addition, he works as a professor at the Faculty of Engineering of the UNCPBA.



**Gerardo G. Acosta** graduated as Engineer in Electronics at the National University of La Plata, Argentina (1988), and as Ph.D. in Computer Science, at the University of Valladolid, Spain (1995). He is currently a Full Professor in Control Systems at the Eng. Faculty at the *Universidad Nacional del Centro de la Provincia de Buenos Aires* (UNCPBA), Argentina. He is also Principal researcher of the Argentinean National Research Council (CONICET) and Director of the INTELYMEC nucleus, part of the *Centro de Investigaciones en Física e Ingeniería del Centro – CIFICEN*. His working interests comprise the use of intelligent control techniques in many domains, including process control and terrestrial and marine robotics. He is Senior Member of the IEEE, Chairman of the IEEE Computational Intelligence Society Argentinean Chapter (2007-2008), and current Chairman of the IEEE Oceanic Engineering Society Argentinean Chapter, being one of its founders, and member of the Administrative Committee of IEEE OES (2015-2016; 2018-2023).