

Analogía entre el modelo semifísico de base fenomenológica y las redes neuronales utilizando como objeto de estudio el crecimiento de biomasa de *Chlorella sorokiniana*.

Analogy between phenomenology-based semi physical modeling and neural networks using *Chlorella sorokiniana* biomass growth as the object of study.

Nicole D. Longa^a, Manuela Londoño Cruz^a, Edison A. Acosta^{b,*}.

a Estudiante de Ingeniería Bioquímica, Facultad de ingeniería - Dpto. Ingeniería Bioquímica, Universidad Icesi, Calle 18 No. 122-135 Pance, Cali - Colombia
b Mg. Ing. Profesor del Departamento de Ingeniería Bioquímica, Universidad Icesi, Calle 18 No. 122-135 Pance, Cali – Colombia.

* edison.acostal@u.icesi.edu.co

Resumen

El interés hacia el uso de modelos matemáticos como, los semifísicos de base fenomenológica (MSBF) y los de redes neuronales (RNA), con el propósito de predecir y analizar el comportamiento de microorganismos como *Chlorella sorokiniana*, cada vez aumenta. Esto se debe a que escalar, diseñar y controlar un cultivo de este microorganismo a gran escala representa grandes gastos en tiempo y dinero. Sin embargo, el modelado de estos sistemas en la actualidad utilizando los modelos matemáticos en cuestión es un proceso en mejora, debido a que aún se desconoce qué tan precisos son estimando el crecimiento de biomasa de *Chlorella sorokiniana*. Por tal motivo, en este estudio se comparó un modelo de base fenomenológica y uno de aprendizaje automatizado en la estimación de biomasa de *Chlorella sorokiniana* teniendo en cuenta el efecto de la irradiancia y una fuente de carbono orgánica, y tomando como punto de referencia los datos experimentales obtenidos en una investigación de la Universidad Icesi. Como resultado, el MSBF obtuvo un porcentaje de error promedio del 11,82%, mientras que el modelo de RNA logró estimar de manera más precisa el comportamiento de la concentración de biomasa microalgal con un porcentaje de error promedio del 6,01%.

Palabras Clave: Microbiología predictiva, Modelamiento, Aprendizaje automático, *Chlorella sorokiniana*.

Abstract

The interest in the use of mathematical models, such as semi-physical phenomenological based models (SPBM) and neural network models (ANN), for the purpose of predicting and analyzing the behavior of microorganisms such as *Chlorella sorokiniana*, is increasing. This is because scaling up, designing, and controlling a large-scale culture of this microorganism represents great expenses in time and money. However, the modeling of these systems at present using the mathematical models in question is an improving process because it is still unknown how accurate they are in estimating the biomass growth of *Chlorella sorokiniana*. For this reason, this study compared a phenomenological-based model and an automated learning model in the estimation of biomass of *Chlorella sorokiniana* considering the effect of irradiance and an organic carbon source and taking as a reference point the experimental data obtained in research of the Universidad Icesi. As a result, the SPBM obtained an average error percentage of 11.82%, while the ANN model was able to estimate more accurately the behavior of the microalgal biomass concentration, with average error percentage was 6.01%.

Keywords: Predictive microbiology, Modeling, Machine learning, *Chlorella sorokiniana*.

1. Introducción

Construir y simular un modelo matemático es un proceso del que se dispone actualmente para analizar y predecir el comportamiento de un sistema biológico. En el proceso de construcción de un modelo lo que se busca es realizar la traducción de un problema del mundo real a un problema matemático equivalente [1], lo que facilita el diseño del bioproceso. Así mismo, se debe tener en cuenta que en la actualidad hay modelos metodológicamente distintos entre sí.

Uno de los modelos matemáticos comúnmente utilizados son los modelos semi físicos de base fenomenológica (MSBF). Estos son modelos de caja gris producto de la combinación de modelos fenomenológicos y empíricos; son de base fenomenológica porque toma su estructura de los balances de materia, energía y cantidad de movimiento del proceso, y son semi físicos porque adicionan a la estructura ecuaciones empíricas para varios de sus parámetros llamadas ecuaciones constitutivas.

Ahora bien, otro tipo de modelo matemático son las redes neuronales (RNA) que se basa en las ciencias informáticas y está relacionado con el desarrollo de la Inteligencia Artificial. Un sistema informático de red neuronal se compone de experiencias y evidencias en forma de datos, con los que se puede comprender por sí mismo patrones o comportamientos. A partir de esto se puede elaborar predicciones de escenarios o iniciar operaciones que son la solución para una tarea específica [2].

Chlorella sorokiniana es una de las microalgas con mayor potencial biotecnológico, se caracteriza por ser cosmopolita, lo que permite que se adapte a diversas condiciones ambientales y nutricionales [3]. Esta microalga es una de las más estudiadas respecto a su fisiología, genética y por su producción de biomasa con alto contenido de lípidos, especialmente ácidos grasos insaturados omega 3, 6 y 9 empleados en áreas de cosmética, fármacos y biocombustibles [4], sin embargo, escalar, diseñar y controlar un cultivo de este microorganismo a gran escala representa grandes gastos en tiempo y dinero.

El aumento creciente de los costos asociados al escalado de estos sistemas dirige la atención hacia el uso de modelos matemáticos como los mencionados anteriormente, con el propósito de predecir y analizar el comportamiento de este microorganismo, como Vásquez y Villalobos en 2015 [5], donde evaluaron los modelos matemáticos de Gompertz y logístico en la cinética de crecimiento de *Spirulina Sp.* Sin embargo, el modelado de estos sistemas en la actualidad utilizando los modelos matemáticos en cuestión es un proceso en mejora, debido a que se desconocen antecedentes donde comparen estos modelos matemáticos con un proceso real y por esta razón, no se sabe que tan precisos son estimando el crecimiento de biomasa de *Chlorella sorokiniana*, además de ser modelos complejos y con diversas restricciones.

Por ello, este trabajo tiene como finalidad realizar una analogía entre los modelos semifísicos de base

fenomenológica (MSBF) y las RNA, utilizando como objeto de estudio el crecimiento de biomasa de *Chlorella sorokiniana*, y tomando como referencia datos experimentales provenientes de una investigación en la Universidad Icesi.

2. Metodología y resultados

La metodología llevada a cabo para los modelos semi físicos de base fenomenológica se basó en la propuesta de Hernán Álvarez y colaboradores en 2009. [6]

2.1. Modelo semifísico de base fenomenológica

2.1.1. Descripción verbal del proceso.

Para el proceso de producción de biomasa de *Chlorella sorokiniana* se emplea un volumen de operación de 250ml. Este es un proceso fermentativo en modo batch de digestato y glucosa, donde la microalga *Chlorella sorokiniana* convierte, en fase líquida y con presencia de oxígeno, estos azúcares en biomasa, agua y dióxido de carbono. Esta fermentación se lleva a cabo en Erlenmeyer de 1 L, con exposición constante a la luz UV y agitación por aireación.

Las corrientes de entrada al reactor son dos: una, en fase líquida, formada por el medio de cultivo inoculado; y otra, en fase gaseosa, formada por aire de los alrededores. Al igual que la anterior, las corrientes de salida son dos: una, en fase líquida, formada por el medio exhausto con la biomasa formada y agua; y otra, en fase gaseosa, formada principalmente por dióxido de carbono.

Figura 1. Diseño conceptual MSBF



2.1.2. Caracterización del modelo.

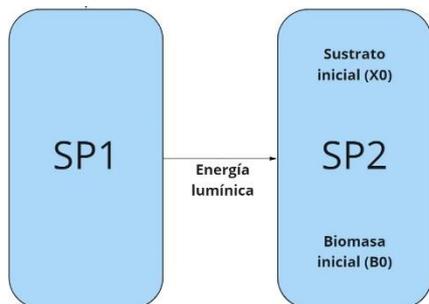
El modelo semifísico que describe este proceso fermentativo, tiene como objetivo evaluar el crecimiento de biomasa de la microalga *Chlorella sorokiniana*. Este modelo se construye con base en la aplicación del principio de conservación, y cinéticas de los procesos asociados, usando una recopilación de datos empíricos reportados en otras investigaciones sobre fermentaciones con dicha microalga. Por último, el modelo responde a la siguiente pregunta: ¿Cuál es el comportamiento de la concentración de biomasa de *Chlorella sorokiniana* durante el tiempo de reacción?

2.1.3. Definición de los sistemas de proceso.

Se han definido dos sistemas de proceso:

- SP1: Bombillas de luz UV que irradian el reactor proporcionando energía lumínica.
- SP2: Medio de cultivo homogéneo y agitado.

Figura 2. Sistemas de proceso MSBF



2.1.4. Principio de conservación.

- **Balance de biomasa:** La ecuación 1 representa la formación de biomasa durante el tiempo de reacción, donde μ es la velocidad específica de crecimiento de *Chlorella sorokiniana*, X es la

2.1.5. Definición de parámetros, constantes y variables del proceso.

Los parámetros cinéticos de crecimiento son las herramientas básicas para escalar los procesos biotecnológicos evaluados en el laboratorio, puesto que permiten predecir el desarrollo de la fermentación y evaluar los rendimientos y las productividades en los procesos [8]. Sin embargo, para este caso en particular, los datos experimentales que fueron provistos para la realización de modelo solo permitieron calcular el parámetro Y_{xs} (Ecuación 3) [9]. Por esta razón, y con el objetivo de que los datos obtenidos por el modelo fueran lo más cercano posible a los datos experimentales de referencia, se realizó un proceso de parametrización u optimización en MATLAB (Anexo 2), con el fin de determinar el valor de los parámetros cinéticos faltantes (u_{max} , k_s , I_{av} , I_k , m_s , n) que hacían que el modelo se ajustara mejor al proceso real. En la tabla 1 se muestran los valores de las constantes y parámetros del proceso.

Ecuación 3. Cálculo rendimiento biomasa/sustrato

$$Y_{xs} = \frac{X_f - X_0}{S_0 - S_f} = \frac{0,509361 - 0,121884}{9,14 - 1,7156} = 0,05219$$

Tabla 1. Constantes, variables y parámetros para MSBF.

	Símbolo	Valor	Unidades
Velocidad máxima de crecimiento	u_{max}	0,074927	1/h
Constante de limitación de sustrato	k_s	5,1799	g/L
Irradiancia promedio en el interior del reactor	I_{av}	280	uE/m^2*s
Irradiancia en la mitad de la velocidad máxima	I_k	25	uE/m^2*s
constante para microalgas	n	9	
Volumen del reactor	V	0,25	L
Rendimiento biomasa sustrato	Y_{xs}	0,05219	g/g
Volumen de muerte	V_k	0,0002	L
Constante de mantenimiento	m_s	0,9184	
Sustrato	S	-	g/L
Biomasa	X	-	g/L

concentración acumulada de biomasa y K_d es la constante de muerte celular. [7]

Ecuación 1. Balance de formación de biomasa.

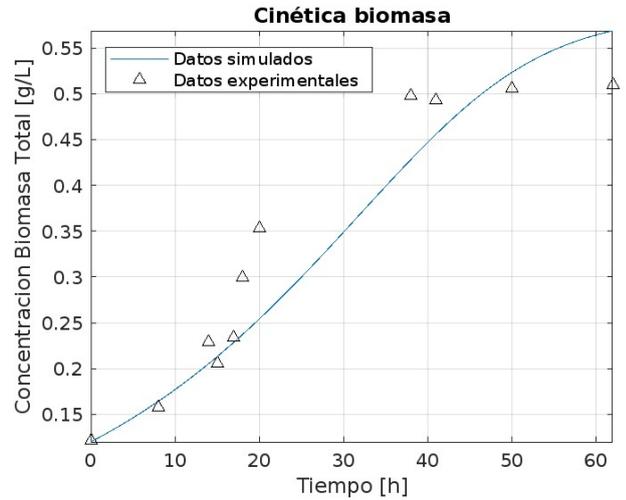
$$\frac{dX}{dT} = \mu * X - K_d * X$$

- **Balance de sustrato:** La ecuación 2 representa el consumo de sustrato, en este caso glucosa, durante el tiempo de reacción, donde Y_{xs} es el coeficiente de rendimiento que representa la cantidad de biomasa celular formada por unidad de masa de sustrato consumido, X es la concentración acumulada de biomasa y m_s es el coeficiente de mantenimiento. [7]

Ecuación 2. Balance de consumo de sustrato.

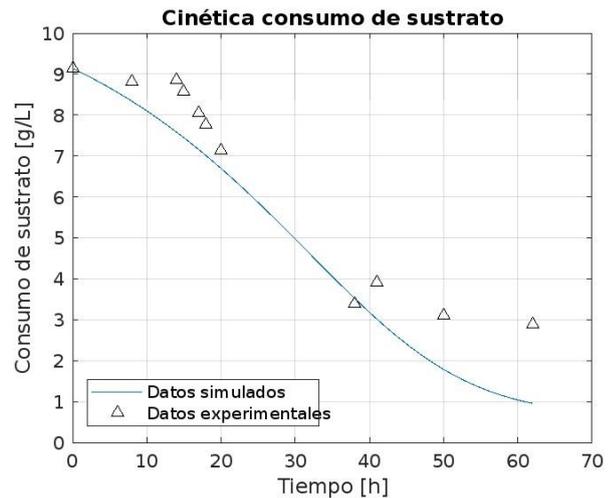
$$\frac{dS}{dT} = \left(\frac{1}{-Y_{xs}} \right) + m_s * X$$

Gráfica 1. Estimación producción de biomasa



Además de estimar la producción de biomasa, también se realizó la estimación del consumo de sustrato con el objetivo de que el modelo matemático fuera lo más cercano posible al cultivo real.

Gráfica 2. Estimación consumo de sustrato



2.1.6. Definición de las ecuaciones constitutivas.

- Velocidad específica de crecimiento:** La ecuación 4 es la primera ecuación constitutiva del modelo, esta representa la velocidad específica de crecimiento de la microalga, la cual se construyó empíricamente teniendo en cuenta dos de los factores que inciden en su crecimiento, sustrato e irradiancia. El parámetro μ_{max} representa la velocidad máxima de crecimiento de biomasa, la cual fue determinada en el proceso de parametrización. El segundo componente de esta ecuación representa la dependencia cinética de la tasa de crecimiento específica sobre la radiación fotosintéticamente, donde I_{av} representa la irradiancia promedio dentro del reactor, I_k representa la irradiancia en la mitad de la velocidad máxima y n es una constante [9]. El tercer y último componente de esta ecuación representa el comportamiento poblacional del microorganismo en función de la cantidad de sustrato (Monod).

Ecuación 4. Velocidad específica de crecimiento

$$\mu = \mu_{max} * \left(\frac{I_{av}^n}{I_k^n + I_{av}^n} \right) * \left(\frac{S}{S + K_s} \right)$$

- Constante de muerte:** La ecuación 5 representa la constante de muerte celular, donde V es el volumen del reactor, V_k es volumen de muerte que corresponde al 0.8% del volumen total del reactor.

Ecuación 5. Constante de muerte

$$K_d = (4 * V_k) / V$$

2.1.7. Simulación.

El modelo matemático fue simulado en MATLAB R2019b usando una función ode45. Los datos experimentales usados constan de 11 muestras en un periodo total de 62 horas (Anexo 1). A continuación, se muestra la curva solución que estima la producción de biomasa.

2.2. Modelamiento con Redes Neuronales

2.2.1. Elección de la Red neuronal y uso de datos.

Considerando que los datos experimentales provistos estaban desajustados debido a saltos de tiempo entre 10 y 15 horas por dato, se utilizó una aplicación de red neuronal, Net Fitting, preestablecida por MATLAB, la cual tiene como función resolver problemas de ajuste de datos. Esta red utiliza el algoritmo de entrenamiento y ajuste de datos con Levenberg Marquardt (LM).

2.2.2. Construcción de Datos para entrenamiento

Los datos de crecimiento de *Chlorella sorokiniana* y consumo de sustrato provistos por la Universidad Icesi se muestra en la tabla 2. Los cultivos de esta microalga fueron realizados en fotobiorreactores de

250 mL con temperatura y pH estables, agitación por aireación y luz UV.

Tabla 2. Resultados experimentales

Tiempo (h)	Azúcares (g/L)	Biomasa (g/L)
0	9.14	0.13
14	8.80	0.23
16	15.76	0.16
17	8.09	0.27
20	7.12	0.35
23	6.58	0.40
26	3.64	0.46
40	2.13	0.29
50	3.10	0.50
62	1.71	0.51

Teniendo en cuenta que, la implementación de redes neuronales de MATLAB acepta una secuencia de datos continuos, discretos y cuya separación de tiempo sea secuencial y, que a partir de la tabla 2 se observó que los intervalos de tiempo entre valores variaban entre 10 y 15 horas, no siendo secuenciales, se utilizó la función fill missing de MATLAB (Ecuación 6), la cual rellena los posibles valores de entrada y salida faltantes realizando interpolaciones, utilizando una media y una ventana móvil con una longitud de ventana, que para este proceso fue 5 pues fue con la cual se logró completar los valores necesarios para el entrenamiento de la red.

Ecuación 6. Función de interpolación por el software MATLAB

F = fill missing (A, movmean, window)

De esta manera se construyeron los datos y se determinaron los valores intermedios necesarios para la adaptación del modelo sea mucho óptima, considerando que se realizaron pocas interpolaciones para evitar una mayor dispersión con los datos reales. Al final, se obtuvo un 39% de los datos por interpolación los cuales fueron necesarios para lograr perfilar el modelo, donde la mayoría de estos se utilizaron en el testeo adicional. (Anexo 6) Finalmente, la cantidad de datos utilizados para el entrenamiento y testeo adicional se muestran en la tabla 3

Tabla 3. Datos experimentales reconstruidos.

Tiempo (h)	Azúcares (g/L)	Biomasa (g/L)
0	9.14	0.12
0	9.30	0.12
2	9.22	0.13
5	9.30	0.15
8	8.80	0.22
11	8.82	0.19
14	8.80	0.22
14	8.84	0.15
15	8.58	0.23

16	8.32	0.22
17	8.09	0.31
17	8.05	0.20
18	7.75	0.29
18	7.65	0.27
20	7.12	0.37
20	7.79	0.23
23	6.58	0.41
23	7.00	0.26
26	6.24	0.29
38	3.39	0.49
41	3.90	0.49
50	3.10	0.50
62	1.71	0.51

Para el entrenamiento se utilizaron 23 datos separados en tres grupos, uno destinado al entrenamiento de la red preestablecida, otro para la evaluación del modelo y el último para test, fraccionados en una proporción 70%, 15% y 15% respectivamente, los 19 datos restantes se utilizaron en un testeo adicional. Aquí se ingresaban datos experimentales de sustrato como entrada (X) correspondientes a datos de salida de biomasa (Y) experimentales, tal como se muestra en la figura 3.

Figura 3. Esquema de entrenamiento de la Red Neuronal



2.2.3. Entrenamiento de la máquina.

Se utilizó el primer set de datos (23 datos), para entrenar la máquina y a su vez lograr una mejora incremental en el grado de predicción. Se ingresaron aleatoriamente los valores del modelo en formato matriz o cell array y se tuvieron en consideración la relación entre las entradas (predictors) y las salidas (Responses), ya que estas varían según los datos que se le suministren (Figura 4). En este caso el algoritmo de entrenamiento usado fue Levenberg-Marquardt. Además, se realizó un testeo adicional con 19 datos más para evitar un sobre ajuste en la red preestablecida, como se muestra en la figura 5.

Figura 4. Resultados del entrenamiento arrojado por el software.

	Observations	MSE	R
Training	17	1.8734	0.9950
Validation	3	5.3242	0.9929
Test	3	135.5040	0.9733

Figura 5: Resultados del entrenamiento arrojado por el software

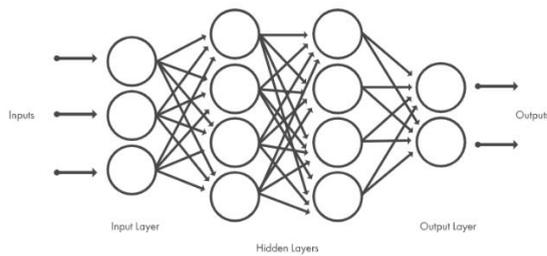
	Observations	MSE	R
Additional test	19	2.4385	0.9847

2.2.4. Evaluación.

La red neuronal combina diversas capas de procesamiento y realiza operaciones en paralelo en cada una de las neuronas a partir de los valores de los Layers obtenidos del entrenamiento, estas se encuentran interconectadas entre sí y cada capa utiliza la salida de la capa anterior como entrada, tal como se muestra en la Figura 6.

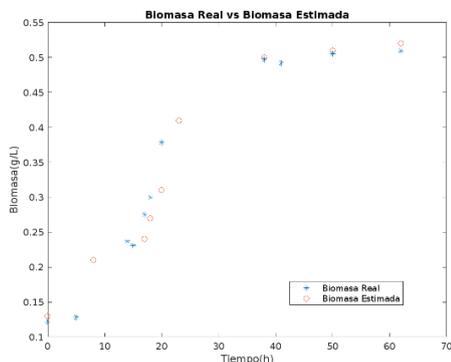
El modelo matemático fue simulado en MATLAB R2019b. Esta red neuronal fue un modelo básico de caja negra donde se evaluó el error cuadrático medio (MSE) de los valores de entradas y salidas. (Anexo 3).

Figura 6: Arquitectura típica de una red neuronal



A partir del entrenamiento se obtuvieron los Layers de cada neurona (Anexo 4), estos permitieron que se procesaran los datos experimentales y se fijaran para la validación. En el proceso de validación se ingresaron datos de entrada (x) de sustrato en el tiempo y a partir de cada Layers se estimó los datos de salida (Y) de biomasa, los cuales se compararon con los datos reales del proceso y así conocer el grado de predicción calculando el porcentaje de error entre cada punto experimental y estimado mostrados en la Tabla 4. Así, se graficó la cinética de crecimiento de biomasa experimental y estimada por el modelo como se muestra en la gráfica 3.

Gráfica 3: Cinética de crecimiento de *Chlorella sorokiniana* (Red neuronal)



3. Discusiones y conclusiones

Tabla 4: Cálculo de % error promedio de MSBF

Tiempo (h)	Biomasa (Experimentación)	Biomasa (Simulación)	% Error
0	0,12	0,12	0,00
8	0,22	0,20	9,09
17	0,23	0,23	0,00
18	0,24	0,24	0,00
20	0,38	0,25	34,21
23	0,41	0,27	34,15
38	0,49	0,42	14,29
50	0,50	0,52	4,00
62	0,51	0,56	10,65
Error promedio (%)			11,82

Teniendo en cuenta la gráfica 1 y la tabla 4, donde se evidencia el poco ajuste de los datos obtenidos por el MSBF con los datos experimentales, es pertinente traer a colación las posibles causas por las cuales se presenta esta situación.

Existen parámetros de cultivo que se deben tener en cuenta al momento de modelar un cultivo, en este caso, la fuente de carbono, fuente de nitrógeno, irradiancia, pH, temperatura y la agitación son factores relevantes para el crecimiento de biomasa de *Chlorella sorokiniana*. Considerando que para el MSBF sólo se tuvieron en cuenta la irradiancia y la fuente de carbono debido a la falta de datos experimentales relacionados con los factores restantes, es importante resaltar la trascendencia de estos en el crecimiento del microorganismo.

El pH es uno de los factores más importantes en el cultivo microalgal. Las membranas plasmáticas de las células microalgales no son libremente permeables a los iones hidrógeno e hidroxilo, por lo tanto, las concentraciones de hidrogeniones intracelular y extracelular no están necesariamente equilibradas y existe un gradiente de concentración de hidrogeniones a través de la membrana. Este parámetro también se ve afectado por el consumo de nutrientes y la solubilidad de ciertos componentes como el CO₂, por ello se acostumbra a usar algún tipo de buffer, con el fin de provocar un efecto tampón y mantener el pH en valores adecuados por el microorganismo. [10]

La temperatura es otro parámetro fundamental para el crecimiento de las microalgas. La biomasa microalgal responde continuamente a la temperatura ambiental. Además de afectar a las reacciones celulares, la temperatura también afecta a la naturaleza del metabolismo, los requerimientos nutricionales y la composición de la biomasa, si bien dentro de los rangos óptimos tiene poca influencia sobre la concentración final de biomasa, así como sobre la producción y la composición bioquímica de las microalgas.

Por último, la agitación. Una agitación suficiente del medio de cultivo es necesaria e incide directamente en el cultivo de este microorganismo. Cuando los

requerimientos están satisfechos y las condiciones ambientales son satisfactorias, la agitación constituye el requisito más importante para la obtención de altos rendimientos de biomasa microalgal [10]. La agitación produce el movimiento del agua, lo que implica una serie de efectos positivos tales como: Asegurar una distribución homogénea de las células y los nutrientes dentro del cultivo, con los nutrientes rápidamente disponibles para las células; mejorar la distribución de la luz a las células asegurando que permanezcan fotosintéticamente activas; y, por último, evitar que las células sedimenten en el fondo del recipiente de cultivo, produciendo una estimulación general del metabolismo celular.

Lo anterior permite entender el poco ajuste del modelo matemático con los datos experimentales, además es pertinente explicar que, al no conocer el comportamiento de la biomasa con respecto a los factores no incluidos, se hace muy complicada la construcción de ecuaciones que puedan definirlos.

Tabla 5: Cálculo de % error promedio de RNA

Tiempo (h)	Biomasa (Experimentación)	Biomasa (Simulación)	% Error
0	0,12	0,13	8,33
8	0,22	0,21	4,55
17	0,23	0,24	4,35
18	0,24	0,27	12,50
20	0,38	0,31	8,42
23	0,41	0,41	0,00
38	0,49	0,5	2,04
50	0,50	0,51	2,00
62	0,51	0,52	1,96
		Error promedio	6,01 (%)

La tabla 5 muestra que el porcentaje de error el porcentaje de error promedio obtenido por el modelo de RNA es de 8,08 %. Este porcentaje de error puede ser producto de la poca cantidad de datos con que se entrenó la red neuronal, ya que su efectividad depende del volumen de información con que se alimenta, y de esta forma mejorar la eficiencia en cada predicción. (Gisi, 2018) Al haber tan poca cantidad de datos para la entrada (42 datos), hace que el resultado en la ejecución del algoritmo sea limitado. Es precisamente debido a estas situaciones que surge la necesidad de entrenar la red neuronal con bases de datos apropiadas y de tamaño suficiente para poder generar un modelo que logre caracterizar de manera efectiva el fenómeno en cuestión. (Gisi, 2018).

Teniendo en cuenta las gráficas 1 y 3, y basadas en los cálculos del error promedio, se puede evidenciar que el modelo matemático de RNA presenta un mejor ajuste que el MSBF, en los datos estimados con respecto a los datos experimentales. Esto debido a que, para los MSBF no se tuvo en cuenta gran cantidad de factores influyentes para la simulación de la cinética de crecimiento de *Chlorella sorokiniana*, además los valores de los parámetros fueron definidos a partir una parametrización, la cual tuvo como objetivo buscar los valores óptimos de cada parámetro, pero no garantiza que estos sean los valores cinéticos que definen el sistema.

A partir de esto se logró construir la tabla 6, la cual muestra diferencias y/o similitudes entre ambos modelos matemáticos, teniendo en cuenta criterios importante como: Principios de estimación, revisión de literatura, adaptación del modelo y precisión de estimación.

Tabla 5. Diferencias y/o similitudes entre los modelos matemáticos

Comparativa		
Criterio	MSBF	RNA
Principios de estimación	Combinación de modelos fenomenológicos y empíricos. Son de base fenomenológica porque toman su estructura de los balances de materia, energía y cantidad de movimiento, y son semifísicos porque para varios de sus parámetros adicionan a la estructura ecuaciones empíricas llamadas ecuaciones constitutivas	Se relaciona con el desarrollo de la Inteligencia Artificial, y que sirve para crear sistemas que pueden aprender por sí solos. Un sistema informático de aprendizaje automático se sirve de experiencias y evidencias en forma de datos, con los que comprender autónomamente patrones o comportamientos. De este modo, puede elaborar predicciones de escenarios o iniciar operaciones que son la solución para una tarea específico
Revisión de literatura	Se encuentra mucha información sobre usos y aplicaciones del modelo, no obstante, para una especie de microalga en un sistema de cultivo en	Se encuentra información sobre el uso e ingreso de data sets en RNA, sin embargo, es complejo encontrar la cantidad de datos

	particular es complejo encontrar ecuaciones y parámetros que definan todo el sistema, lo cual hace que el proceso de revisión bibliográfica sea más extenso.	adecuada para realizar el entrenamiento de la máquina.
Adaptación del modelo	Es más complejo ya que, al no tener los parámetros cinéticos de crecimiento es necesario hacer procesos de optimización (parametrización) con el fin de determinar los valores que permitan que el modelo se asemeje al proceso real.	Solo se deben adaptar los datos a la red neuronal ya existente. Sin embargo, es necesario un alto volumen de datos.
Precisión de estimación	A mayor cantidad de ecuaciones que representen los factores incidentes o relevantes en el sistema, mayor precisión.	A mayor cantidad de datos para el entrenamiento (aprendizaje de la máquina), mayor precisión.

3.1. Conclusiones y recomendaciones.

Con el proyecto se concluye que es posible estimar la concentración de biomasa de *Chlorella sorokiniana* a partir de ambos modelos matemáticos. Además, las redes neuronales mostraron un mayor grado de predicción que los modelos semifísicos de base fenomenológica, esto a partir del error promedio los cuales fueron 6.01% y 11.82% respectivamente.

Para futuras investigaciones se recomienda tener mayor cantidad de datos experimentales para el entrenamiento de la red neuronal, pues depende de estos datos el aprendizaje mismo de la máquina (Backpropagation), buscando así mejores estimaciones para la cinética de crecimiento. Así mismo, se recomienda utilizar este artículo en futuras investigaciones realizando el proceso experimental donde se pueda tener control de todos los factores necesarios del proceso y cantidad de toma de datos necesaria para los modelos.

4. Agradecimientos

Los realizadores de este trabajo agradecen a la Ingeniera Bioquímica Claudia M. Sichel y PhD Erika Y. Ortiz, por compartir los datos experimentales del crecimiento de *Chlorella sorokiniana* de su investigación llevada en la Universidad Icesi.

5. Referencias

- [1] M. R. Botero, «Modelo Semifísico de Base Fenomenológica de la transferencia de oxígeno para el tratamiento de aguas residuales en una planta piloto aireada por difusión,» Medellín, 2019.
- [2] «Redacción APD,» 2019. [En línea]. Available: <https://www.apd.es/que-es-machine-learning/>.
- [3] S. R. Wehr J, «Freshwater algae of North America: Ecology and classification,» San diego, 2003.
- [4] O. J. T. H. Ugwu C, «Characterization of light utilization and biomass yields of *Chlorella sorokiniana* in inclined outdoor tubular photobioreactors equipped with static mixers.,» 2005, pp. 3406-3411.
- [5] V. A. G. P. C. A. K. & E. T. A. (. Vázquez-Villalobos, «Mathematical modeling and through artificial neural networks of the spirulina sp. growth in a photobioreactor with fluorescent light source and solid state light,» Perú, 2013.
- [6] R. L. Hernand D. Alvarez, «Metodología para la Obtención de Modelos Semifísicos de Base Fenomenológica Aplicada a una Sulfitadora de Jugo de Caña de Azúcar,» Medellín, 2009.
- [7] J. A. C. G. Alejandro RINCÓN SANTAMARÍA, «CINÉTICA DE CRECIMIENTO DE GLUCONACETOBACTER DIAZOTROPHICUS USANDO MELAZA DE CAÑA Y SACAROSA: EVALUACIÓN DE MODELOS CINÉTICOS,» *ACTA BIOLÓGICA COLOMBIANA*, 2018.
- [8] M. H. R. José E ZAPATA, «PARÁMETROS CINÉTICOS DE CRECIMIENTO DE *Saccharomyces cerevisiae* EN PRESENCIA DE UN CAMPO MAGNÉTICO VARIABLE DE BAJA INTENSIDAD Y ALTA FRECUENCIA,» *REVISTA DE LA FACULTAD DE QUÍMICA FARMACÉUTICA*, 2005.
- [9] M. P. Alessandro Concas, «Mathematical Modelling of *Chlorella Vulgaris* Growth in Semi-Batch Photobioreactors Fed with Pure CO₂,» *CHEMICAL ENGINEERING TRANSACTIONS*, 2013.
- [10] A. C. J. F. E. T. y. C. H. Julio Abalde, «Microalgas: cultivos y aplicaciones,» Coruña, pp. 20-39.
- [11] Gisi, «ACERVO PARA EL MEJORAMIENTO DEL APRENDIZAJE DE ALUMNOS DE INGENIERÍA, EN INTELIGENCIA ARTIFICIAL,» 2018. [En línea]. Available: http://virtual.cuautitlan.unam.mx/intar/?page_id=1060.
- [12] M. Jimenez, «Microalgal biomass with high potential for the biofuels production,» Trujillo, 2021.
- [13] C. Y., «Biodiesel from microalgae,» 2007.
- [14] J.-C. C. D. E. I. A. Spolaore P, «Commercial applications of microalgae.,» *Journal of Biosciences and Bioengineering.*, 2006.
- [15] H. Pérez, «Microalgae, culture and benefits.,» Chile, 2014.
- [16] G. D. Singh S, «Microbial and plant derived biomass for removal of heavy metals from wastewater.,» *Bioresources Technology.*, pp. 2343-2357, 2007.

6. Anexos



Código Modelo Semifísico.zip

Anexo 1: Código MSBF



Código Parametrización.zip

Anexo 2: Código Parametrización



Código RNA.m

Anexo 3: Código RNA



LAYERS.m

Anexo 4: Layers



Tablas de datos
experimentales.docx

Anexo 5: Datos para entrenamiento