



SEP

SES

TecNM

INSTITUTO TECNOLÓGICO DE TOLUCA

**PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS EN INGENIERÍA AMBIENTAL**

TESIS

**“MODELADO NEURONAL ARTIFICIAL PARA EL PROCESO
DE ADSORCIÓN POR LOTES DEL COLORANTE
AMARILLO NO. 5 EN SOLUCIÓN ACUOSA MEDIANTE EL
HIDROGEL DE QUITOSANO-POLI(VINIL ALCOHOL)”**

PRESENTA:

SALVADOR DOMÍNGUEZ BELTRÁN

No. CONTROL:

1928M0256

DIRECTOR:

DR. JOSÉ LUIS GARCÍA RIVAS

CODIRECTOR:

DR. ROBERTO ALEJO ELEUTERIO

METEPEC, ESTADO DE MÉXICO, JUNIO DE 2026



Instituto Tecnológico de Toluca
División de Estudios de Posgrado e Investigación

Metepec, Edo. de México, 05/Junio/2026
DEPI-3200-359/2026

SALVADOR DOMINGUEZ BELTRÁN
CANDIDATO AL GRADO DE MAESTRO
EN CIENCIAS EN INGENIERÍA AMBIENTAL

PRESENTE

De acuerdo con los Lineamientos para la Operación de los Estudios de Posgrado en el Tecnológico Nacional de México y las disposiciones en este Instituto, habiendo cumplido con todas las indicaciones que la Comisión Revisora realizó con respecto a su Trabajo de Tesis titulado **"MODELADO NEURONAL ARTIFICIAL PARA EL PROCESO DE ADSORCIÓN POR LOTES DEL COLORANTE AMARILLO No. 5 EN SOLUCIÓN ACUOSA MEDIANTE EL HIDROGEL DE QUITOSANO-POLI(VINIL ALCOHOL)"**, la División de Estudios de Posgrado e Investigación de este Instituto, concede la Autorización para que proceda al envío del trabajo final vía correo Institucional.

Sin más por el momento, quedo de usted.

ATENTAMENTE

Excelencia en Educación Tecnológica®
"Educación, integridad y ciencia"


JORGE SÁNCHEZ JAIME
JEFE DE LA DIVISIÓN DE ESTUDIOS
DE POSGRADO E INVESTIGACIÓN



ccp. Archivo
JSJ/mccl



2026
año de
Margarita
Maza





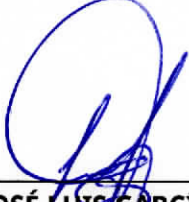
Metepec, Edo. de México, 05/junio/2026
DEPI-3200-360/2026

DR. JORGE SÁNCHEZ JAIME
JEFE DE LA DIVISIÓN DE ESTUDIOS
DE POSGRADO E INVESTIGACIÓN

PRESENTE

Por este medio comunicamos a usted que la comisión Revisora designada para analizar la Tesis denominada **"MODELADO NEURONAL ARTIFICIAL PARA EL PROCESO DE ADSORCIÓN POR LOTES DEL COLORANTE AMARILLO No. 5 EN SOLUCIÓN ACUOSA MEDIANTE EL HIDROGEL DE QUITOSANO-POLI(VINIL ALCOHOL)"**, que como parte de los requisitos para obtener el grado académico de **MAESTRO EN CIENCIAS EN INGENIERÍA AMBIENTAL** presenta el **C. SALVADOR DOMINGUEZ BELTRÁN** con número de control **1928M0256** para sustentar el acto de Recepción Profesional, ha dictaminado que dicho trabajo reúne las características de contenido y calidad para proceder a la impresión del mismo.

ATENTAMENTE


DR. JOSÉ LUIS GARCÍA RIVAS
DIRECTOR DE TESIS


DR. ROBERTO ALEJO ELEUTERIO
CODIRECTOR DE TESIS


DR. FRANCISCO JAVIER ILLESCAS MARTÍNEZ
REVISOR DE TESIS


DRA. MARIA SONIA MIREYA MARTÍNEZ
GALLEGOS
REVISORA DE TESIS



ccp. Archivo
JSJ/mccl



2026
año de
Margarita
Maza



RESUMEN

La liberación de aguas residuales industriales que contienen colorantes sintéticos plantea un problema ambiental importante debido a su toxicidad y persistencia. Entre las opciones de tratamiento, los materiales naturales, específicamente el hidrogel de quitosano-alcohol polivinílico (quitosano-PVA), han demostrado una alta eficacia en la eliminación de colorantes debido a sus abundantes grupos funcionales y capacidad de adsorción comprobada. Sin embargo, la optimización experimental de estos sistemas suele ser laboriosa y requiere muchos recursos. Este estudio introduce un modelo de red neuronal artificial (RNA) para predecir la capacidad de adsorción (q_e) y el tiempo necesario para alcanzar el equilibrio durante la eliminación del colorante tartrazina utilizando perlas de hidrogel de quitosano-PVA de diferentes tamaños promedio, categorizados como pequeños, medianos y grandes (2,1, 2,5 y 3,2 mm, respectivamente) a temperaturas de 10, 30 y 50 °C. El modelo RNA se comparó con modelos cinéticos tradicionales: pseudo-primer orden, pseudo-segundo orden y Elovich. Los resultados mostraron que la ANN superó a los modelos convencionales en la predicción de q_e y el tiempo de equilibrio, especialmente para perlas pequeñas a 10 °C, donde $q_e = 945$ mg/g en 40 h con un R^2 de 0.9428. En todas las condiciones, la ANN logró altos coeficientes de correlación ($R^2 > 0.94$) y acortó significativamente los tiempos de predicción. Aunque el modelo de pseudo-segundo orden alcanzó altos valores de R^2 (hasta 0.9929), tardó más de 72 h en alcanzar la predicción de equilibrio. Estos resultados demuestran que el modelado basado en ANN puede reducir el esfuerzo experimental hasta en un 50 % en el tiempo de predicción, manteniendo una alta capacidad predictiva ($R^2 > 0.94$), lo que ofrece un enfoque sostenible y eficiente para el diseño de procesos de tratamiento de aguas residuales.

Abstract

The release of industrial wastewater containing synthetic dyes poses a major environmental issue because of their toxicity and persistence. Among treatment options, natural materials, specifically chitosan–polyvinyl alcohol (chitosan–PVA) hydrogel, have shown high effectiveness in dye removal due to their abundant functional groups and proven adsorption capacity. However, optimizing these systems experimentally is often time-consuming and requires many resources. This study introduces an artificial neural network (ANN) model to predict the adsorption capacity (q_e) and the time needed to reach equilibrium during the removal of tartrazine dye using chitosan–PVA hydrogel beads of different mean sizes, categorized as small, medium and large (2.1, 2.5, and 3.2 mm, respectively) at temperatures of 10, 30, and 50 °C. The ANN model was compared with traditional kinetic models: pseudo-first-order, pseudo-second-order, and Elovich. Results showed that the ANN outperformed conventional models in predicting q_e and equilibrium time, especially for small beads at 10 °C, where it predicted $q_e = 945$ mg/g in 40 h with an R^2 of 0.9428. Across all conditions, the ANN achieved strong correlation coefficients ($R^2 > 0.94$) and significantly shortened prediction times. Although the pseudo-second-order model achieved high R^2 values (up to 0.9929), it took over 72 h to reach equilibrium prediction. These results demonstrate that ANN-based modeling can reduce experimental effort by up to 50% in prediction time while maintaining high predictive accuracy ($R^2 > 0.94$), offering a sustainable and efficient approach for designing wastewater treatment processes.

CONTENIDO

ÍNDICE DE FIGURAS.....	7
ÍNDICE DE TABLAS.....	8
INTRODUCCIÓN	9
1 FUNDAMENTOS	11
1.1 Colorantes sintéticos en la industria.....	11
1.1.1 Colorante amarillo No. 5.....	12
1.1.2 Tratamientos para la remoción de colorantes sintéticos en aguas residuales	13
1.1.3 Adsorción	14
1.1.4 Quitosano.....	15
1.1.5 Cinéticas de adsorción	16
1.2 Red neuronal artificial (RNA).....	18
2 METODOLOGÍA	24
2.1 Base de datos (BD) de entrenamiento	28
2.2 Identificación de entradas y salidas.....	31
2.3 Adaptación de la arquitectura de la RNA.....	32
2.4 Diseño de la RNA	33
2.5 Entrenamiento y validación de la RNA	35
2.6 Modelo de la RNA	37
3 RESULTADOS Y DISCUSIÓN	40
3.1 Comparativo de RNA con modelos tradicionales	41
Resultados del valor <i>qe</i> para la esfera chica	52
Resultados del valor <i>qe</i> para la esfera mediana.....	54
Resultados del valor <i>qe</i> para la esfera grande	55
Discusión de resultados	57
4 CONCLUSIONES	59
5 REFERENCIAS	62

ÍNDICE DE FIGURAS

Figura 1. Industrias responsables de la presencia de colorantes en el agua (Katheresan et al., 2018).....	12
Figura 2. Estructura química de la quitina, quitosano y quitano (Sahnoun & Boutahala, 2018).	16
Figura 3. (a) Estructura de una neurona y (b) una Red Neuronal Artificial (Pacheco-Sánchez, Alejo, Cruz-Reyes, & Álvarez-Ramírez, 2019).	20
Figura 4. Diagrama de flujo de los procesos de formación de la RNA.....	28
Figura 6. Arquitectura de la red.	35
Figura 7. Resultados de predicción para la esfera chica.....	42
Figura 8. Resultados de predicción para la esfera mediana.	43
Figura 9. Resultados de predicción para la esfera grande.....	44
Figura 10. Modelo de RNA para esfera chica a 10 ° C. Con 10 y 10 neuronas en las capas ocultas.	45
Figura 11. Modelo de RNA para esfera chica a 10, 30 y 50 °C, con 2 y 2 neuronas en las capas ocultas.....	46
Figura 12. Modelo de RNA para esfera chica a 10, 30 y 50 °C, con 10 y 10 neuronas en las capas ocultas.	47
Figura 13. Modelo de RNA para esfera chica a 10, 30 y 50 °C, con 20 y 20 neuronas en las capas ocultas.	48
Figura 14. Modelo de RNA para esfera chica, mediana y grande a 10, 30 y 50 °C, con 2 y 2 neuronas en las capas ocultas.	49
Figura 15. Modelo de RNA para esfera chica, mediana y grande a 10, 30 y 50 °C, con 10 y 10 neuronas en las capas ocultas.	50
Figura 16. Modelo de RNA para esfera chica, mediana y grande a 10, 30 y 50 °C, con 20 y 20 neuronas en las capas ocultas.	51

ÍNDICE DE TABLAS

Tabla 1. Propiedades del colorante amarillo No. 5 (Goscianska & Ciesielczyk, 2019).	13
Tabla 2. Modelos cinéticos y ecuaciones de los modelos.	17
Tabla 3. Datos experimentales.	30
Tabla 4. Comparativa de resultados de q_e para la esfera chica con la temperatura 10°. ...	52
Tabla 5. Comparativa de resultados de q_e para la esfera chica con la temperatura 30°. ...	53
Tabla 6. Comparativa de resultados de q_e para la esfera grande con temperatura 50°.	53
Tabla 7. Comparativa de resultados de q_e para la esfera mediana con la temperatura 10°.	54
Tabla 8. Comparativa de resultados de q_e para la esfera mediana con la temperatura 30°.	54
Tabla 9. Comparativa de resultados de q_e para la esfera mediana con la temperatura 50°.	55
Tabla 10. Comparativa de resultados de q_e para la esfera grande con la temperatura 10°.	56
Tabla 11. Comparativa de resultados de q_e para la esfera grande con la temperatura 30°.	56
Tabla 12. Comparativa de resultados de q_e para la esfera grande con la temperatura 50°.	57

INTRODUCCIÓN

La descarga de aguas residuales industriales en ambientes acuáticos representa un problema ambiental de relevancia mundial debido a la presencia de contaminantes persistentes y tóxicos, entre ellos los colorantes sintéticos (Ullah et al., 2020; Wong et al., 2020). Estos compuestos pueden afectar negativamente a los ecosistemas acuáticos y a la salud humana, incluso cuando se encuentran en bajas concentraciones. Para reducir este impacto, se han desarrollado diferentes métodos de tratamiento, como la adsorción, la fotodegradación, la ozonización, la filtración por membranas y la ósmosis inversa (Ghaedi & Vafaei, 2017; Abdi et al., 2017). Entre estos métodos, la adsorción destaca por su alta eficiencia, simplicidad operativa y bajo costo para la remoción de contaminantes como colorantes, metales pesados y compuestos orgánicos (Karam et al., 2020).

Los adsorbentes naturales, tales como residuos agrícolas, arcillas y biopolímeros como el quitosano, han recibido creciente atención debido a su sostenibilidad, biodegradabilidad y baja toxicidad (Ullah et al., 2020; Karam et al., 2020). El quitosano, derivado de la quitina, ha demostrado un desempeño favorable en la remoción de colorantes debido a la presencia de grupos funcionales amino e hidroxilo, los cuales favorecen interacciones electrostáticas con colorantes aniónicos como la tartrazina (Crini & Badot, 2008). Por ejemplo, se ha reportado una capacidad de adsorción de hasta 584 mg/g para la remoción de tartrazina mediante un compuesto de quitosano/polianilina (Sahnoun & Boutahala, 2018). Sin embargo, la optimización experimental de los procesos de adsorción suele requerir una cantidad considerable de tiempo y recursos, lo que limita el desarrollo rápido de estos sistemas (Katheresan et al., 2018). En este sentido, el modelado computacional se ha convertido en una herramienta útil para predecir el comportamiento del sistema y disminuir la carga experimental.

Las redes neuronales artificiales (RNA) han despertado interés por su capacidad para modelar relaciones complejas y no lineales sin requerir ecuaciones físicas explícitas. A diferencia de los modelos cinéticos tradicionales, como los modelos de

pseudo-primer orden y pseudo-segundo orden, o de isothermas como Langmuir y Freundlich, los cuales dependen de supuestos simplificados, las RNA pueden incorporar múltiples variables, tales como pH, temperatura, tiempo de contacto, concentración inicial y dosis de adsorbente, para generar predicciones precisas (Ghaedi & Vafaei, 2017).

Diversos estudios han utilizado con éxito las RNA en procesos de adsorción. Pauletto et al. (2020) desarrollaron una RNA para el modelado simultáneo de adsorción multicomponente, mejorando la precisión predictiva en sistemas complejos. Iftikhar et al. (2023) emplearon redes neuronales para analizar la capacidad de adsorción de colorantes industriales utilizando materiales carbonáceos, obteniendo valores de R^2 superiores a 0.95 y bajos errores cuadráticos medios. Alardhi et al. (2023) mostraron una alta concordancia entre las predicciones de RNA y los datos experimentales en la adsorción de colorantes sobre materiales naturales. Asimismo, Al-Hameed et al. (2023) combinaron la metodología de superficie de respuesta (RSM) con RNA para optimizar la remoción del colorante Reactive Yellow 105 mediante un adsorbente basado en materiales zeolíticos, alcanzando valores de R^2 superiores a 0.96. De manera similar, Karam et al. (2020) demostraron la efectividad de las RNA para comparar adsorbentes en la remoción de colorantes textiles, con eficiencias cercanas al 100% y coeficientes de determinación superiores a 0.90. Además, da Costa et al. (2025) destacaron el potencial de las RNA para reducir el número de ensayos experimentales y apoyar el diseño sostenible de procesos de adsorción.

En este contexto, el presente estudio propone un modelo de RNA para predecir la adsorción de tartrazina, también conocida como FD&C Yellow No. 5, utilizando perlas de hidrogel de quitosano–alcohol polivinílico (quitosano–PVA). El modelo tiene como objetivo estimar tanto la capacidad de adsorción en equilibrio (q_e) como el tiempo necesario para alcanzar dicho equilibrio bajo diferentes condiciones de operación, específicamente el tamaño de las perlas y la temperatura. Este enfoque permite reducir la dependencia de una experimentación extensa y contribuye al diseño de procesos de tratamiento de aguas residuales más sostenibles y eficientes.

La relevancia científica de este trabajo radica en que integra dos dimensiones complementarias del problema: por un lado, el comportamiento fisicoquímico de un sistema de adsorción basado en un hidrogel de quitosano-PVA y, por otro, la construcción de un modelo neuronal capaz de representar la evolución cinética del proceso. Esta integración es pertinente porque la adsorción de colorantes en matrices poliméricas hidratadas no depende únicamente de la disponibilidad de sitios activos, sino también de fenómenos de transporte, hinchamiento, difusión interna, heterogeneidad superficial e interacciones electrostáticas. En consecuencia, el uso de una RNA permite abordar la no linealidad del sistema sin imponer desde el inicio una estructura cinética rígida, lo cual fortalece la capacidad predictiva del modelo frente a modelos empíricos convencionales (Ghaedi & Vafaei, 2017; Dutta et al., 2021; Rápó & Tonk, 2021; Domínguez Beltrán et al., 2025).

1 FUNDAMENTOS

1.1 Colorantes sintéticos en la industria

Los colorantes sintéticos son una necesidad en varias industrias importantes, como la industria textil, alimenticia, farmacéutica, papelera, fabricantes de colorantes entre otras, por sus propiedades de coloración. Se estima que cada año se fabrican 700,000 toneladas de diversos colorantes a partir de unos 100,000 colorantes comercialmente accesibles. A menudo, una vez que los colorantes han cumplido su propósito, la mayoría de ellos se desechan sin mayor cuidado en los cuerpos de agua naturales o en sitios de disposición final.

Se sabe de la existencia de cinco industrias principales que son responsables de la presencia de colorantes en cuerpos de agua naturales. (Abdi et al., 2017) & (Katheresan et al., 2018). La industria textil (54%) libera la mayor cantidad de colorantes, lo que contribuye a más de la mitad de los colorantes en efluentes existentes en el medio ambiente en todo el mundo. También se sabe que la industria del teñido (21%), la industria del papel y la pulpa (10%), la industria de curtiembres y pinturas (8%) y la industria de la fabricación de pintura (7%) producen altas

cantidades de colorantes en efluentes a partir de varios procesos asociados (ver Figura 1).

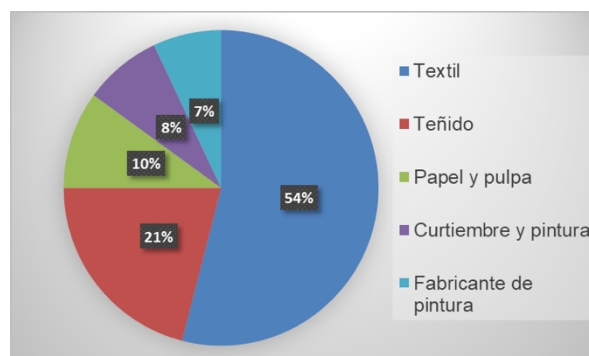


Figura 1. Industrias responsables de la presencia de colorantes en el agua (Katheresan et al., 2018).

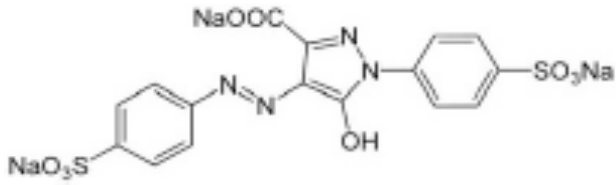
La cantidad exacta de colorantes en efluentes expulsados por cada industria en el medio ambiente es desconocida, pero se puede decir que el número es bastante grande como resultado de un problema ambiental significativo. Entre otras industrias que utilizan colorantes, se dice que la industria textil utiliza la mayor cantidad de colorante aproximadamente 10,000 toneladas por año en todo el mundo.

1.1.1 Colorante amarillo No. 5

Es un colorante amarillo sintético también llamado tartrazina, de fórmula molecular $C_{16}H_9N_4Na_3O_9S_2$, su masa molar es 534.4 g/mol, pertenece al grupo de los colorantes azoicos, caracterizados por la presencia del grupo azo ($-N=N-$) unido a anillos aromáticos.

Se presenta en forma de polvo brillante de color amarillo-limón, es inodoro, higroscópico, estable en ácidos, soluble en agua y poco soluble en etanol. Su uso está ampliamente difundido en la industria alimentaria, en los productos que presentan una coloración amarilla. En la Tabla 1 se muestran la estructura química y algunas propiedades físicas la tartrazina o amarillo No. 5. Que se utilizará en esta investigación.

Tabla 1. Propiedades del colorante amarillo No. 5 (Goscianska & Ciesielczyk, 2019).

Tartrazina	
Estructura	
Otros nombres	Amarillo 5, Amarillo ácido 23,
	Amarillo 4 en comida, C.I. 19140
Nombre corto	E102
Tamaño de molécula	1.68 nm x 0.848 nm
Color	Amarillo limón

1.1.2 Tratamientos para la remoción de colorantes sintéticos en aguas residuales

Se han utilizado diversas tecnologías de tratamiento de aguas residuales fisicoquímicas y biológicas, como sistemas de biorremediación aeróbica o anaeróbica, proceso oxidativo, reactivo de Fenton, ozonización, destrucción electroquímica, coagulación electrocinética, intercambio de iones, irradiación, filtración de membrana, decoloración por hongos de la pudrición blanca y adsorción, para la eliminación de colorantes, metales pesados, materia orgánica e inorgánica que contamina el agua en efluentes industriales, ya que cada método tiene sus propias ventajas y desventajas (Wong et al., 2020) (Karam et al., 2020).

Entre todas estas tecnologías, la adsorción es uno de los métodos más efectivos para el tratamiento de aguas que contienen contaminantes (colorantes, metales pesados, gases y microorganismos) (Wong et al., 2020) (Karam et al., 2020) (Ghaedi & Vafaei, 2017).

1.1.3 Adsorción

La adsorción es un proceso en el que los contaminantes presentes en solución se transfieren hacia la superficie de un adsorbente sólido, que se usa ampliamente en el control de la contaminación ambiental, como el tratamiento de aguas residuales. Para el uso de un sistema de adsorción, es esencial estudiar su cinética, que se puede utilizar para predecir la tasa de remoción de contaminantes, establecer la dependencia del tiempo del adsorbato residual en solución y revelar los mecanismos involucrados en el proceso de adsorción (Guo & Wang, 2019).

La adsorción es la técnica más empleada en la eliminación de estos compuestos (colorantes, metales pesados, materia orgánica, inorgánica y entre otros) debido a su facilidad de operación, capacidad para eliminar incluso niveles bajos de contaminantes y la ausencia de cualquier subproducto peligroso. La adsorción es un fenómeno interfacial complejo y tiene aplicaciones en procesos tecnológicos, biológicos y ambientales, como el tratamiento de aguas residuales, catálisis, bio-implantes, sistemas de enfriamiento, por nombrar algunos. Los sistemas de adsorción son relativamente fáciles y económicos de operar (Gopinath, Aravamudan, Muthukkumaran, & Ganesan, 2019).

Desde el punto de vista del diseño de procesos, la adsorción por lotes es una etapa experimental indispensable porque permite identificar la capacidad máxima del adsorbente, la velocidad de aproximación al equilibrio y la influencia de variables operativas como tiempo, temperatura, concentración inicial, pH y dosis de adsorbente. En sistemas con hidrogeles, estos factores adquieren especial importancia debido a que la red polimérica hidratada puede generar una trayectoria de adsorción en varias etapas: una etapa inicial dominada por la transferencia externa y la ocupación rápida de sitios accesibles, seguida por una etapa más lenta asociada con difusión intrapartícula, reorganización de cadenas poliméricas y aproximación al equilibrio. Esta complejidad explica por qué el ajuste cinético no debe interpretarse únicamente a partir de un valor elevado de R^2 , sino considerando

también la coherencia del tiempo de equilibrio y la utilidad del modelo para la predicción del proceso (Guo & Wang, 2019; Fang et al., 2020; Dutta et al., 2021).

1.1.4 Quitosano

El quitosano es un biopolímero ecológico que se obtiene de la desacetilación de la quitina. Se extrae del componente principal exoesqueleto de las conchas de los crustáceos, como los camarones y los cangrejos. El quitosano es biodegradable, biocompatible, no irritante y exhibe buenas propiedades de formación de película, alta resistencia mecánica y adhesión.

Es relativamente barato y se ha usado ampliamente como adsorbente para colorantes, metales de transición y especies orgánicas porque los grupos amino (NH_2) e hidroxilo (OH) en las cadenas de quitosano pueden servir como sitios de coordinación y reacción. Los grupos amino del quitosano pueden ser catiónizados y disponibles para la adsorción de colorantes aniónicos fuertemente por atracción electrostática en los medios ácidos, como se muestra en la molécula de la figura 2 (Sahnoun & Boutahala, 2018).

La combinación quitosano-PVA resulta especialmente adecuada para la preparación de perlas de hidrogel porque el PVA contribuye a mejorar la estabilidad mecánica, la flexibilidad y la resistencia estructural del material, mientras que el quitosano aporta grupos amino e hidroxilo capaces de interactuar con especies aniónicas. Esta sinergia permite obtener matrices adsorbentes que conservan la afinidad química del quitosano, pero con mayor integridad física durante tiempos prolongados de contacto en medio acuoso. En colorantes azoicos como la tartrazina, la presencia de grupos sulfonato favorece interacciones electrostáticas con grupos amino protonados del quitosano, particularmente bajo condiciones en las que dichos grupos se encuentran disponibles para la unión del contaminante (Crini & Badot, 2008; Sahnoun & Boutahala, 2018; Saheed & Suah, 2023; Li et al., 2025).

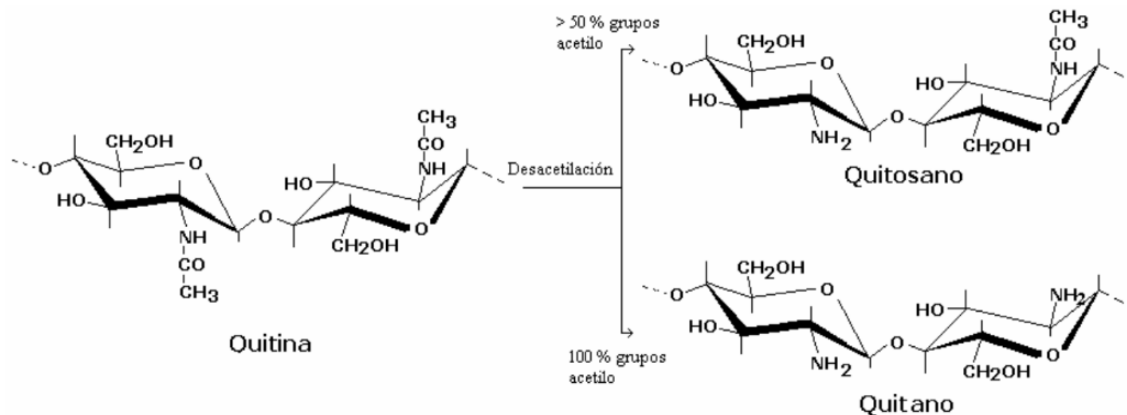


Figura 2. Estructura química de la quitina, quitosano y quitano (Sahnoun & Boutahala, 2018).

1.1.5 Cinéticas de adsorción

Explica la rapidez de la reacción química que se produce en el proceso de adsorción entre un material (adsorbente) y un contaminante (adsorbato), proporcionando algunos factores posibles que pueden afectar la velocidad de reacción (Crini & Badot, 2008) y que puede ser expresada con la ecuación (1.1):



Donde a es el adsorbato en la solución; n son los sitios vacíos del adsorbente; A es el coeficiente estequiométrico de la reacción (Ho & McKay, 1999). Se han aplicado varios tipos de modelos cinéticos para describir la cinética de adsorción, por ejemplo, el modelo de pseudo primer orden (PPO), modelo de pseudo segundo orden (PSO), y el modelo de Elovich (ver tabla 2), los cuales se usan con mayor frecuencia. A medida que han avanzado los estudios cinéticos, la mayoría de estos han descrito que la cinética de adsorción empleando los modelos PPO y PSO en paralelo, podrían describir mejor los datos cinéticos experimentales mediante el análisis del coeficiente de correlación (R^2) u otros parámetros estadísticos (Guo & Wang, 2019).

Tabla 2. Modelos cinéticos y ecuaciones de los modelos.

No.	Modelo	Ecuación
1.2	Lagergren	$q_t = q_e(1 - e^{-k_1 t})$
1.3	Ho and Mckay	$q_t = \frac{k_2 q_e^2 t}{1 + k_2 q_e t}$
1.4	Elovich	$q_t = \alpha \ln(\alpha \beta) + \beta \ln(t)$

Donde:

q_e : Cantidad del colorante removido en el equilibrio por el material sorbente (mg/g)

q_t : Cantidad del colorante removido en el tiempo t por el material sorbente (mg/g)

k_1 : Constante de velocidad en el equilibrio de una sorción de primer orden (1/h)

k_2 : Constante de velocidad en el equilibrio de una adsorción de segundo orden (g/mg h)

α : Constante de adsorción inicial

β : Constante de desorción durante el experimento

t : Tiempo de contacto (h)

La expresión matemática 1.2 corresponde al modelo cinético de pseudo-primer orden o modelo de Lagergren, se basa en la suposición de que a cada ión se le asigna un sitio de sorción del material sorbente, lo que refiere un proceso de fisisorción, debido a la asociación del adsorbato por medio de las fuerzas de van der Waals, lo cual en términos de velocidad de reacción se expresa como (Kumar, Tripathi, & Shahi, 2009). El modelo cinético de pseudo-segundo orden (ecuación 1.3) fue desarrollado por Ho-McKay por lo que también es conocido como modelo de Ho; en el se supone que el sorbato se sorbe en dos sitios activos del sorbente. En este caso, la ecuación de velocidad de la cinética de sorción se expresa como (Gerente, Lee, Cloirec, & McKay, 2007). Por otro lado, la ecuación 1.4 corresponde al modelo de Elovich el cual es de aplicación general a la cinética de quimisorción,

este modelo se ha utilizado satisfactoriamente en sistemas en los que la superficie de adsorción es heterogénea, La ecuación cinética de Elovich es (Vargas et al., 2012):

No obstante, los modelos de Lagergren, Ho-McKay y Elovich deben utilizarse con cautela cuando se aplican a hidrogeles poliméricos. El modelo de pseudo-primer orden se asocia frecuentemente con procesos controlados por ocupación de sitios disponibles; el modelo de pseudo-segundo orden suele emplearse para describir sistemas con fuerte interacción adsorbato-adsorbente; y Elovich es útil en superficies heterogéneas. Sin embargo, estos modelos pueden ofrecer buenos coeficientes de correlación aun cuando el mecanismo real involucra simultáneamente difusión externa, difusión interna, hinchamiento del polímero y adsorción en sitios con distinta energía. Por ello, la interpretación rigurosa requiere comparar no sólo el ajuste estadístico, sino también la capacidad de cada modelo para estimar q_e y el tiempo de equilibrio de forma operativamente útil (Ho & McKay, 1999; Guo & Wang, 2019; Hubbe et al., 2019; Fang et al., 2020).

1.2 Red neuronal artificial (RNA)

Las redes neuronales artificiales se pueden definir de diferentes formas: existen definiciones cortas, genéricas y las que intentan explicar más detalladamente lo que significa una RNA o computación neuronal.

La RNA fue introducida por McCulloch y Pitts, ésta actúa como un modelo de "caja negra" derivado de un concepto simplificado del cerebro humano. Ha sido ampliamente aceptada como un enfoque para la predicción, los sistemas de control, la clasificación, la optimización y la toma de decisiones en varios campos. Las RNA están inspiradas en el procesamiento nervioso biológico, puede emplearse para resolver y modelar muchos de los sistemas de ingeniería complejos debido a su simplicidad, robustez, confiabilidad y no linealidad. El modelo RNA puede aprender a partir de datos experimentales para resolver las complejas relaciones no lineales

y funcionales multidimensionales sin ningún supuesto previo sobre su naturaleza (Ghaedi & Vafaei, 2017).

Las redes neuronales artificiales simulan principios del funcionamiento del cerebro humano y realizan tareas de aprendizaje y predicción. La arquitectura de la RNA consta de una capa de entrada, una o más capas ocultas y una capa de salida (El Bardiji, Ziat, Naji, & Saidi, 2020). Cada capa de la red consiste en un conjunto de elementos de procesamiento interconectados llamados neuronas o también llamados nodos artificiales en cada capa todos ellos interconectados mediante un conjunto de lazos sinápticos (Haykin, 2009). Estas neuronas interactúan entre sí con la ayuda del peso. Cada neurona está conectada a todas las neuronas en la siguiente capa. Los datos se presentan a la red neuronal en la capa de entrada. La salida de la red neuronal está dada por la capa de salida para los datos de entrada dados. Las capas ocultas permiten que estas redes calculen relaciones complicadas entre entradas y salidas (Aghav, Kumar, & Mukherjee, 2011).

La arquitectura o topología de RNA es una característica importante que representa la forma en que se conectan las entradas, las neuronas artificiales y las salidas como se muestra en la figura 3(b). Estas conexiones se conocen como pesos y sesgos, que representan los parámetros principales del modelo RNA. En el marco de una red neuronal artificial, las variables de entrada se agrupan en la capa de entrada. A su vez, las neuronas artificiales (ver figura 3(a)) se agrupan en varias capas, como las capas ocultas y de salida. Para resolver problemas de regresión múltiple (aproximaciones de funciones), la topología más empleada es la arquitectura de alimentación de múltiples capas, designada también como perceptrón de múltiples capas (Cojocar, Samoilă, & Pascariu, 2019). El perceptrón multicapa (PM) es la arquitectura más convencional de la RNA (Li et al., 2019).

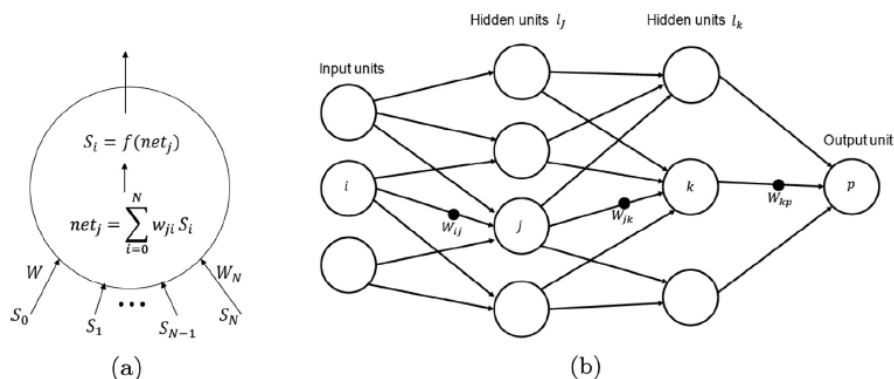


Figura 3. (a) Estructura de una neurona y (b) una Red Neuronal Artificial (Pacheco-Sánchez, Alejo, Cruz-Reyes, & Álvarez-Ramírez, 2019).

Las RNA han sido un tema de estudio que ha beneficiado muchas áreas debido a su gran capacidad que han ido adquiriendo para dar solución a un problema o varios, en el área ambiental no ha sido la excepción pues se han implementado satisfactoriamente, todo esto con la finalidad de tener un mayor conocimiento sobre el comportamiento del fenómeno de adsorción con los diferentes contaminantes y adsorbentes a continuación se ven varios ejemplos de la implantación de RNA.

Ghaedi y Vafaei (Ghaedi & Vafaei, 2017), hicieron una revisión de aplicaciones de métodos de RNA como predicción, donde encontraron cuatro categorías en los reportes de adsorción: redes neuronales de alimentación de múltiples capas (RNAMC), sistemas de inferencia neuronal adaptativa (SINA), máquina de vectores de soporte (MVS), y el híbrido con algoritmos genéticos (AG) y optimización de enjambre de partículas (OEP).

Dentro de los estudios realizados en la síntesis de nuevos materiales para la remoción de colorantes, se encuentra el realizado por Tanzifi y colaboradores (Tanzifi et al., 2017) en el que sintetizaron un adsorbente de polianilina para adsorber el naranja de metilo de una solución acuosa. Analizaron los efectos de varios parámetros, como el pH, la temperatura, el tiempo de adsorción, la concentración inicial y la dosis de adsorbente, los cuales optimizaron por medio de un modelo de RNA. La red neuronal de alimentación avanzada de varias capas con

cinco entradas y una salida la entrenaron con ocho neuronas en la capa oculta. Una comparación de los datos experimentales con la eficiencia de adsorción del colorante predicho por el modelo de RNA les mostró que este modelo puede estimar el comportamiento del proceso de adsorción del colorante naranja de metilo en el nano-adsorbente de polianilina en diferentes condiciones.

Por otro lado, Gadekar y Ahammed utilizaron el método de superficie de respuesta (MSR) con una RNA para optimizar el modelo de remoción por adsorción de un colorante disperso en un efluente líquido de un tratamiento de agua residual y así conocer las condiciones óptimas para tal fin. El MSR fue utilizado en una primera etapa para evaluar el proceso considerando tres parámetros operativos controlables, la cantidad del agua residual tratada, el pH inicial y la concentración de colorante. En una segunda etapa, los resultados experimentales de los datos de diseño obtenido en el MSR los usaron para entrenar la RNA junto con un parámetro no controlable, el pH final. Una vez entrenada la RNA, se observó una remoción del colorante de $52.6 \pm 2.0\%$ obtenida experimentalmente en condiciones optimizadas (pH inicial 3.0, dosis de adsorbente 30 g/L y concentración de colorante 75 mg/L) fue comparable a 52.0% y 52.2% pronosticado por MSR y el método de superficie de respuesta con redes neuronales artificiales (MSR-RNA), respectivamente. Este estudio demostró que es posible optimizar/predecir el proceso de remoción de colorante utilizando el enfoque de MSR-RNA (Gadekar & Ahammed, 2019).

Cojocarú y colaboradores desarrollaron un modelo de RNA para explorar ampliamente los efectos de los factores en el rendimiento de adsorción. Demostraron que el acoplamiento del modelo RNA con un algoritmo genético proporcionó condiciones óptimas de adsorción (Cojocarú et al., 2019).

Gopinath y colaboradores propusieron una RNA rápida, confiable y eficiente para el modelo de difusión de superficie homogénea (MDSH), demostrada para numerosas simulaciones y estimaciones para superficies difusas individuales, sistemas binarios y ternarios involucrando moléculas de colorantes y farmacéuticas, también

desarrollaron una RNA para estimar los parámetros de adsorción haciendo su función en reversa para simular la cinética de múltiples componentes y rastrear el aumento en el porcentaje de eliminación o remoción de los adsorbatos con el tiempo en diferentes condiciones del proceso (Gopinath et al., 2019).

También se ha evaluado la aplicabilidad de las nanopartículas de Ho-CaWO₄ preparadas utilizando el método hidrotérmico para la eliminación del azul de metileno (AM) de la solución acuosa mediante el proceso de adsorción. La metodología de superficie de respuesta (MSR) y las técnicas de modelado de redes neuronales artificiales (RNA) fueron aplicadas para modelar el proceso y también se examinó su rendimiento y las capacidades predictivas de la respuesta (eficiencia de eliminación). El proceso de adsorción lo optimizaron utilizando la MSR y determinaron las condiciones óptimas. El proceso también se modeló utilizando la isoterma de adsorción y los modelos cinéticos. El modelo de RNA-MSR les mostró una predicción adecuada de la respuesta (Igwegbe et al., 2019).

Karam y colaboradores (Karam et al., 2020), simularon el proceso de adsorción en una red neuronal artificial (RNA) y obtuvieron que fueron compatibles los resultados del análisis de regresión derivado de la optimización de la MSR, mostrando que la mayor capacidad para nanopartículas de hierro cero-valentes (nZVI), carbón activado (AC) y nanopartículas de hierro cero-valentes por síntesis en química verde (GT-nZVI) en la eliminación de color de aguas residuales de industrias textiles. A pH 5, tiempo de contacto 50 minutos y velocidad de agitación 150 rpm, las nZVI mostraron una buena eficacia de eliminación de color de aproximadamente 71% y 99% para concentraciones de color iniciales de 350 y 50 mg/L Pt/Co unidad de color, respectivamente, mientras que para el AC se lograron capacidades de eliminación de color ligeramente mayor de aproximadamente 72% y 100% a pH 8, con tiempo de contacto 70 minutos y velocidad de agitación 250 rpm. Finalmente, registraron la mayor capacidad de eliminación de color de aproximadamente 85% y 100% para las nanopartículas GT-nZVI a pH 7, tiempo de contacto 40 minutos y velocidad de agitación 150 rpm. Su trabajo muestra la capacidad mejorada de eliminación de

color del GT-nZVI como un potencial material de decoloración de aguas residuales textiles, abriendo el camino para muchas aplicaciones industriales y ambientales.

Wong y colaboradores (Wong et al., 2020), compararon las eficacias de la red neuronal artificial (RNA), el sistema adaptativo de inferencia neuro difusa (SAIND) y los modelos de regresión lineal múltiple (MRL) y evaluaron su capacidad para estimar la eficiencia de adsorción de un biocarbón de cáscara de rambután (*Nephelium lappaceum*) para la eliminación de iones Cu (II) basado en 480 juegos experimentales obtenidos en un estudio de laboratorio por lotes. Investigaron los efectos de parámetros operativos tales como el tiempo de contacto, la temperatura de operación, la dosis de biocarbón y la concentración inicial de iones Cu (II) en la eliminación de iones Cu (II). Compararon estadísticamente once algoritmos de entrenamiento diferentes en RNA y 8 funciones pertenencia diferentes en SAIND y se evaluaron en términos de errores de estimación, que fueron el error cuadrático medio (ECM), error absoluto medio (EAM) y la precisión. Estudiaron los efectos del número de neuronas ocultas en el modelo RNA y la combinación de conjuntos difusos en SAIND. En este estudio, se encontró que el modelo SAIND con la función de pertenencia gaussiana y la combinación de conjuntos difusos es el mejor método, con una precisión del 90.24% y 87.06% para el conjunto de datos de entrenamiento y prueba, respectivamente. La contribución de su estudio es que las técnicas de modelado RNA, SAIND y MRL las utilizaron por primera vez para estudiar la adsorción de iones Cu (II) a partir de soluciones acuosas usando biochar de cáscara de rambután.

En las investigaciones mencionadas aluden a las RNA para predecir el comportamiento del proceso de adsorción, se ajustan mejor a los datos experimentales, superando a los modelos y métodos estadísticos. Las RNA se han aplicado ampliamente para la predicción de la adsorción de colorantes durante la última década. Estos modelos de RNA están obteniendo popularidad como enfoques, que pueden emplearse con éxito para la adsorción de colorantes con una precisión aceptable (Ghaedi & Vafaei, 2017).

En términos metodológicos, una RNA de tipo perceptrón multicapa se justifica porque se comporta como un aproximador universal de funciones y puede aprender relaciones no lineales entre variables experimentales sin requerir la formulación explícita de una ecuación mecánica cerrada. En procesos ambientales, esta capacidad es útil cuando los fenómenos son multivariantes y parcialmente acoplados. Aun así, el uso de RNA debe acompañarse de criterios de validación rigurosos, tales como separación de datos de entrenamiento, validación y prueba, normalización de variables, evaluación mediante métricas estadísticas y comparación con modelos de referencia. Sin estos elementos, una RNA puede presentar sobreajuste o una aparente precisión que no necesariamente se traduce en capacidad de generalización (Haykin, 2009; Ekman, 2022; Kim et al., 2025; Vasiliauskaite & Antulov-Fantulin, 2024).

2 METODOLOGÍA

Para el desarrollo del modelo se utilizaron datos experimentales que fueron obtenidos en el Laboratorio de Investigación de Ingeniería Ambiental (LIIA) del Instituto Tecnológico de Toluca.

Metodológicamente, el trabajo se estructuró como un estudio de modelado predictivo basado en datos experimentales de adsorción por lotes. Esto significa que la tesis no parte de simulaciones hipotéticas, sino de mediciones obtenidas bajo condiciones controladas de laboratorio, las cuales posteriormente se transformaron en una base de datos apta para el entrenamiento de una red neuronal artificial. La lógica general del procedimiento fue: obtener datos cinéticos confiables, organizar las variables relevantes del proceso, seleccionar entradas y salidas del modelo, entrenar diferentes arquitecturas neuronales, validar su desempeño y comparar sus predicciones con modelos cinéticos tradicionales. Esta secuencia es coherente con los enfoques actuales de ingeniería ambiental basada en datos, en los que los modelos de aprendizaje automático se emplean para reconocer relaciones no lineales cuando los modelos convencionales resultan insuficientes para describir

todos los fenómenos involucrados (Ghaedi & Vafaei, 2017; Pauletto et al., 2020; Domínguez Beltrán et al., 2025).

El uso de datos obtenidos en el LIIA permitió mantener trazabilidad experimental y coherencia entre la fase fisicoquímica y la fase computacional. En estudios de adsorción, esta trazabilidad es indispensable porque la capacidad de adsorción no depende únicamente de la concentración del contaminante, sino también de la masa y morfología del adsorbente, el volumen de solución, el tiempo de contacto, la temperatura, la agitación y la estabilidad estructural del material. Por ello, antes de entrenar la RNA se consideró necesario conservar la correspondencia entre cada medición experimental y la condición bajo la cual fue generada, evitando mezclar observaciones que pertenecen a tamaños de perla o temperaturas distintas sin una identificación clara del contexto experimental (Rápó & Tonk, 2021; Dutta et al., 2021).

Se obtuvo la base de datos que contiene los datos experimentales del proceso de adsorción discontinuo del colorante amarillo 5 en solución acuosa usando como material adsorbente perlas de hidrogel de quitosano-alcohol polivinílico, desarrollado por el grupo de investigación responsables del proyecto de donde emana este trabajo de tesis.

La base de datos constituye el eje del modelado, por lo que su organización debe responder a una lógica experimental y estadística. En este caso, las observaciones se derivan de curvas de adsorción por lotes generadas a tres temperaturas y tres tamaños de perla, lo cual permite capturar variaciones asociadas tanto con la energía térmica del sistema como con la relación superficie-volumen del adsorbente. La construcción de la base con 297 puntos experimentales proporciona un soporte razonable para entrenar un modelo supervisado, siempre que los datos sean tratados con procedimientos consistentes de normalización y partición. La normalización es relevante porque evita que una variable con mayor magnitud numérica domine el proceso de aprendizaje y mejora la estabilidad del descenso de

gradiente durante el entrenamiento (Kim et al., 2025; Domínguez Beltrán et al., 2025).

El experimento de adsorción consistió en colocar una masa de hidrogel en formas de perlas (adsorbente) de tres tamaños (2, 2.6 y 3.1 mm) por separado en viales de vidrio, y ponerlas en contacto con un volumen fijo de solución de colorante con concentración C_0 , velocidad de agitación constante, a temperaturas de 10°C, 30°C y 50°C y pH. La capacidad de adsorción (q) se expresa en mg de colorante adsorbido por gramos de adsorbente (mg/g). Se tomaron alícuotas de la solución de colorante tomadas en los tiempos: 0.5, 1, 4, 8, 16, 24, 32, 40, 48, 63 y 72 horas. Los resultados de estos experimentos permitieron conformar una base de datos de gráficas o curvas de cinética de adsorción de hidrogel base quitosano independientes. En la etapa 2 se detallan los materiales y métodos utilizados para el proceso de transformación de cada experimento a curva de cinética de adsorción por RNA, las cuales son visualizadas en un sistema de coordenadas (vista en la sección correspondiente). Para la implementación de la RNA se utilizó un sistema operativo de software libre llamado CentOS 7, la RNA fue programada en C++, la que se puede ejecutar desde la terminal de CentOS 7, el código es susceptible de modificación, ya que está inspirado en ser de código abierto y un software libre, por ejemplo, se puede cambiar la función de activación, razón de aprendizaje, máximo o mínimo número de neuronas en cada capa, etc. Por otro lado, para graficar los resultados obtenidos se utilizó Gnuplot, esto con el fin de observar el comportamiento de los datos y poder predecir el fenómeno de adsorción en estudio. El sistema experimental corresponde a una adsorción discontinua o por lotes, en la cual una cantidad conocida de adsorbente se pone en contacto con un volumen determinado de solución de colorante y se monitorea la disminución de la concentración en la fase líquida con respecto al tiempo. La ventaja de este arreglo experimental es que permite construir curvas cinéticas completas y observar tres etapas características del proceso: una etapa inicial de adsorción rápida, asociada con la disponibilidad de sitios activos externos; una etapa intermedia en la que la velocidad disminuye por el avance de la difusión hacia el interior de la perla; y una

etapa final de aproximación al equilibrio, en la que los cambios de q_t se vuelven pequeños debido a la saturación progresiva de los sitios disponibles (Crini & Badot, 2008; Guo & Wang, 2019).

La capacidad de adsorción a cada tiempo se interpreta como la cantidad de tartrazina retenida por unidad de masa de hidrogel. En términos experimentales, esta variable se calcula a partir del balance de masa entre la concentración inicial y la concentración medida en cada alícuota, considerando el volumen de solución y la masa de adsorbente: $q_t = ((C_0 - C_t)V)/m$. Esta relación es central para el trabajo porque transforma una medición analítica de concentración en una variable de desempeño del adsorbente. Así, cada valor de q_t resume la respuesta del material bajo una condición específica de tiempo, temperatura y tamaño de perla, y constituye la información que posteriormente alimenta al modelo neuronal y a los modelos cinéticos tradicionales (Ho & McKay, 1999; Vargas et al., 2012; Domínguez Beltrán et al., 2025).

La selección de los tiempos de muestreo, desde los primeros puntos de contacto hasta las 72 h, permite representar adecuadamente la forma de la curva cinética. Los tiempos cortos son relevantes porque capturan la velocidad inicial de adsorción, mientras que los tiempos prolongados permiten identificar la tendencia hacia el equilibrio. Esta distribución temporal es importante para el entrenamiento de la RNA, ya que un conjunto de datos concentrado únicamente en la zona final de la curva no permitiría reconocer la etapa inicial rápida, mientras que un conjunto concentrado en los primeros tiempos impediría estimar con solidez la meseta de equilibrio. Por ello, la inclusión de puntos intermedios como 16, 24, 32, 40, 48 y 63 h fortalece la capacidad del modelo para aprender la transición entre adsorción rápida y estabilización del sistema (Guo & Wang, 2019; Fang et al., 2020).

En la figura 4 se presenta el diagrama de flujo que permite visualizar la construcción de la RNA que predice el comportamiento del proceso de adsorción del Ama5 así como la obtención del modelo general para calcular la capacidad de adsorción (q_e).

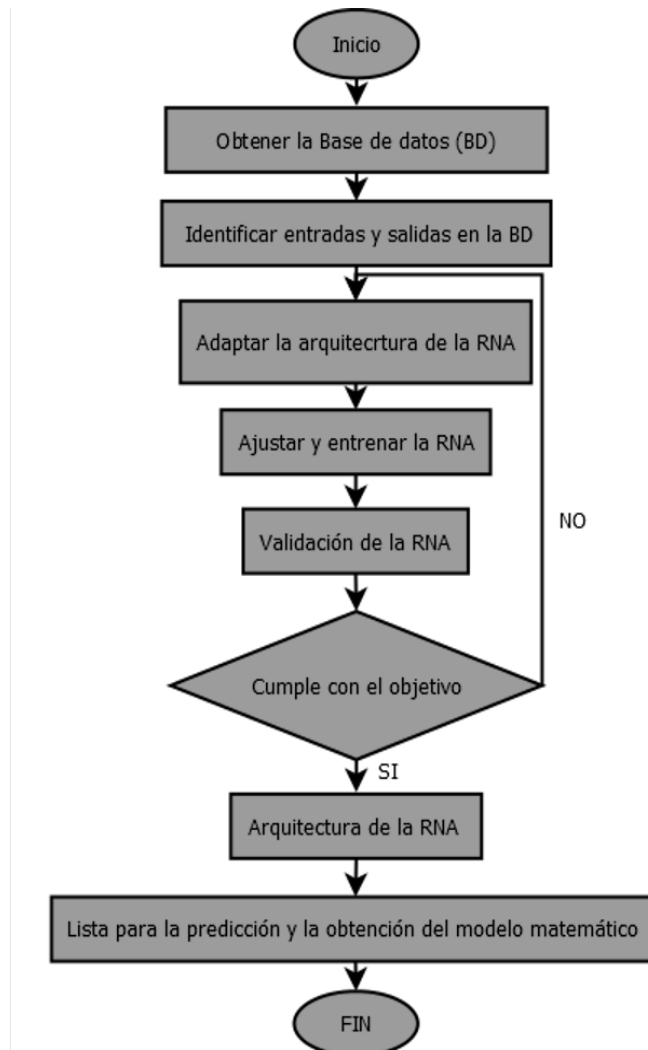


Figura 4. Diagrama de flujo de los procesos de formación de la RNA.

2.1 Base de datos (BD) de entrenamiento

Se utilizó para entrenar la RNA, la base de datos tiene un total de 297 muestras experimentales, con las características: concentración inicial de adsorbato: $C_0 = 300 \text{ mg/L}$, $V = 0.01 \text{ L}$, para tres tamaños diferentes de perlas (material adsorbente) que se denominaron como **chica, mediana y grande**, las demás características se muestran en la Tabla 2.1. donde también se ingresaron algunos ejemplos de datos experimentales para hacer pruebas de cinéticas de adsorción. Donde $t(h)$ es el

tiempo expresado en horas, T ($^{\circ}C$) es la temperatura en escala de Celsius, m (g) es masa del adsorbente en gramos, C_f (mg/L) es la concentración final y q_t (mg/g) es la capacidad de adsorción.

Cada registro de la base de datos representa una condición experimental individual y debe entenderse como una observación multivariable. La concentración inicial y el volumen se mantuvieron constantes para que la comparación entre tratamientos se concentrara principalmente en el efecto del tamaño de perla, la temperatura, la masa real de adsorbente y el tiempo de contacto. Esta decisión metodológica reduce la variabilidad no deseada y permite que el análisis se enfoque en los factores que se busca modelar. Al trabajar con tres tamaños de perla y tres temperaturas, la base de datos no solo describe una curva individual, sino una familia de curvas cinéticas bajo condiciones controladas.

La masa de adsorbente incluida en la base de datos es relevante porque pequeñas variaciones en m modifican directamente el valor de q_t . En materiales poliméricos hidratados, como los hidrogeles de quitosano-PVA, la masa puede estar asociada con el grado de hidratación y con la integridad de las perlas durante el ensayo. Por esta razón, registrar la masa permite conservar el balance de materia y evita que las diferencias de capacidad sean atribuidas únicamente al efecto de la temperatura o del tamaño de partícula. Esta consideración es importante en hidrogeles, donde la estructura tridimensional, la hinchazón y la disponibilidad de grupos funcionales influyen en la captura de colorantes aniónicos (Crini & Badot, 2008; Li et al., 2025).

Desde el punto de vista computacional, la base de datos fue preparada para que la RNA pudiera identificar patrones de variación entre t y q_t , y asociarlos con la capacidad de equilibrio. En este tipo de modelado, la calidad de la base de datos es tan importante como la arquitectura de la red. Una RNA entrenada con datos no normalizados, desordenados o inconsistentes puede producir ajustes aparentemente aceptables, pero con baja capacidad de generalización. Por ello, el preprocesamiento de los datos, la revisión de unidades, la consistencia de las variables y la separación entre entrenamiento, validación y prueba son pasos

metodológicos indispensables para evitar sobreajuste y mejorar la reproducibilidad del modelo (Haykin, 2009; Kim et al., 2025).

Tabla 3. Datos experimentales.

t (h)	T (°C)	m (g)	C _f (mg/L)	q _t (mg/g)
0.30	10	0.0319	280.89	181.53
0.30	30	0.0319	278.12	207.85
0.30	50	0.0315	273.60	253.96
1.00	10	0.0310	270.84	285.04
1.00	30	0.0316	262.00	364.40
1.00	50	0.0317	264.60	338.40
4.00	10	0.031	241.24	574.39
4.00	30	0.0318	239.88	572.90
4.00	50	0.0319	246.74	505.94
8.00	10	0.0315	228.00	692.64
8.00	30	0.0309	218.36	800.63
8.00	50	0.0317	240.16	572.03
16.00	10	0.0316	220.96	757.96
16.00	30	0.0315	210.33	862.63
16.00	50	0.0318	235.00	619.40
24.00	10	0.0317	220.64	758.63
24.00	30	0.0315	205.52	908.90
24.00	50	0.0321	225.68	701.60
32.00	10	0.0311	210.96	867.58
32.00	30	0.0315	203.40	929.29
32.00	50	0.0310	220.16	780.45
40.00	10	0.0314	207.16	895.97
40.00	30	0.0315	204.24	921.21
40.00	50	0.0317	214.60	816.37
48.00	10	0.0321	199.20	951.57
48.00	30	0.0309	203.68	944.59
48.00	50	0.0318	210.40	853.82
63.00	10	0.0319	198.04	968.56
63.00	30	0.0310	203.16	946.63
63.00	50	0.0317	203.21	925.25
72.00	10	0.0313	198.07	986.83
72.00	30	0.0312	202.44	947.55
72.00	50	0.0318	206.00	895.75

2.2 Identificación de entradas y salidas

De todas las variables que se establecieron para el entrenamiento de la RNA fue necesario identificar las variables de entrada y salida, esto con el objetivo de obtener una salida deseada ya que dependiendo la entrada que tenga la RNA va a depender la salida. Se hizo una revisión y aprendizaje de cuáles son las variables de entrada y salida de la RNA, lo cual permitió tomar la decisión de realizar una representación general de los datos como lo que se utiliza para ajustar los datos experimentales a los modelos cinéticos de adsorción de pseudo primer orden (PPO), pseudo segundo orden (PSO) y Elovich en la capa de entrada la red neuronal recibe el tiempo (h) y q_t (mg/g) en la capa de salida la RNA lo que proporciona una predicción de la capacidad de adsorción.

La variable de salida se definió como q_e porque representa el parámetro de mayor utilidad para comparar el desempeño del adsorbente. Mientras q_t describe el comportamiento transitorio, q_e permite estimar la capacidad máxima alcanzada bajo una condición experimental específica. En términos de diseño de procesos, q_e es útil para dimensionar dosis de adsorbente, estimar tiempos de operación y comparar materiales. Por ello, predecir q_{te} a partir de la evolución temporal de q_t permite que la RNA funcione como una herramienta de apoyo para anticipar el comportamiento del sistema sin depender exclusivamente de ensayos prolongados hasta 72 h (Ghaedi & Vafaei, 2017; Karam et al., 2020).

La decisión de utilizar t y q_t como entradas implica que el modelo aprende a partir de la forma de la curva experimental, más que de una lista amplia de variables fisicoquímicas. Esta delimitación es apropiada para la tesis porque los experimentos fueron realizados con concentración inicial y volumen constantes, y porque el objetivo principal fue comparar la capacidad predictiva de la RNA frente a modelos cinéticos. No obstante, debe reconocerse que, para una generalización más amplia, futuras versiones del modelo podrían incorporar explícitamente temperatura, diámetro de perla, pH, concentración inicial y dosis de adsorbente como entradas adicionales. Esta ampliación permitiría convertir el modelo en una herramienta

predictiva más general para diseño y optimización (Pauletto et al., 2020; Alardhi et al., 2023; da Costa et al., 2025).

La selección de entradas y salidas en una RNA no es un paso meramente operativo, sino una decisión conceptual que define qué relación física se pretende aprender. Al emplear el tiempo de contacto y q_t como variables de entrada, el modelo se orienta a reconocer la trayectoria cinética observada durante el experimento y a estimar la capacidad de equilibrio asociada con esa trayectoria. Esta estrategia es coherente con el propósito del trabajo, ya que no se busca sustituir la caracterización fisicoquímica completa del adsorbente, sino construir un modelo predictivo a partir de la información experimental disponible. En este sentido, la RNA se utiliza como un modelo de regresión no lineal entrenado con datos de cinética, comparable con los modelos de Lagergren, Ho-McKay y Elovich (Ghaedi & Vafaei, 2017; Pauletto et al., 2020; Domínguez Beltrán et al., 2025).

2.3 Adaptación de la arquitectura de la RNA

Las redes neuronales artificiales al ser sistemas dinámicos auto adaptativos tienen la capacidad de ajustar las neuronas que constituyen la red y realizar constantes cambios con base a las nuevas condiciones. Sin embargo, no existe una arquitectura de red neuronal artificial que sea aplicable a cualquier problemática, razón por la cual el diseño e implementación de una RNA está condicionada al propósito de la tarea y lo que el investigador desee mostrar.

En esta tesis, la adaptación de la arquitectura se abordó mediante la evaluación de configuraciones con distinto número de neuronas en las capas ocultas. Esta etapa fue necesaria porque una red con pocas neuronas puede ser incapaz de representar la curvatura de las trayectorias cinéticas, mientras que una red con demasiadas neuronas puede ajustar ruido o variaciones puntuales de los datos experimentales. Por ello, la selección de la arquitectura no debe interpretarse como una decisión arbitraria, sino como un proceso de búsqueda de equilibrio entre capacidad de representación y estabilidad predictiva. Esta consideración es fundamental en

problemas de adsorción, donde los datos suelen ser limitados y la respuesta experimental puede presentar dispersión asociada con medición, transferencia de masa y heterogeneidad del adsorbente (Haykin, 2009; Ekman, 2022).

La arquitectura de perceptrón multicapa fue pertinente porque la relación entre tiempo de contacto y capacidad de adsorción no es estrictamente lineal. En las primeras horas, el incremento de q_t puede ser pronunciado; posteriormente, la pendiente disminuye hasta aproximarse a una meseta. Esta forma sigmoideal o asintótica es difícil de representar con modelos lineales simples, pero puede ser aprendida por redes con funciones de activación no lineales. De este modo, la RNA no requiere imponer previamente una ecuación cinética única, sino que ajusta los pesos internos para reproducir el patrón observado en los datos (Li et al., 2019; Ghaedi & Vafaei, 2017).

2.4 Diseño de la RNA

Para el diseño de la RNA se consideraron dos variables de entrada $t(h)$ y q_t , donde el $t(h)$ representa los diferentes tiempos a lo que se llevó la experimentación y q_t la cantidad de adsorbato adsorbido a diferentes tiempos, se propusieron dos capas ocultas con un número variable de neuronas y solo se identificó una salida que representada por la cantidad de adsorbato en el equilibrio (q_e).

Operativamente, la red recibe los valores de entrada en la primera capa, los transforma en las capas ocultas mediante pesos, sesgos y funciones de activación, y finalmente genera una salida asociada con q_e . Cada conexión ponderada representa un parámetro ajustable del modelo; durante el entrenamiento, estos parámetros se modifican de manera iterativa para reducir la diferencia entre la salida calculada y el valor esperado. Esta estructura permite que la RNA represente interacciones no lineales entre variables, lo cual resulta útil en sistemas donde la adsorción puede involucrar simultáneamente atracción electrostática, difusión

intraparticular, hidratación del polímero y disponibilidad progresiva de sitios activos (Crini & Badot, 2008; Cojocar et al., 2019).

La utilización de dos capas ocultas ofrece una mayor flexibilidad para aproximar funciones complejas que una red con una sola capa oculta, especialmente cuando la respuesta experimental presenta cambios de pendiente. Sin embargo, esta ventaja debe evaluarse con precaución: una arquitectura más compleja no necesariamente produce un modelo más útil si no mejora la predicción frente a datos no utilizados durante el entrenamiento. Por esta razón, el diseño de la RNA se complementó con la comparación frente a modelos cinéticos tradicionales y con el análisis de R^2 como indicador de ajuste. La finalidad no fue construir la red más grande, sino identificar una configuración capaz de reproducir adecuadamente la evolución de q_t y estimar q_e de forma práctica (Haykin, 2009; Domínguez Beltrán et al., 2025).

La elección de dos capas ocultas responde a la necesidad de representar una relación no lineal suficientemente flexible entre las variables de entrada y la salida. En un perceptrón multicapa, las capas ocultas transforman las entradas mediante funciones de activación y permiten que la red capture curvaturas, mesetas y transiciones propias de los perfiles cinéticos de adsorción. El empleo de funciones sigmoides resulta congruente con procesos que presentan crecimiento inicial y posterior aproximación asintótica al equilibrio, aunque su desempeño depende de la escala de los datos, la tasa de aprendizaje y el número de neuronas. Por ello, la arquitectura debe evaluarse empíricamente y contrastarse con modelos cinéticos convencionales para evitar conclusiones basadas únicamente en la complejidad del modelo (Haykin, 2009; Ekman, 2022; Li et al., 2019).

Después de saber las variables que afectan el proceso de adsorción y lo que se quiere predecir se planteó una arquitectura para la RNA como se muestra en la Figura 6.

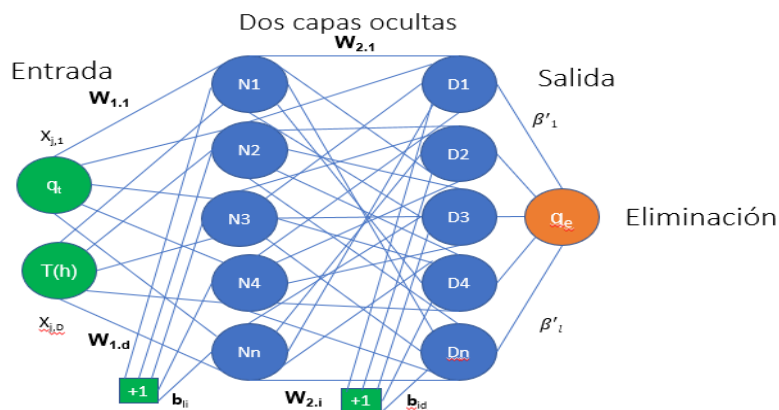


Figura 6. Arquitectura de la red.

2.5 Entrenamiento y validación de la RNA

La formación de la RNA consistió en encontrar una arquitectura la cual ayude a realizar una buena predicción sobre el proceso de adsorción, para después pasar a la fase de extraer el modelo matemático que represente dicho proceso fisicoquímico de separación en estudio, aunque se pretende que el modelo sea general, se tomaron como base el número de capas, neuronas y los pesos, que después de las pruebas de validación presentaron las mejores condiciones operativas.

Durante el entrenamiento, los datos se procesan en ciclos iterativos. En una primera fase, denominada propagación hacia adelante, las entradas atraviesan la red y generan una predicción. Posteriormente, el error entre la predicción y el valor de referencia se propaga hacia atrás mediante el algoritmo de retropropagación, ajustando los pesos en función de la contribución de cada conexión al error total. Este procedimiento se repite hasta que el error disminuye o hasta completar el número establecido de iteraciones. La tasa de aprendizaje controla la magnitud de los cambios en los pesos; si es demasiado alta, el modelo puede oscilar y no converger, mientras que si es demasiado baja, el aprendizaje puede ser excesivamente lento (Haykin, 2009; Ekman, 2022; Ngwenyama & Gitau, 2024).

La validación del modelo se entiende como una etapa distinta del entrenamiento. El entrenamiento permite ajustar los pesos, pero la validación permite valorar si la red

conserva capacidad predictiva más allá de los datos usados para aprender. En el artículo base se reporta una partición de los datos en subconjuntos de entrenamiento, validación y prueba, lo cual fortalece la evaluación del desempeño porque reduce el riesgo de interpretar un sobreajuste como si fuera una predicción real. Además, la normalización de las variables en el intervalo [0,1] favorece la convergencia numérica y evita que las diferencias de escala entre variables afecten el aprendizaje (Kim et al., 2025; Domínguez Beltrán et al., 2025).

La etapa de entrenamiento y validación es crítica para distinguir entre ajuste y predicción. El algoritmo de retropropagación utiliza la regla de la cadena para calcular gradientes y ajustar los pesos de la red con el fin de reducir el error entre la salida estimada y la respuesta experimental. No obstante, un modelo con demasiados parámetros puede aprender ruido experimental si no se controla adecuadamente. Por esta razón, la comparación con modelos cinéticos tradicionales cumple una doble función: proporcionar una línea base de desempeño y verificar si la RNA ofrece una ventaja práctica real en términos de estimación de q_e y del tiempo de equilibrio (Ekman, 2022; Ngwenyama & Gitau, 2024; Domínguez Beltrán et al., 2025).

Para obtener la mejor arquitectura se validaron los resultados realizando una comparación con los modelos cinéticos tradicionales: Lagergren, Ho-McKay y Elovich.

La comparación con Lagergren, Ho-McKay y Elovich fue indispensable porque estos modelos constituyen una referencia ampliamente aceptada en estudios de cinética de adsorción. Lagergren permite representar procesos donde la velocidad depende de los sitios no ocupados; Ho-McKay suele emplearse cuando el ajuste sugiere una cinética de pseudo-segundo orden; y Elovich se utiliza con frecuencia en superficies heterogéneas. Sin embargo, estos modelos no siempre ofrecen simultáneamente una predicción precisa de q_e y del tiempo de equilibrio, y sus parámetros pueden ser sensibles a la forma de ajuste. Por ello, la RNA se evaluó no solo por su R^2 , sino también por su capacidad de proporcionar una estimación útil del tiempo requerido

para alcanzar la meseta de adsorción (Ho & McKay, 1999; Guo & Wang, 2019; Fang et al., 2020).

En este trabajo, el coeficiente de determinación R^2 se utilizó como métrica principal porque permite cuantificar la proporción de variabilidad explicada por el modelo. No obstante, su interpretación debe realizarse junto con el sentido físico de la predicción. Un modelo puede alcanzar un R^2 elevado y, al mismo tiempo, estimar un tiempo de equilibrio poco práctico o mayor al intervalo experimental. Por esta razón, la discusión de resultados se centró en el equilibrio entre ajuste estadístico, capacidad de adsorción estimada y utilidad operativa del tiempo predicho. Esta lectura evita reducir la comparación de modelos a un único indicador numérico (Hubbe et al., 2019; Fang et al., 2020).

2.6 Modelo de la RNA

El modelo matemático a obtener por medio de la RNA que represente el proceso de adsorción en un sistema por lotes para la remoción de un colorante disuelto en agua y por el cual se puede obtener el punto de equilibrio de q_e , así como el tiempo en el que no hay variación de la cantidad adsorbida por el adsorbato.

El modelo neuronal puede entenderse como una función matemática parametrizada por pesos y sesgos. A diferencia de una ecuación cinética clásica, cuyos parámetros suelen tener una interpretación directa dentro de un mecanismo supuesto, los pesos de una RNA representan transformaciones internas que permiten aproximar la relación entre entradas y salida. En este caso, el interés principal no radica en interpretar cada peso de forma aislada, sino en utilizar el conjunto de pesos entrenados para estimar q_e y localizar el punto en el que la curva deja de presentar variaciones significativas. Esta aproximación es congruente con el uso de RNA como modelos de caja negra o modelos sustitutos en procesos ambientales complejos (Ghaedi & Vafaei, 2017; Gopinath et al., 2019).

La derivación mediante regla de la cadena es el fundamento matemático que permite ajustar los pesos del perceptrón multicapa. Cada derivada parcial expresa cómo cambia la salida de la red ante una modificación en un peso específico. Al multiplicar las derivadas de las capas sucesivas, el algoritmo determina la dirección en la que deben corregirse los pesos para reducir el error. Por tanto, las expresiones incluidas en esta sección no son únicamente un desarrollo algebraico, sino la base del aprendizaje supervisado utilizado para convertir los datos experimentales en un modelo predictivo (Haykin, 2009; Ekman, 2022).

La identificación del equilibrio mediante la RNA se realizó a partir de la estabilización de la capacidad de adsorción predicha. En términos prácticos, el equilibrio se interpreta como el intervalo en el que el incremento de q_t con el tiempo ya no representa un cambio significativo para el sistema. Esta interpretación es relevante porque, desde el punto de vista experimental, continuar el ensayo después de la estabilización consume tiempo y recursos sin aportar una mejora proporcional en la información obtenida. Por ello, una de las aportaciones de la tesis es mostrar que el modelo neuronal puede sugerir rangos temporales de equilibrio más cortos que los modelos tradicionales, orientando mejor la planeación experimental (Domínguez Beltrán et al., 2025).

Para hallar el punto de equilibrio se realizaron las derivadas parciales respecto a z_j tomando como referencia las siguientes ecuaciones:

$$\begin{aligned} a_1 &= f(p_1) = (1 + e^{-p_1})^{-1}, \\ y_m &= h(r_m) = (1 + e^{-r_m})^{-1} \\ z_j &= g(s_j) = (1 + e^{-s_j})^{-1} \end{aligned}$$

Obteniendo las derivadas parciales para cada variable, la expresión final, juntando todas las derivadas parciales, luce de la siguiente forma:

$$q_e = \left(\frac{\partial z_j}{\partial s_j}\right) \left(\frac{\partial s_j}{\partial y_m}\right) \left(\frac{\partial y_m}{\partial r_m}\right) \left(\frac{\partial r_m}{\partial a_l}\right) \left(\frac{\partial a_l}{\partial p_l}\right) \left(\frac{\partial p_l}{\partial X_n}\right)$$

A continuación, se muestra el desarrollo del cálculo de las derivadas:

$$\left(\frac{\partial z_j}{\partial s_j}\right) = \frac{\partial(1 + e^{-s_j})^{-1}}{\partial s_j} = -1(1 + e^{-s_j})^{-2} \cdot (e^{-s_j}) = -z_j^2 j(e^{-s_j}) = -z_j^2 j\left(\frac{1}{z_j} - 1\right) = -z_j + z_j^2$$

$$\frac{\partial s_j}{\partial y_m} = \frac{\partial(\sum_m y_m U_{mj})}{\partial y_m} = U_{mj}$$

$$\begin{aligned} \frac{\partial y_m}{\partial r_m} &= \frac{\partial(1 + e^{-r_m})^{-1}}{\partial r_m} = -1(1 + e^{-r_m})^{-2} \cdot (e^{-r_m}) = -y_m^2 (e^{-r_m}) = -y_m^2 (y_m^{-1} - 1) \\ &= -y_m^2 \left(\frac{1}{y_m} - 1\right) = -y_m + y_m^2 \end{aligned}$$

$$\frac{\partial r_m}{\partial a_l} = \frac{\partial(\sum_l a_l W_{lm})}{\partial a_l} = W_{lm}$$

$$\begin{aligned} \frac{\partial a_l}{\partial p_l} &= \frac{\partial(1 + e^{-p_l})^{-1}}{\partial p_l} = -1(1 + e^{-p_l})^{-2} \cdot (e^{-p_l}) = -a_l^2 (e^{-p_l}) = -a_l^2 (a_l^{-1} - 1) = -a_m^2 \left(\frac{1}{a_l} - 1\right) \\ &= -a_l + a_l^2 \end{aligned}$$

$$\frac{\partial p_l}{\partial X_n} = \frac{\partial(\sum_n X_n V_{nl})}{\partial X_n} = V_{nl}$$

Finalmente, al aplicar todos los resultados de cada una de las derivadas parciales se obtiene una expresión como la siguiente:

$$qe = z_j^2 (z_j^{-1} - 1) \cdot U_{mj} \cdot y_m^2 (y_j^{-1} - 1) \cdot W_{ml} \cdot a_i (-1 + a_i) \cdot V_{nl}$$

Simplificando la expresión anterior:

$$qe = z_j (-1 + z_j) \cdot U_{mj} \cdot Y_m (-1 + Y_m) \cdot W_{ml} \cdot a_i (-1 + a_i) \cdot V_{nl}$$

Donde:

z_j	valores calculados por la red
a_i	salidas de la primera capa oculta
Y_m	salida de la segunda capa oculta
V_{nl}	pesos iniciales del modelo
W_{ml}	pesos de la segunda capa oculta
U_{mj}	pesos de la capa de salida

3 RESULTADOS Y DISCUSIÓN

El estudio de materiales adsorbentes en el tratamiento de aguas residuales en las últimas décadas ha sido de gran interés, sobre todo aquellos de origen natural como minerales arcillosos, zeolitas, óxidos o biopolímeros. El quitosano es un biopolímero que ha sido estudiado ampliamente con fines de adsorción dando buenos resultados.

En esta investigación se utilizaron perlas o esferas de tres tamaños diferentes de un compósito de quitosano – poli (vinil alcohol) para la remoción del colorante amarillo 5 para generar conocimiento del equilibrio de adsorción con este tipo de materiales novedosos y puedan compararse con los materiales tradicionales como el carbón activado.

Para ello a continuación se presentan los resultados experimentales en un proceso de adsorción por lotes, en donde se muestra la capacidad adsorbente del material a diferentes tiempos de contacto con el colorante en solución acuosa.

La interpretación de los resultados se realizó considerando tres criterios complementarios: la magnitud de la capacidad de adsorción en equilibrio, el tiempo estimado para alcanzar dicho equilibrio y el coeficiente de determinación. Esta lectura integral es necesaria porque un modelo puede presentar un R^2 elevado y, al mismo tiempo, estimar un tiempo de equilibrio poco útil para la operación del proceso. En sistemas por lotes, el valor de q_e informa sobre la eficiencia potencial del material, mientras que el tiempo de equilibrio define la viabilidad operativa, el tiempo de residencia requerido y la posibilidad de reducir ensayos prolongados. De acuerdo con el enfoque reportado para este sistema, la RNA se valoró no sólo por su ajuste estadístico, sino por su capacidad para anticipar la región de equilibrio con menor tiempo experimental que los modelos cinéticos clásicos (Ghaedi & Vafaei, 2017; Fang et al., 2020; Domínguez Beltrán et al., 2025).

Para describir los resultados, cada conjunto de datos se analizó de forma separada según el tamaño de perla y la temperatura. Esta organización permite reconocer

tendencias que podrían ocultarse si todos los datos se interpretaran de manera global. El tamaño de perla modifica la relación superficie-volumen y la distancia de difusión hacia el interior del hidrogel, mientras que la temperatura puede alterar la movilidad molecular del colorante, la flexibilidad de la red polimérica y la fuerza relativa de las interacciones entre tartrazina y grupos funcionales del quitosano. Por ello, la discusión no se limita a reportar valores de R^2 , sino que relaciona la respuesta del modelo con posibles efectos de transferencia de masa y disponibilidad de sitios activos (Crini & Badot, 2008; Saheed & Suah, 2023; Li et al., 2025).

La lectura de las tablas se realizó considerando que un modelo útil para el diseño del proceso debe cumplir dos condiciones: reproducir adecuadamente los datos experimentales y entregar parámetros interpretables para la operación. En ese sentido, la RNA se valora por estimar simultáneamente q_e y tiempo de equilibrio, mientras que Elovich, aunque puede alcanzar R^2 altos, no proporciona de forma directa ambos parámetros. Esta diferencia es importante porque en ingeniería ambiental no basta con ajustar una curva; también se requiere que el modelo apoye decisiones sobre duración del tratamiento, comparación entre tamaños de adsorbente y selección de condiciones de operación (Ghaedi & Vafaei, 2017; Karam et al., 2020; Domínguez Beltrán et al., 2025).

3.1 Comparativo de RNA con modelos tradicionales

En las figuras presentadas en esta sección se muestra el comportamiento cinético de adsorción del colorante amarillo No. 5 sobre perlas de hidrogel de quitosano-poli(vinil alcohol), comparando los datos experimentales con los modelos cinéticos tradicionales y con el modelo de red neuronal artificial. El propósito de este comparativo es evaluar la capacidad predictiva de la RNA frente a modelos ampliamente utilizados en estudios de adsorción, tales como Lagergren, Ho-McKay y Elovich, considerando diferentes tamaños de esfera y temperaturas de operación.

De manera general, las curvas muestran un comportamiento característico de los procesos de adsorción por lotes: una etapa inicial de rápida remoción del colorante,

seguida por una fase de adsorción más lenta hasta aproximarse al equilibrio. La primera etapa se asocia con la disponibilidad inicial de sitios activos libres en la superficie del hidrogel, lo que permite una interacción rápida entre las moléculas de tartrazina y los grupos funcionales del quitosano, principalmente grupos amino e hidroxilo. Posteriormente, conforme los sitios activos disponibles disminuyen y el gradiente de concentración entre la solución y el adsorbente se reduce, la velocidad de adsorción decrece gradualmente hasta alcanzar una zona de estabilización.

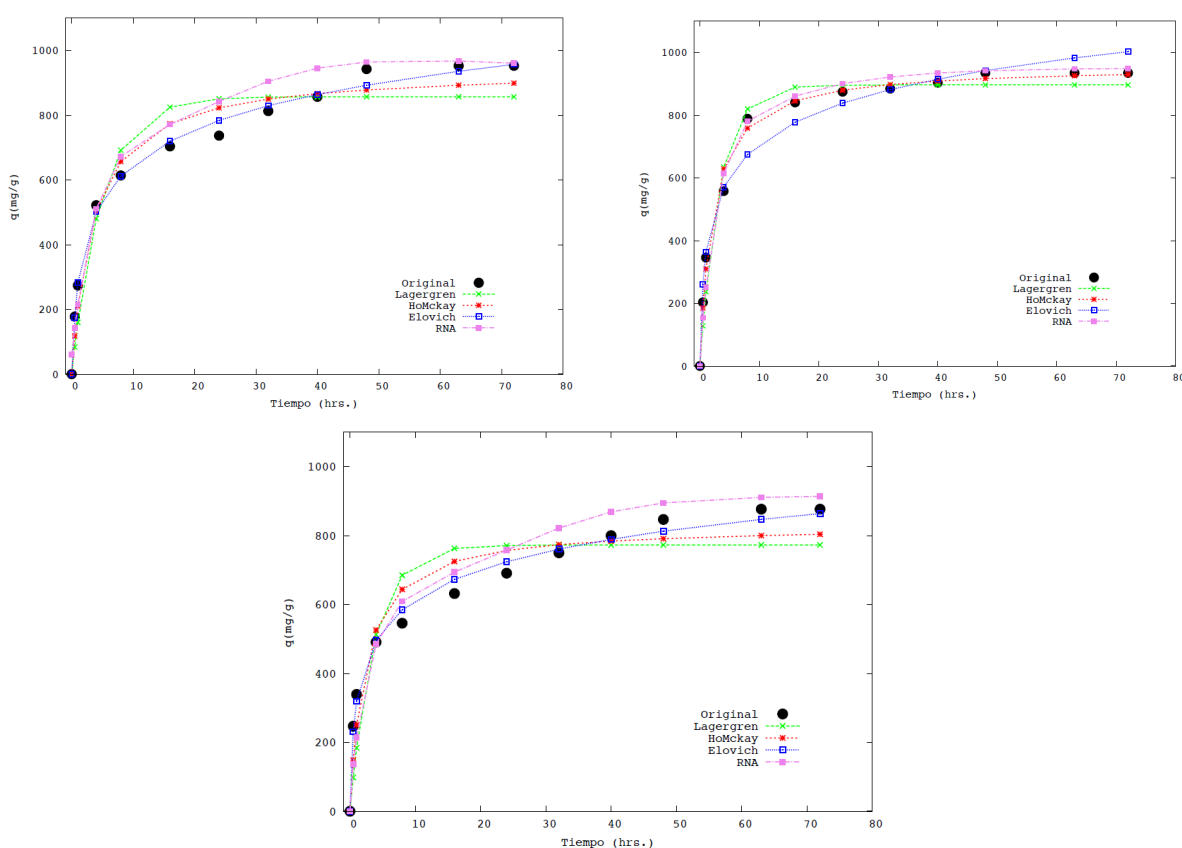


Figura 5. Resultados de predicción para la esfera chica.

En el caso de las esferas chicas evaluadas a 10, 30 y 50 °C, se observa que la red neuronal artificial reproduce adecuadamente la tendencia de los datos experimentales durante la mayor parte del proceso. Las curvas correspondientes a los modelos tradicionales también describen la tendencia general de la adsorción; sin embargo, presentan diferencias en ciertas regiones, principalmente en la etapa inicial y en la zona cercana al equilibrio. Esta diferencia es relevante, ya que en los

procesos de adsorción no sólo interesa alcanzar un valor alto de ajuste estadístico, sino también estimar correctamente el momento en el que el sistema se estabiliza y la capacidad de adsorción deja de cambiar de manera significativa.

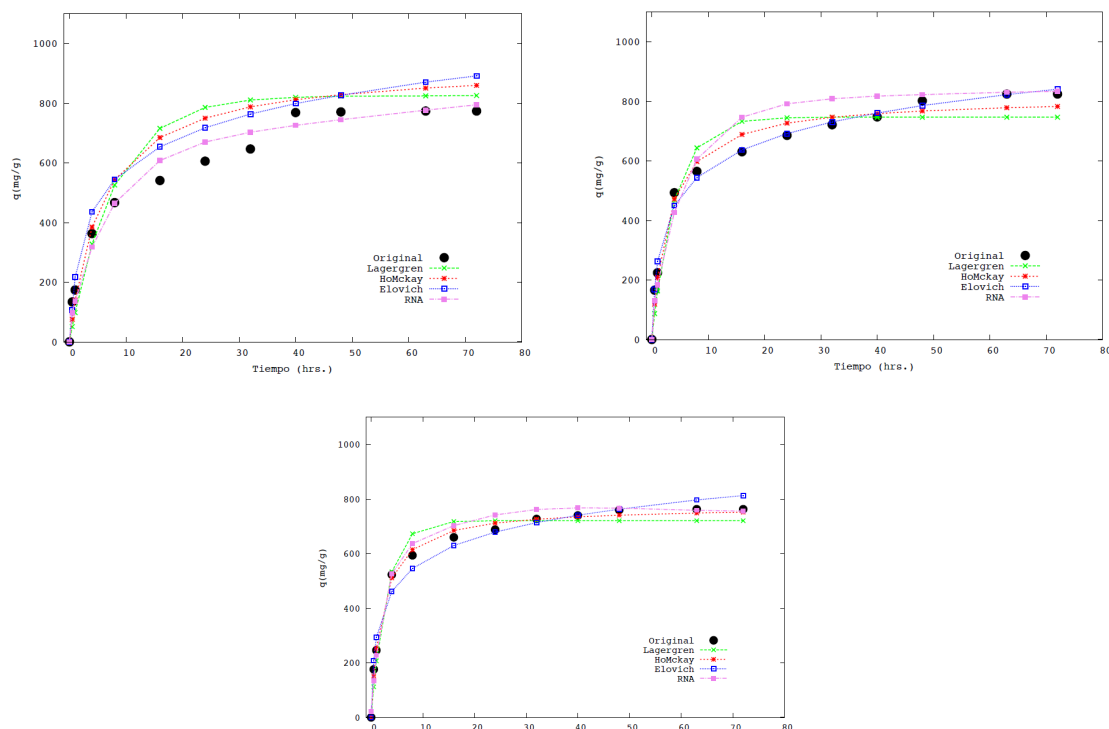


Figura 6. Resultados de predicción para la esfera mediana.

Para las esferas medianas, las imágenes muestran un comportamiento similar, aunque con variaciones en la velocidad con la que se alcanza la zona de equilibrio. La RNA mantiene una tendencia cercana a los datos experimentales, especialmente en las regiones intermedias de la curva, donde ocurre una parte importante del proceso de transferencia de masa. En este tamaño de esfera, la influencia de la difusión interna puede ser más evidente que en las esferas chicas, debido a que el recorrido que debe realizar el colorante hacia el interior del material es mayor. Por ello, las diferencias entre los modelos no deben interpretarse únicamente como un problema de ajuste matemático, sino como una manifestación de la complejidad del fenómeno de adsorción en materiales poliméricos hidratados.

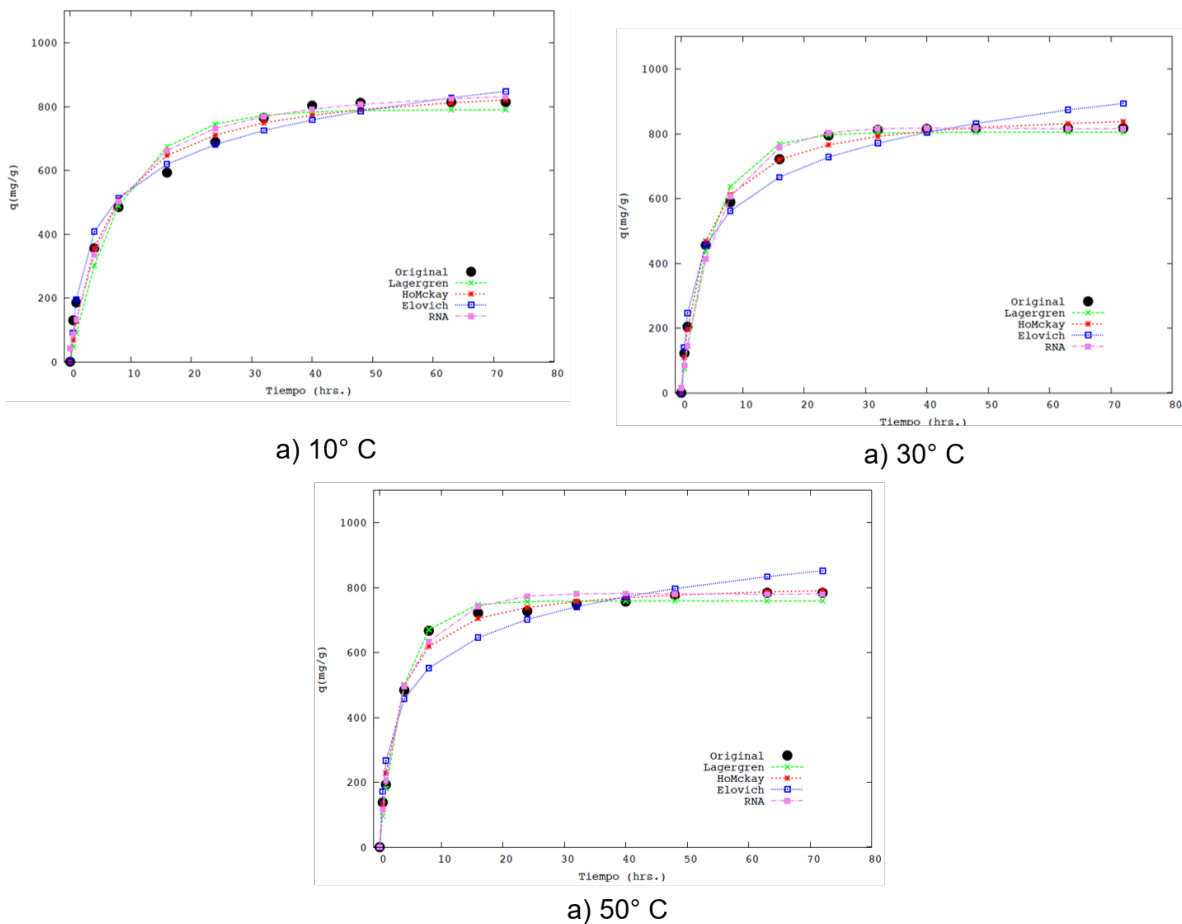


Figura 7. Resultados de predicción para la esfera grande.

En las esferas grandes, el ajuste de los modelos muestra que el proceso de adsorción continúa siendo dependiente del tiempo y de la temperatura, pero con una posible mayor resistencia a la transferencia de masa. Al aumentar el tamaño de la esfera, disminuye la relación área superficial/volumen, lo que puede reducir la disponibilidad inmediata de sitios activos expuestos al colorante. En consecuencia, aunque la capacidad de adsorción puede mantenerse en valores relevantes, el sistema puede requerir mayor tiempo para que las moléculas de tartrazina difundan hacia regiones internas del hidrogel. Las imágenes permiten observar que la RNA y algunos modelos tradicionales, particularmente Ho-McKay, presentan un comportamiento cercano a los datos experimentales; no obstante, la RNA ofrece la ventaja de estimar directamente el comportamiento del sistema sin imponer una forma cinética rígida.

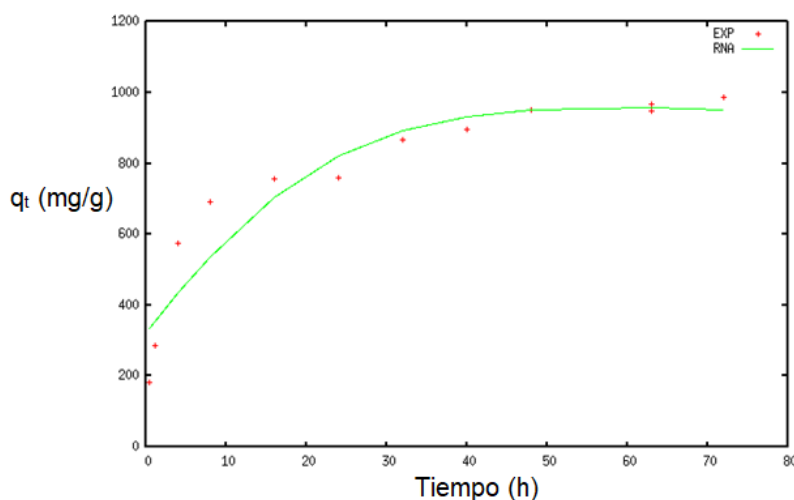


Figura 8. Modelo de RNA para esfera chica a 10 ° C. Con 10 y 10 neuronas en las capas ocultas.

En la Figura 3.1 se observa que para la esfera chica a 10 °C la RNA predice una capacidad de adsorción en equilibrio de 945 mg/g en un tiempo de 40 h, con un coeficiente de determinación de $R^2 = 0.9428$. Este resultado indica que el modelo neuronal logra representar adecuadamente la tendencia experimental del sistema, identificando una zona de equilibrio antes de concluir las 72 h de experimentación. En comparación, el modelo de Lagergren estima una capacidad menor, de 857 mg/g, y requiere 57 h para aproximarse al equilibrio, lo cual sugiere una representación más limitada de la etapa final del proceso. Por su parte, el modelo de Ho-McKay predice un valor de Q_e muy cercano al de la RNA, con 943 mg/g, y presenta un R^2 superior de 0.9701; sin embargo, requiere 72 h para alcanzar el equilibrio, por lo que su ajuste estadístico no necesariamente implica una mejor utilidad práctica en términos de reducción del tiempo experimental. El modelo de Elovich presenta el mayor valor de R^2 con 0.9946, lo cual evidencia un buen ajuste a la tendencia de los datos; no obstante, al no proporcionar directamente una estimación de Q_e ni del tiempo de equilibrio, su utilidad para el diseño operativo del proceso es limitada. En conjunto, los resultados muestran que, bajo esta condición, la RNA ofrece una predicción equilibrada entre capacidad adsorbida, tiempo de equilibrio y ajuste estadístico.

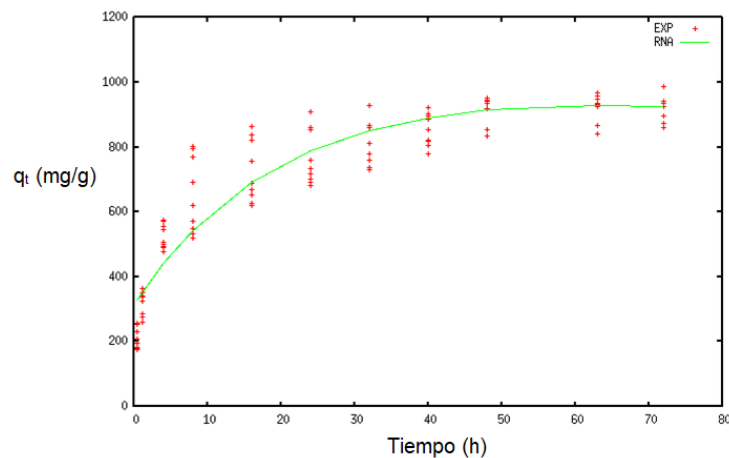


Figura 9. Modelo de RNA para esfera chica a 10, 30 y 50 °C, con 2 y 2 neuronas en las capas ocultas.

Para la esfera chica a 30 °C, en la Figura 3.2 se observa que la RNA estima una capacidad de adsorción en equilibrio de 934 mg/g en 40 h, con un coeficiente de determinación de $R^2 = 0.9720$. Este valor refleja una alta concordancia entre el modelo neuronal y los datos experimentales, lo que indica que la RNA logra capturar de manera adecuada el comportamiento cinético del sistema a esta temperatura. El modelo de Lagergren predice una capacidad de 897 mg/g en 44 h, con un $R^2 = 0.9729$, muy cercano al obtenido por la RNA; sin embargo, su capacidad estimada es menor, lo que puede indicar una ligera subestimación del potencial adsorbente del hidrogel. En contraste, el modelo de Ho-McKay predice una capacidad mayor, de 957 mg/g, con el coeficiente de determinación más alto de esta condición, $R^2 = 0.9929$, pero requiere más de 72 h para alcanzar el equilibrio. Esto sugiere que, aunque el modelo de pseudo-segundo orden ajusta bien los datos, tiende a prolongar la estimación del tiempo necesario para estabilizar el proceso. El modelo de Elovich también presenta un ajuste aceptable, con $R^2 = 0.9752$, pero nuevamente no permite obtener directamente la capacidad ni el tiempo de equilibrio. Por lo tanto, para 30 °C, la RNA resulta particularmente útil porque mantiene un ajuste alto y permite reducir el tiempo estimado de equilibrio a 40 h.

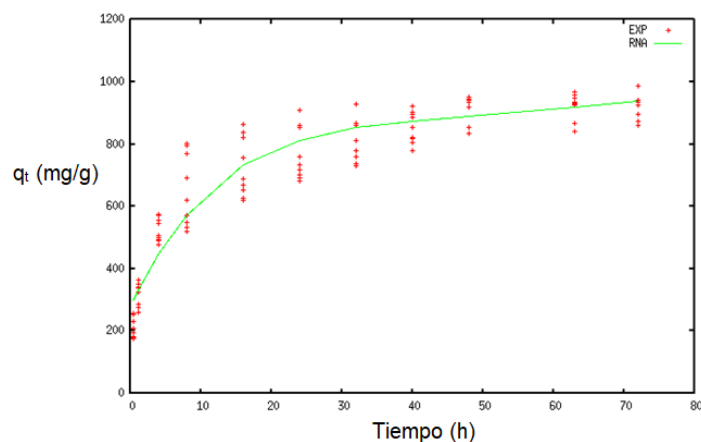


Figura 10. Modelo de RNA para esfera chica a 10, 30 y 50 °C, con 10 y 10 neuronas en las capas ocultas.

En la Figura 3.3 se muestra que para la esfera chica a 50 °C, la RNA predice una capacidad de adsorción de 869 mg/g en 40 h, con un $R^2 = 0.8830$. Aunque este coeficiente de determinación es menor que el observado a 10 y 30 °C, el modelo neuronal conserva una representación razonable del comportamiento general del sistema. Esta disminución en el ajuste puede asociarse con una mayor variabilidad experimental a temperaturas elevadas, así como con posibles cambios en el comportamiento del hidrogel, tales como modificaciones en su hinchamiento, movilidad de las cadenas poliméricas o interacción entre el colorante y los grupos funcionales del adsorbente. El modelo de Lagergren estima una capacidad menor, de 773 mg/g, en 48 h, con $R^2 = 0.8614$, por lo que presenta un desempeño inferior al de la RNA tanto en capacidad estimada como en ajuste. El modelo de Ho-McKay predice 830 mg/g, con $R^2 = 0.9335$, pero requiere más de 72 h para alcanzar el equilibrio, lo cual nuevamente limita su aplicación cuando se busca reducir el tiempo experimental. El modelo de Elovich alcanza un $R^2 = 0.9913$, indicando un buen ajuste de la forma de la curva, aunque sin proporcionar valores directos de Q_e ni tiempo de equilibrio. De esta forma, aun cuando la RNA presenta menor ajuste estadístico a 50 °C, continúa siendo útil para estimar de forma directa la capacidad y el tiempo de estabilización del sistema.

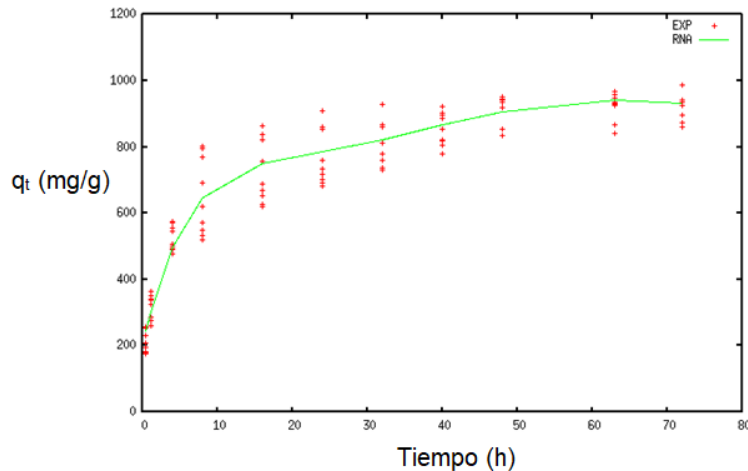


Figura 11. Modelo de RNA para esfera chica a 10, 30 y 50 °C, con 20 y 20 neuronas en las capas ocultas.

En la Figura 3.4 se presenta el modelo de RNA para la esfera chica considerando simultáneamente las temperaturas de 10, 30 y 50 °C, con una arquitectura de 20 neuronas en la primera capa oculta y 20 neuronas en la segunda capa oculta. En esta gráfica se observa una mayor dispersión de los datos experimentales debido a que se integran diferentes condiciones térmicas en una misma representación. A pesar de esta variabilidad, la curva generada por la RNA mantiene la tendencia general del proceso de adsorción: un incremento rápido de la capacidad de adsorción durante las primeras horas, seguido de una etapa de crecimiento más lento hasta aproximarse a una zona de equilibrio. Este comportamiento confirma que el modelo neuronal logra identificar la relación no lineal entre el tiempo de contacto y la cantidad adsorbida, aun cuando los datos provienen de condiciones experimentales distintas. El uso de 20 y 20 neuronas en las capas ocultas permite aumentar la capacidad de aprendizaje del modelo, ya que la red cuenta con un mayor número de parámetros ajustables para representar la variabilidad de los datos. Sin embargo, también se observa que el incremento en la complejidad de la arquitectura no necesariamente implica una reproducción perfecta de todos los puntos experimentales, por ello, la red describe de manera adecuada el comportamiento promedio de la esfera chica bajo las tres temperaturas evaluadas, aunque persisten diferencias asociadas a la dispersión experimental y a la influencia de la temperatura sobre el proceso de adsorción.

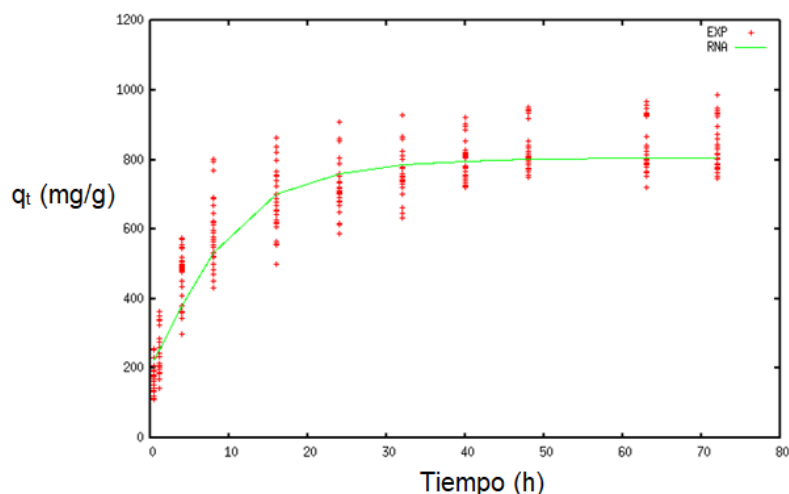


Figura 12. Modelo de RNA para esfera chica, mediana y grande a 10, 30 y 50 °C, con 2 y 2 neuronas en las capas ocultas.

En la Figura 3.5 se muestra el modelo de RNA aplicado de manera conjunta a las esferas chica, mediana y grande, evaluadas a 10, 30 y 50 °C, utilizando una arquitectura de 2 neuronas en la primera capa oculta y 2 neuronas en la segunda capa oculta. En esta representación se integra una mayor heterogeneidad experimental, ya que además de la temperatura se incorpora el efecto del tamaño de esfera. La dispersión de los puntos experimentales es más evidente, debido a que cada tamaño de esfera presenta una relación área superficial/volumen distinta y, por lo tanto, una respuesta cinética particular frente al colorante. Aun así, la RNA logra generar una curva que representa la tendencia general del sistema.

El comportamiento mostrado indica que una arquitectura con pocas neuronas puede ser suficiente para describir el patrón global de la adsorción, pero puede resultar limitada para capturar variaciones más específicas entre tamaños de esfera y temperaturas. La curva obtenida reproduce la etapa inicial de adsorción rápida y la posterior estabilización, pero al tratarse de una red de baja complejidad, su capacidad para ajustarse a cambios locales de los datos es menor. Esta figura es importante porque muestra el punto de partida del análisis arquitectónico de la RNA: una red sencilla puede representar la tendencia principal del fenómeno, aunque con menor sensibilidad ante la variabilidad experimental.

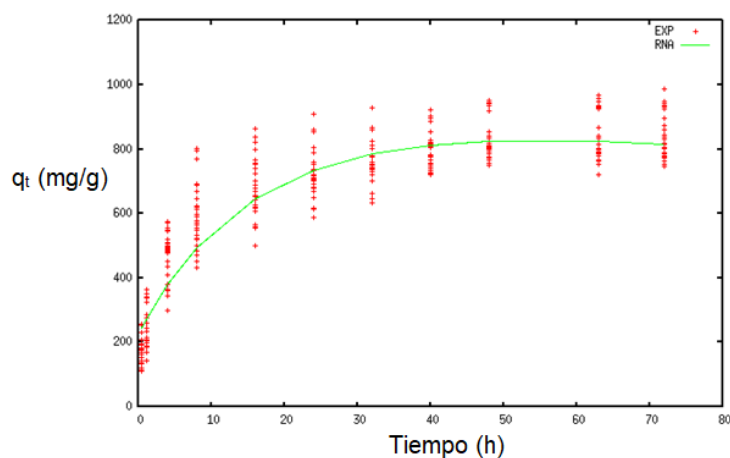


Figura 13. Modelo de RNA para esfera chica, mediana y grande a 10, 30 y 50 °C, con 10 y 10 neuronas en las capas ocultas.

La Figura 3.6 presenta nuevamente el modelo general para esferas chica, mediana y grande a 10, 30 y 50 °C, pero con una arquitectura de 10 neuronas en la primera capa oculta y 10 neuronas en la segunda capa oculta. En comparación con la Figura 3.5, el incremento en el número de neuronas permite que la RNA tenga mayor flexibilidad para representar la dispersión de los datos experimentales. La curva generada conserva la forma típica del proceso de adsorción por lotes, con una fase inicial de incremento pronunciado y una etapa posterior de aproximación al equilibrio. Esta forma es coherente con la disponibilidad inicial de sitios activos en el hidrogel y con la disminución progresiva de la fuerza impulsora de transferencia de masa conforme avanza el tiempo.

La arquitectura de 10 y 10 neuronas representa un equilibrio entre simplicidad y capacidad predictiva. A diferencia de una red demasiado pequeña, esta configuración puede capturar con mayor precisión la influencia combinada del tamaño de esfera y la temperatura, sin incrementar excesivamente la complejidad del modelo. La dispersión de los datos alrededor de la curva de la RNA refleja que el comportamiento del sistema no depende únicamente del tiempo, sino también de factores físicos asociados al tamaño del adsorbente, como la difusión interna y la disponibilidad de área superficial. Por ello, la Figura 3.6 evidencia que la RNA puede

funcionar como una herramienta de predicción general cuando se incorporan diferentes condiciones operativas en una misma base de datos.

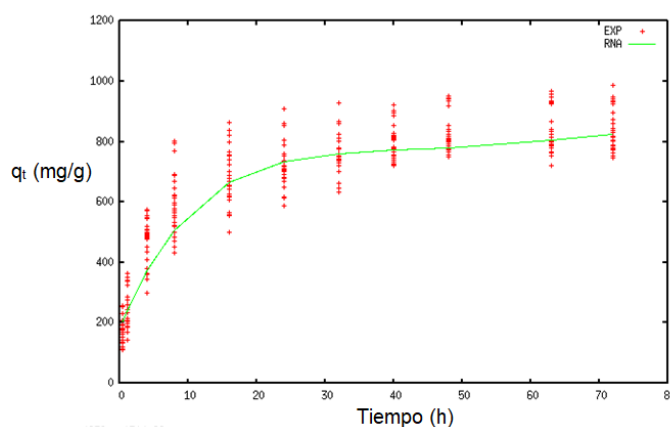


Figura 14. Modelo de RNA para esfera chica, mediana y grande a 10, 30 y 50 °C, con 20 y 20 neuronas en las capas ocultas.

El análisis de la Figura 3.7 permite observar que el aumento en el número de neuronas puede mejorar la capacidad de la RNA para aproximarse a los datos experimentales; sin embargo, también debe considerarse que una arquitectura más compleja no siempre garantiza una mejora proporcional en la predicción. En redes neuronales, el exceso de neuronas puede aumentar el riesgo de sobreajuste, es decir, que el modelo aprenda detalles particulares de los datos de entrenamiento sin mejorar necesariamente su capacidad de generalización. Por ello, la selección de la arquitectura debe realizarse considerando no sólo el ajuste visual de la curva, sino también indicadores cuantitativos como el coeficiente de determinación, el error de predicción y la coherencia fisicoquímica del comportamiento modelado.

En conjunto, las Figuras 3.4 a 3.7 permiten comparar el efecto de la arquitectura de la RNA sobre la representación del proceso de adsorción. Las arquitecturas con menor número de neuronas muestran una tendencia general adecuada, pero con menor capacidad para reproducir la variabilidad experimental. Por su parte, las arquitecturas con mayor número de neuronas ofrecen una mayor flexibilidad para aproximarse a los datos, aunque requieren una evaluación cuidadosa para evitar una interpretación excesivamente dependiente del ajuste gráfico. Este análisis es

relevante porque demuestra que el desempeño de una RNA no depende únicamente de la cantidad de datos disponibles, sino también de la correcta selección de entradas, salidas, capas ocultas y número de neuronas.

Asimismo, estas figuras muestran que el fenómeno de adsorción evaluado posee un comportamiento no lineal, influido por el tiempo de contacto, la temperatura y el tamaño de las esferas de hidrogel. La RNA resulta adecuada para este tipo de sistemas porque no requiere imponer una ecuación cinética fija, sino que aprende la tendencia a partir de los datos experimentales. De esta manera, el modelo neuronal permite representar de forma integrada el comportamiento de diferentes condiciones operativas y proporciona una base para estimar el equilibrio de adsorción en menor tiempo que el requerido por una experimentación prolongada.

Por lo tanto, la comparación entre las arquitecturas de 2-2, 10-10 y 20-20 neuronas evidencia que la RNA puede adaptarse progresivamente a la complejidad del sistema. No obstante, la arquitectura más conveniente debe seleccionarse con base en el equilibrio entre precisión, estabilidad y capacidad de generalización. En este trabajo, las figuras permiten visualizar que la red neuronal es capaz de reproducir la tendencia cinética del proceso de adsorción de tartrazina sobre hidrogel de quitosano-PVA, tanto para una condición específica como para el conjunto de tamaños de esfera y temperaturas evaluadas.

Resultados del valor q_e para la esfera chica

Tabla 4. Comparativa de resultados de q_e para la esfera chica con la temperatura 10°.

Temperatura 10 °C				
Modelo	Q_e (mg/g)	Tiempo	R ²	K
RNA	945	40	0.9428	23.625
Lagergren	857	57	0.9258	14.70237
Ho-Mckay	943	72	0.9701	20.69924
Elovich	-	-	0.9946	-

La tabla de esfera chica a 10 °C muestra que la RNA predice una capacidad de equilibrio de 945 mg/g en 40 h, con un R^2 de 0.9428. Este resultado es relevante

porque se aproxima al valor estimado por Ho-McKay, pero con un tiempo de equilibrio considerablemente menor. Lagergren subestima la capacidad de adsorción y requiere más tiempo para aproximarse al equilibrio, mientras que Elovich presenta un ajuste estadístico alto, aunque no proporciona directamente q_e ni tiempo de equilibrio. En conjunto, la comparación indica que, para esta condición, la RNA ofrece una predicción operativamente más útil, ya que identifica una región de equilibrio antes de las 72 h del experimento.

Tabla 5. Comparativa de resultados de q_e para la esfera chica con la temperatura 30°.

Temperatura 30 °C				
Modelo	Q_e (mg/g)	Tiempo	R ²	K
RNA	934	40	0.9720	23.35
Lagergren	897	44	0.9729	21.93497
Ho-Mckay	957	72+	0.9929	33.47255
Elovich	-	-	0.9752	-

La tabla de esfera chica a 30 °C evidencia el mejor desempeño global de la RNA dentro de las condiciones evaluadas para este tamaño, con $q_e = 934$ mg/g, tiempo de equilibrio de 40 h y $R^2 = 0.9720$. Aunque Ho-McKay alcanza el R^2 más alto y predice una capacidad ligeramente mayor, su estimación se desplaza más allá de 72 h, lo que reduce su utilidad para definir tiempos prácticos de operación. La cercanía entre los valores de Lagergren y RNA muestra que ambos siguen adecuadamente la tendencia experimental, pero la RNA conserva la ventaja de ofrecer una lectura directa y rápida del equilibrio.

Tabla 6. Comparativa de resultados de q_e para la esfera grande con temperatura 50°.

Temperatura 50 °C				
Modelo	Q_e (mg/g)	Tiempo	R ²	K
RNA	869	40	0.8830	21.725
Lagergren	773	48	0.8614	19.45236
Ho-Mckay	830	72+	0.9335	32.67626
Elovich	-	-	0.9913	-

La tabla de esfera chica a 50 °C muestra una disminución de la capacidad de equilibrio estimada por la RNA hasta 869 mg/g, con equilibrio a 40 h y $R^2 = 0.8830$. Aunque este valor de R^2 es menor que el de otras condiciones, la RNA mantiene una estimación de q_e superior a Lagergren y cercana al comportamiento experimental observado. Ho-McKay mejora el ajuste estadístico, pero nuevamente proyecta el equilibrio por encima de 72 h. Esta condición sugiere que el aumento de temperatura modifica el perfil cinético de la esfera chica y puede intensificar diferencias entre modelos, especialmente en la zona intermedia y final de la curva.

Resultados del valor q_e para la esfera mediana

Tabla 7. Comparativa de resultados de q_e para la esfera mediana con la temperatura 10°.

Temperatura 10 °C				
Modelo	Q_e (mg/g)	Tiempo	R ²	K
RNA	777	63	0.9658	12.3333
Lagergren	729	72+	0.9405	9.0381
Ho-Mckay	819	72+	0.9732	11.9633
Elovich	-	-	0.9865	-

La tabla de esfera mediana a 10 °C indica que la RNA predice $q_e = 777$ mg/g en 63 h, con $R^2 = 0.9658$. El valor de q_e se ubica entre Lagergren y Ho-McKay, lo que sugiere una estimación intermedia y consistente con la tendencia general de los datos. A diferencia de Ho-McKay, que requiere más de 72 h para alcanzar el equilibrio, la RNA delimita una zona de estabilización dentro del intervalo experimental.

Tabla 8. Comparativa de resultados de q_e para la esfera mediana con la temperatura 30°.

Temperatura 30 °C				
Modelo	Q_e (mg/g)	Tiempo	R ²	K
RNA	823	48	0.9114	17.1458
Lagergren	747	49	0.9465	17.6125
Ho-Mckay	815	72+	0.9829	24.7606
Elovich	-	-	0.9946	-

Este comportamiento es congruente con un tamaño de perla mayor que el de la esfera chica, donde la transferencia de masa puede requerir tiempos más prolongados.

La tabla de esfera mediana a 30 °C muestra que la RNA estima la mayor capacidad para este tamaño, con $q_e = 823$ mg/g en 48 h. Aunque su $R^2 = 0.9114$ es inferior al de los modelos tradicionales, su predicción de capacidad y tiempo resulta razonable para fines de operación, ya que ubica el equilibrio antes de 72 h. Lagergren predice un tiempo cercano, pero con menor capacidad de adsorción, mientras que Ho-McKay conserva un ajuste elevado a costa de un tiempo de equilibrio más largo. El resultado refleja que la evaluación del modelo no debe depender únicamente del R^2 , sino también del significado práctico de la predicción.

Tabla 9. Comparativa de resultados de q_e para la esfera mediana con la temperatura 50°.

Temperatura 50 °C				
Modelo	Q_e (mg/g)	Tiempo	R²	K
RNA	762	32	0.9754	23.8125
Lagergren	721	35	0.9716	24.09825
Ho-Mckay	773	72+	0.9958	34.31186
Elovich	-	-	0.9817	-

La tabla de esfera mediana a 50 °C presenta uno de los mejores desempeños de la RNA para este tamaño, con $q_e = 762$ mg/g en 32 h y $R^2 = 0.9754$. Este resultado indica una aproximación rápida al equilibrio y una representación adecuada del perfil cinético. Lagergren ofrece valores cercanos tanto en q_e como en tiempo, mientras que Ho-McKay muestra el R^2 más alto, pero mantiene la tendencia a estimar el equilibrio después de 72 h. La comparación sugiere que, bajo esta condición, la RNA permite reducir el intervalo experimental requerido para identificar la estabilización del sistema.

Resultados del valor q_e para la esfera grande

La tabla de esfera grande a 10 °C muestra que la RNA predice $q_e = 807$ mg/g en 48 h, con $R^2 = 0.9810$. En esta condición, el modelo neuronal presenta un ajuste

robusto y una capacidad estimada menor que Ho-McKay, pero con una reducción importante del tiempo de equilibrio.

Tabla 10. Comparativa de resultados de q_e para la esfera grande con la temperatura 10°.

Temperatura 10 °C				
Modelo	Q_e (mg/g)	Tiempo	R ²	K
RNA	807	48	0.981	16.8125
Lagergren	790	72+	0.9682	8.60486
Ho-Mckay	890	72+	0.9869	10.90084
Elovich	-	-	0.9883	-

Lagergren también ofrece una buena aproximación, aunque su predicción temporal se extiende más allá de 72 h. La lectura conjunta indica que la esfera grande conserva una capacidad considerable, pero su comportamiento requiere analizar con atención el equilibrio entre difusión interna, disponibilidad de sitios y estabilidad de la curva.

Tabla 11. Comparativa de resultados de q_e para la esfera grande con la temperatura 30°.

Temperatura 30 °C				
Modelo	Q_e (mg/g)	Tiempo	R ²	K
RNA	818	48	0.9876	17.04167
Lagergren	805	62	0.9890	14.0003
Ho-Mckay	879	72+	0.9973	18.99397
Elovich	-	-	0.9792	-

La tabla de esfera grande a 30 °C evidencia el mayor R^2 obtenido por la RNA para este tamaño, con $q_e = 818$ mg/g, equilibrio a 48 h y $R^2 = 0.9876$. Ho-McKay alcanza el R^2 más alto de toda la serie, pero predice el equilibrio por encima de 72 h. Lagergren estima una capacidad cercana y un tiempo de 62 h. En términos de aplicación, la RNA resulta ventajosa porque mantiene un ajuste alto y reduce el tiempo estimado para alcanzar una condición estable, lo que favorece la planeación de ensayos y el diseño preliminar del proceso.

Tabla 12. Comparativa de resultados de q_e para la esfera grande con la temperatura 50°.

Temperatura 50 °C				
Modelo	Q_e (mg/g adsorbido)	Tiempo	R ²	K
RNA	781	40	0.9893	19.525
Lagergren	759	45	0.9938	19.29374
Ho-Mckay	819	72+	0.9951	26.42141
Elovich	-	-	0.9621	-

La tabla de esfera grande a 50 °C muestra que la RNA estima $q_e = 781$ mg/g en 40 h, con $R^2 = 0.9893$. Esta condición confirma la estabilidad predictiva del modelo neuronal para las perlas de mayor tamaño, ya que ofrece un ajuste alto y un tiempo de equilibrio menor que Ho-McKay. Aunque Lagergren también predice un tiempo cercano y un R^2 elevado, la RNA conserva una estimación equilibrada entre capacidad, rapidez y ajuste. El conjunto de resultados para esfera grande sugiere que la RNA puede describir adecuadamente sistemas donde la difusión interna y el tamaño de partícula influyen de manera más marcada.

Discusión de resultados

Si bien los modelos cinéticos tradicionales, como los modelos de pseudo-primer orden, pseudo-segundo orden y Elovich, se emplean ampliamente en estudios de adsorción debido a su versatilidad y relativa facilidad de aplicación, están inherentemente limitados por una serie de supuestos simplificadores que restringen su poder predictivo y su interpretabilidad mecanicista.

Una limitación principal radica en su suposición de homogeneidad superficial y energía de activación uniforme en todos los sitios de adsorción. Estas condiciones rara vez se cumplen en sistemas reales que involucran materiales heterogéneos, como las perlas de hidrogel de quitosano y alcohol polivinílico. Estos modelos a menudo no consideran fenómenos complejos como la fisisorción y la quimisorción simultáneas, los efectos de difusión por poros o los mecanismos de adsorción multietapa.

En el caso específico de las perlas de quitosano-PVA, la heterogeneidad del adsorbente se relaciona con la coexistencia de regiones más hidratadas, dominios poliméricos de distinta accesibilidad y grupos funcionales distribuidos de manera no uniforme. Esta condición favorece que el proceso no se comporte como una reacción simple de un solo paso. Por ello, modelos como Elovich pueden obtener valores altos de R^2 al representar una superficie heterogénea, pero no aportan directamente parámetros de diseño como q_e o el tiempo de equilibrio. Esta limitación es importante porque en ingeniería ambiental la utilidad del modelo depende de su capacidad para orientar decisiones experimentales y operativas, no únicamente de reproducir una curva (Guo & Wang, 2019; Hubbe et al., 2019; Fang et al., 2020).

La influencia del tamaño de perla puede interpretarse a partir de la relación superficie-volumen. Las perlas chicas tienen una mayor área externa relativa y una menor longitud característica de difusión, por lo que tienden a presentar capacidades elevadas y una aproximación más rápida a la zona de equilibrio. En contraste, al aumentar el diámetro de la perla, la difusión hacia el interior del hidrogel adquiere mayor peso y puede retrasar la estabilización de la curva. Esta interpretación coincide con estudios sobre perlas de quitosano e hidrogeles, donde el tamaño, la porosidad y la estabilidad estructural condicionan el acceso del adsorbato a los grupos funcionales disponibles (Crini & Badot, 2008; Saheed & Suah, 2023; Li et al., 2025).

La temperatura también modifica la trayectoria de adsorción, aunque su efecto no siempre es lineal. Un incremento de temperatura puede favorecer la movilidad molecular y acelerar etapas de difusión; sin embargo, también puede alterar la afinidad adsorbato-adsorbente, la estructura hidratada del hidrogel o el equilibrio de interacciones electrostáticas. Por ello, la respuesta observada en las tres temperaturas no debe interpretarse como una simple mejora o disminución uniforme, sino como el resultado de interacciones simultáneas entre transporte de masa, disponibilidad de sitios activos y estabilidad del material. Esta complejidad justifica el uso de herramientas de modelado no lineal para complementar los modelos cinéticos convencionales (Rápó & Tonk, 2021; Dutta et al., 2021; Domínguez Beltrán et al., 2025).

Los resultados obtenidos respaldan la pertinencia de la RNA porque el modelo no se limita a ajustar una ecuación preestablecida, sino que aprende la forma de las curvas experimentales bajo distintas combinaciones de tamaño y temperatura. Esta ventaja es especialmente notable cuando se compara el tiempo estimado de equilibrio: mientras el modelo de Ho-McKay alcanza con frecuencia R^2 superiores, su predicción se desplaza hacia tiempos mayores a 72 h, lo cual puede sobreestimar los tiempos requeridos para ensayos de laboratorio o para el dimensionamiento preliminar. La RNA, en cambio, identifica intervalos de equilibrio entre 32 y 63 h, reduciendo el tiempo de exploración experimental sin perder precisión aceptable (Ghaedi & Vafaei, 2017; Pauletto et al., 2020; da Costa et al., 2025).

A pesar de estas ventajas, el modelo neuronal debe considerarse una herramienta predictiva complementaria y no una sustitución absoluta de la interpretación fisicoquímica. Su desempeño depende de la calidad de la base de datos, del rango experimental utilizado para el entrenamiento y de la selección de hiperparámetros. Por ello, una ampliación futura debería incluir validación externa con otros colorantes, concentraciones iniciales, pH y matrices reales de agua residual, así como estrategias de optimización de arquitectura mediante algoritmos genéticos, enjambre de partículas u optimización bayesiana. Este enfoque permitiría aumentar la reproducibilidad del modelo y evaluar su generalización fuera de las condiciones experimentales originales (Cojocar et al., 2019; Al-Hameed et al., 2023; Vasiliauskaite & Antulov-Fantulin, 2024; Salahshoori et al., 2024).

4 CONCLUSIONES

La investigación demuestra que el modelado mediante redes neuronales artificiales es una alternativa para representar la cinética de adsorción de tartrazina sobre perlas de hidrogel de quitosano-PVA. El modelo permitió estimar la capacidad de adsorción en equilibrio y el tiempo necesario para alcanzar la estabilización del sistema bajo distintas condiciones de tamaño de perla y temperatura, manteniendo en la mayoría de los casos coeficientes de determinación elevados y coherentes con el comportamiento experimental.

La RNA mostró una ventaja práctica frente a los modelos cinéticos tradicionales porque identificó tiempos de equilibrio entre 32 y 63 h, mientras que el modelo de pseudo-segundo orden tendió a proyectar el equilibrio por encima de 72 h. Aunque Ho-McKay presentó en varios casos valores de R^2 ligeramente superiores, su estimación temporal resulta menos conveniente para reducir el número de ensayos experimentales. En contraste, la RNA proporcionó una predicción más útil para delimitar intervalos de trabajo y planificar experimentos con menor consumo de tiempo y recursos.

Los resultados también evidencian que el tamaño de la perla influye de manera importante en el desempeño del proceso. Las esferas chicas alcanzaron las capacidades de adsorción más altas, especialmente a 10 y 30 °C, lo que puede asociarse con una mayor superficie disponible y una menor distancia de difusión. Las esferas medianas y grandes conservaron un desempeño relevante, pero presentaron diferencias en los tiempos de equilibrio y en la magnitud de q_e , lo que demuestra que el diseño del adsorbente debe considerar simultáneamente capacidad, estabilidad y rapidez de adsorción.

La temperatura modificó el perfil cinético del sistema, pero su efecto no fue estrictamente lineal. En algunos casos favoreció una aproximación más rápida al equilibrio, mientras que en otros se observaron cambios en la capacidad estimada. Esto confirma que el proceso de adsorción en hidrogeles de quitosano-PVA está gobernado por interacciones simultáneas entre transporte de masa, estructura del material y afinidad química entre el colorante y los grupos funcionales del adsorbente.

El modelo neuronal obtenido no reemplaza por completo a los modelos cinéticos clásicos, ya que estos siguen siendo útiles para interpretar tendencias y comparar resultados con la literatura. Sin embargo, la RNA sí constituye una herramienta complementaria de alto valor para la predicción, optimización preliminar y reducción de trabajo experimental. Su principal aporte consiste en ofrecer una lectura más

flexible de un fenómeno no lineal y heterogéneo, especialmente cuando se busca identificar condiciones cercanas al equilibrio sin extender innecesariamente los ensayos.

Como trabajo futuro, se recomienda ampliar la base de datos con diferentes concentraciones iniciales, valores de pH, dosis de adsorbente, otros colorantes y muestras de agua residual real. También será conveniente automatizar la selección de hiperparámetros mediante técnicas de optimización y validar el modelo con datos externos. Estas mejoras permitirán fortalecer la reproducibilidad, aumentar la capacidad de generalización y consolidar el uso de redes neuronales artificiales como herramienta de apoyo para el diseño de procesos de remoción de colorantes en ingeniería ambiental.

5 REFERENCIAS

- Abdi, J., Vossoughi, M., Mahmoodi, N. M., & Alemzadeh, I. (2017). Synthesis of metal-organic framework hybrid nanocomposites based on GO and CNT with high adsorption capacity for dye removal. *Chemical Engineering Journal*, 326, 1145–1158. <https://doi.org/10.1016/j.cej.2017.06.054>
- Aghav, R. M., Kumar, S., & Mukherjee, S. N. (2011). Artificial neural network modeling in competitive adsorption of phenol and resorcinol from water environment using some carbonaceous adsorbents. *Journal of Hazardous Materials*, 188(1), 67–77. <https://doi.org/10.1016/j.jhazmat.2011.01.067>
- Alardhi, S. M., Salman, A. D., Al-Mashaqbeh, A. M., Al-Ansari, A. M., & Al-Obaidi, A. M. (2023). Prediction of methyl orange dye (MO) adsorption using activated carbon with an artificial neural network optimization modeling. *Heliyon*, 9, e12888. <https://doi.org/10.1016/j.heliyon.2023.e12888>
- Al-Hameed, N. A., Al-Mashaqbeh, M. A., Al-Ansari, M. A., & Al-Obaidi, M. A. (2023). Response surface methodology, and artificial neural network model for removal of textile dye reactive yellow 105 from wastewater using zeolitic materials. *Environmental Progress & Sustainable Energy*, 42, e13478. <https://doi.org/10.1002/ep.13478>
- Cojocaru, C., Samoila, P., & Pascariu, P. (2019). Chitosan-based magnetic adsorbent for removal of water-soluble anionic dye: Artificial neural network modeling and molecular docking insights. *International Journal of Biological Macromolecules*, 123, 587–599. <https://doi.org/10.1016/j.ijbiomac.2018.11.080>
- Crini, G., & Badot, P. M. (2008). Application of chitosan, a natural aminopolysaccharide, for dye removal from aqueous solutions by adsorption processes using batch studies: A review of recent literature. *Progress in Polymer Science*, 33(4), 399–447. <https://doi.org/10.1016/j.progpolymsci.2007.11.001>
- da Costa, M. F. P., Araújo, R. S., Silva, A. R., Pereira, L., & Silva, G. M. M. (2025). Predictive artificial neural networks as applied tools in the remediation of dyes by adsorption—A review. *Applied Sciences*, 15, 2310. <https://doi.org/10.3390/app15042310>
- El Bardiji, N., Ziat, K., Naji, A., & Saidi, M. (2020). Fractal-like kinetics of adsorption applied to the solid/solution interface. *ACS Omega*, 5(10), 5105–5115. <https://doi.org/10.1021/acsomega.9b04088>
- Gadekar, M. R., & Ahammed, M. M. (2019). Modelling dye removal by adsorption onto water treatment residuals using combined response surface methodology-artificial neural network approach. *Journal of Environmental Management*, 231, 241–248. <https://doi.org/10.1016/j.jenvman.2018.10.017>

Gerente, C., Lee, V. K. C., Le Cloirec, P., & McKay, G. (2007). Application of chitosan for the removal of metals from wastewaters by adsorption—Mechanisms and models review. *Critical Reviews in Environmental Science and Technology*, 37(1), 41–127. <https://doi.org/10.1080/10643380600729089>

Ghaedi, A. M., & Vafaei, A. (2017). Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review. *Advances in Colloid and Interface Science*, 245, 20–39. <https://doi.org/10.1016/j.cis.2017.04.015>

Gopinath, A., Aravamudan, K., Muthukkumaran, A., & Ganesan, R. B. (2019). Swift, versatile and a rigorous kinetic model based artificial neural network surrogate for single and multicomponent batch adsorption processes. *Journal of Molecular Liquids*, 111888. <https://doi.org/10.1016/j.molliq.2019.111888>

Guo, X., & Wang, J. (2019). A general kinetic model for adsorption: Theoretical analysis and modeling. *Journal of Molecular Liquids*, 288, 111100. <https://doi.org/10.1016/j.molliq.2019.111100>

Haykin, S. S. (2009). *Neural networks and learning machines* (3rd ed.). Prentice Hall.

Ho, Y. S., & McKay, G. (1999). Pseudo-second order model for sorption processes. *Process Biochemistry*, 34(5), 451–465. [https://doi.org/10.1016/S0032-9592\(98\)00112-5](https://doi.org/10.1016/S0032-9592(98)00112-5)

Iftikhar, S., Zahra, N., Rubab, F., Sumra, R. A., Khan, M. B., Abbas, A., & Jaffari, Z. H. (2023). Artificial neural networks for insights into adsorption capacity of industrial dyes using carbon-based materials. *Separation and Purification Technology*, 326, 124891. <https://doi.org/10.1016/j.seppur.2023.124891>

Igwegbe, C. A., Mohammadi, L., Ahmadi, S., Rahdar, A., Khadkhodaiy, D., Dehghani, R., & Rahdar, S. (2019). Modeling of adsorption of methylene blue dye on Ho-CaWO₄ nanoparticles using response surface methodology (RSM) and artificial neural network (ANN) techniques. *MethodsX*, 6, 1779–1797. <https://doi.org/10.1016/j.mex.2019.07.016>

Karam, A., Zaher, K., & Mahmoud, A. S. (2020). Comparative studies of using nano zerovalent iron, activated carbon, and green synthesized nano zerovalent iron for textile wastewater color removal using artificial intelligence, regression analysis, adsorption isotherm, and kinetic studies. *Air, Soil and Water Research*, 13, 1178622120908273. <https://doi.org/10.1177/1178622120908273>

Katheresan, V., Kansedo, J., & Lau, S. Y. (2018). Efficiency of various recent wastewater dye removal methods: A review. *Journal of Environmental Chemical Engineering*, 6, 4676–4697. <https://doi.org/10.1016/j.jece.2018.06.060>

Kumar, M., Tripathi, B. P., & Shahi, V. K. (2009). Crosslinked chitosan/polyvinyl alcohol blend beads for removal and recovery of Cd(II) from wastewater. *Journal of Hazardous Materials*, 172(2), 1041–1048. <https://doi.org/10.1016/j.jhazmat.2009.07.108>

Li, D., Huang, F., Yan, L., Cao, Z., Chen, J., & Ye, Z. (2019). Landslide susceptibility prediction using particle-swarm-optimized multilayer perceptron: Comparisons with multilayer-perceptron-only, BP neural network, and information value models. *Applied Sciences*, 9(18), 3664. <https://doi.org/10.3390/app9183664>

Pacheco-Sánchez, J. H., Alejo, R., Cruz-Reyes, H., & Álvarez-Ramírez, F. (2019). Neural networks to fit potential energy curves from asphaltene-asphaltene interaction data. *Fuel*, 236, 1117–1127. <https://doi.org/10.1016/j.fuel.2018.09.031>

Pauletto, P. S., Dotto, G. L., & Salau, N. P. (2020). Optimal artificial neural network design for simultaneous modeling of multicomponent adsorption. *Journal of Molecular Liquids*, 320, 114418. <https://doi.org/10.1016/j.molliq.2020.114418>

Sahnoun, S., & Boutahala, M. (2018). Adsorption removal of tartrazine by chitosan/polyaniline composite: Kinetics and equilibrium studies. *International Journal of Biological Macromolecules*, 114, 1345–1353. <https://doi.org/10.1016/j.ijbiomac.2018.02.146>

Tanzifi, M., Hosseini, S. H., Kiadehi, A. D., Olazar, M., Karimipour, K., Rezaiemehr, R., & Ali, I. (2017). Artificial neural network optimization for methyl orange adsorption onto polyaniline nano-adsorbent: Kinetic, isotherm and thermodynamic studies. *Journal of Molecular Liquids*, 244, 189–200. <https://doi.org/10.1016/j.molliq.2017.08.122>

Uddin Ahmad, Z., Yao, L., Wang, J., Gang, D., Islam, F., Lian, Q., & Zappi, M. (2018). Neodymium embedded ordered mesoporous carbon (OMC) for enhanced adsorption of sunset yellow: Characterizations, adsorption study and adsorption mechanism. *Chemical Engineering Journal*, 359, 814–826. <https://doi.org/10.1016/j.cej.2018.11.174>

Ullah, S., Assiri, M. A., Bustam, M. A., Al-Sehemi, A. G., Abdul Kareem, F. A., & Irfan, A. (2020). Equilibrium, kinetics and artificial intelligence characteristic analysis for Zn(II) ion adsorption on rice husks digested with nitric acid. *Paddy and Water Environment*, 18(2), 455–468. <https://doi.org/10.1007/s10333-020-00794-8>

Vargas, A. M. M., Cazetta, A. L., Martins, A. C., Moraes, J. C. G., Garcia, E. E., Gauze, G. F., Costa, W. F., & Almeida, V. C. (2012). Kinetic and equilibrium studies: Adsorption of food dyes Acid Yellow 6, Acid Yellow 23, and Acid Red 18 on activated carbon from flamboyant pods. *Chemical Engineering Journal*, 181–182, 243–250. <https://doi.org/10.1016/j.cej.2011.11.073>

Wong, Y. J., Arumugasamy, S. K., Chung, C. H., Selvarajoo, A., & Sethu, V. (2020). Comparative study of artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and multiple linear regression (MLR) for modeling of Cu(II) adsorption from aqueous solution using biochar derived from rambutan (*Nephelium lappaceum*) peel. *Environmental Monitoring and Assessment*, 192, 439. <https://doi.org/10.1007/s10661-020-08268-4>

Ahmed, M., Mavukkandy, M. O., Giwa, A., Elektorowicz, M., Katsou, E., Khelifi, O., Naddeo, V., & Hasan, S. W. (2022). Recent developments in hazardous pollutants removal from wastewater and water reuse within a circular economy. *NPJ Clean Water*, 5, 1–25. <https://doi.org/10.1038/s41545-022-00154-5>

Domínguez Beltrán, S., Miranda Piña, G., Granda Gutiérrez, E. E., Alejo Eleuterio, R., García Rivas, J. L., & Reyes García, A. (2025). ANN-based prediction of tartrazine adsorption on chitosan–polyvinyl alcohol hydrogel beads: A comparison with kinetic models. *Modelling*, 6, 149. <https://doi.org/10.3390/modelling6040149>

Dutta, S., Gupta, B., Srivastava, S. K., & Gupta, A. K. (2021). Recent advances on the removal of dyes from wastewater using various adsorbents: A critical review. *Materials Advances*, 2, 4497–4531. <https://doi.org/10.1039/D1MA00354B>

Ekman, M. (2022). *Learning deep learning*. Addison-Wesley Professional.

Fang, D., Zhuang, X., Huang, L., Zhang, Q., Shen, Q., Jiang, L., Xu, X., & Ji, F. (2020). Developing the new kinetics model based on the adsorption process: From fitting to comparison and prediction. *Science of the Total Environment*, 725, 138490. <https://doi.org/10.1016/j.scitotenv.2020.138490>

Hubbe, M. A., Azizian, S., & Douven, S. (2019). Implications of apparent pseudo-second-order adsorption kinetics onto cellulosic materials: A review. *BioResources*, 14(3), 7582–7626. <https://doi.org/10.15376/biores.14.3.Hubbe>

Kim, Y. S., Kim, M. K., Fu, N., Liu, J., Wang, J., & Srebric, J. (2025). Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models. *Sustainable Cities and Society*, 118, 105570. <https://doi.org/10.1016/j.scs.2024.105570>

Li, Z., Qin, R., Xue, J., Lin, C., & Jiang, L. (2025). Chitosan-based hydrogel beads: Developments, applications, and challenges. *Polymers*, 17, 920. <https://doi.org/10.3390/polym17070920>

Ngwenyama, M. K., & Gitau, M. N. (2024). Application of back propagation neural network in complex diagnostics and forecasting loss of life of cellulose paper insulation in oil-immersed transformers. *Scientific Reports*, 14, 6080. <https://doi.org/10.1038/s41598-024-56956-z>

Rápó, E., & Tonk, S. (2021). Factors affecting synthetic dye adsorption; desorption studies: A review of results from the last five years (2017–2021). *Molecules*, 26, 5419. <https://doi.org/10.3390/molecules26175419>

Saheed, I. O., & Suah, F. B. M. (2023). Developing nano-micro size chitosan beads using imidazolium-based ionic liquid: A perspective. *International Journal of Biological Macromolecules*, 241, 124610. <https://doi.org/10.1016/j.ijbiomac.2023.124610>

Salahshoori, I., Wang, Q., Nobre, M. A., Mohammadi, A. H., Dawi, E. A., & Khonakdar, H. A. (2024). Molecular simulation-based insights into dye pollutant adsorption: A perspective review. *Advances in Colloid and Interface Science*, 333, 103281. <https://doi.org/10.1016/j.cis.2024.103281>

Vasiliauskaite, V., & Antulov-Fantulin, N. (2024). Generalization of neural network models for complex network dynamics. *Communications Physics*, 7, 348. <https://doi.org/10.1038/s42005-024-01839-w>