UNIVERSIDAD
NACIONAL
DE COLOMBIA

INTEGRACIÓN DE LOS PARÁMETROS FISICOQUÍMICOS DE NANOFLUIDOS POLIMÉRICOS BASADOS EN PUNTOS CUÁNTICOS DE CARBONO (CQDS) PARA LA PREDICCIÓN DE VISCOSIDAD UTILIZANDO MODELOS DE APRENDIZAJE DE MÁQUINAS

As Augusto Ríos Muñoz

Facultad de Minas

Universidad Nacional de Colombia

Medellín, Colombia

21 de febrero de 2025

INTEGRACIÓN DE LOS PARÁMETROS FISICOQUÍMICOS DE NANOFLUIDOS POLIMÉRICOS BASADOS EN PUNTOS CUÁNTICOS DE CARBONO (CQDS) PARA LA PREDICCIÓN DE VISCOSIDAD UTILIZANDO MODELOS DE APRENDIZAJE DE MÁQUINAS

As Augusto Ríos Muñoz

Tesis de maestría presentada como requisito parcial para optar al título de:

Magister en Ingeniería - Analítica

Directores

John Willian Branch Bedoya, Ph.D
Farid Bernardo Cortés Correa, Ph.D
Camilo Andrés Franco Ariza, Ph.D

Facultad de Minas

Universidad Nacional de Colombia

Medellín, Colombia

2024

A mi familia, amigos y compañeros de viaje, su paciencia, entrega y apoyo incondicional fueron el motor de este trabajo de investigación. De ustedes es este logro.

Agradecimientos

A los profesores John Willian Branch, Farid Bernardo Cortés y Camilo Andrés Franco quienes dirigieron, acompañaron y sustentaron el desarrollo de esta tesis de maestría. Agradezco también a todos y cada uno de los miembros del Grupo de Investigación en Fenómenos de Superficie – Michael Polanyi y el Grupo de Investigación y Desarrollo en Inteligencia Artificial (GIDIA), sus retroalimentaciones y comentarios permitieron construir este resultado.

Este estudio de maestría fue posible gracias a la financiación del Fondo Francisco José de Caldas, MINCIENCIAS y la Agencia Nacional de Hidrocarburos (ANH) a través del contrato No. 112721-282-2023 (Proyecto 1118-1035-9300) con la Universidad Nacional de Colombia – Sede Medellín y PAREX RESOURCES COLOMBIA AG SUCURSAL.

Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, excepto donde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando se han presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones de texto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.

As Augusto Rios Muñoz
Nombre

Fecha 22/10/2024

Resumen

Integración de los parámetros fisicoquímicos de nanofluidos poliméricos basados en puntos cuánticos de carbono (CQDs) para la predicción de viscosidad utilizando modelos de aprendizaje de máquinas

Los procesos de recobro mejorado en Colombia toman relevancia para mantener la soberanía energética del país a través de la implementación de nuevas tecnologías. Este trabajo permitió identificar las variables relevantes para la evaluación de nanofluidos poliméricos basados en puntos cuánticos de carbono (CQDs) y poliacrilamida parcialmente hidrolizada (HPAM) para implementar modelos de aprendizaje de máquinas que logren predecir la viscosidad de la solución resultante como estrategia de monitoreo. El conjunto de datos fue generado experimentalmente bajo diferentes condiciones de salinidad, tiempo de añejamiento, orden de adición, tipo de CQD y tasa de corte como predictoras de viscosidad, manteniendo la temperatura, concentración de CQD y polímero fijas. La implementación de cuatro (4) modelos de aprendizaje supervisado de regresión, permitió establecer la capacidad predictiva de los modelos ante un conjunto de datos altamente desbalanceado y de baja correlación lineal entre las variables siendo el de bosques aleatorios el modelo con mejor desempeño con un R^2 cercano al 87%. Los resultados experimentales permitieron establecer nuevos acercamientos para investigar el efecto sinérgico CQD-HPAM en procesos de inyección de polímeros para recobro mejorado utilizando modelos de aprendizaje de máquinas como una estrategia efectiva para generalizar el comportamiento de estos sistemas físicos.

Palabras clave: Recobro mejorado, nanofluidos poliméricos, predicción de viscosidad, aprendizaje de máquinas, puntos cuánticos de carbono.

Abstract

Integration of physicochemical parameters of polymeric nanofluids based on carbon quantum dots (CQDs) for viscosity prediction using machine learning algorithms.

Enhanced recovery processes in Colombia are relevant in maintaining the country's energy sovereignty through implementing innovative technologies. This work allowed us to identify the relevant variables for the evaluation of polymeric nanofluids based on carbon quantum dots (CQDs) and partially hydrolyzed polyacrylamide (HPAM) to implement machine learning models that can predict the viscosity of the resulting solution as a monitoring strategy. The data set was generated experimentally under different salinity conditions, aging time, addition order, CQD type, and shear rate as viscosity predictors, keeping the temperature, polymer, and CQD concentration fixed. The implementation of 4 regression models allowed the establishment of the predictive capacity of the methods in the face of a highly unbalanced dataset and low linear correlation between the variables, with the random forest model being the algorithm with the best performance with an R^2 near 87%. The experiment highlighted novel approaches to investigate the CQD-HPAM synergistic effect in polymer injection processes for enhanced oil recovery using machine learning techniques as an effective tool to generalize the behavior of these physical systems.

Keywords: Enhanced oil recovery, polymer nanofluids, viscosity prediction, machine learning, carbon quantum dots.

Contenido

Pág.

Resumen ..... **XI**

Lista de figuras ..... **XV**

Lista de tablas ..... **XVI**

Abreviaturas ..... **XVII**

1. Introducción **19**
1.1 Motivación 19
1.2 Trabajos previos 20
1.3 Planteamiento del problema 21
1.4 Objetivos 22
1.4.1 Objetivo general 22
1.4.2 Objetivos específicos 22
1.5 Contribución 22
1.6 Estructura del documento 23

2. Marco teórico de referencia **24**
2.1 Fenómenos asociados a la viscosidad de nanofluidos poliméricos y la inclusión de CQDs para recobro mejorado con polímeros 24
2.2 Modelos de aprendizaje de máquinas seleccionados para la predicción de viscosidad de nanofluidos poliméricos 27
2.2.1 Regresión lineal 28
2.2.2 Máquinas de soporte vectorial 28
2.2.3 Árboles de decisión 28
2.2.4 Bosques aleatorios 29

3. Revisión sistemática de literatura **30**
3.1 Introducción 30
3.2 Metodología 30
3.3 Resultados y discusión 31
3.3.1 Efectos de los parámetros fisicoquímicos en la conformación de nanofluidos poliméricos y la inclusión de CQDs 32
3.3.2 Implementación de modelos de aprendizaje de máquinas para la predicción de propiedades físicas de soluciones poliméricas 34

4. Metodología para la predicción de viscosidad de nanofluidos poliméricos utilizando modelos de aprendizaje de máquinas **39**

4.1 Metodología experimental para la caracterización y medición de viscosidad de nanofluidos poliméricos 39
4.2 Conformación y análisis descriptivo del conjunto de datos experimentales. 41
4.3 Modelos de aprendizaje de máquinas para la predicción de viscosidad de nanofluidos poliméricos 42

5. Resultados 45
5.1 Caracterización y evaluaciones estáticas del nanofluido polimérico 45
5.2 Análisis descriptivo del conjunto de datos generado 50
5.3 Desempeño de los modelos de aprendizaje de máquinas implementados 51

6. Conclusiones y trabajo futuro 57
6.1 Conclusiones 57
6.2 Trabajo futuro 58

7. Referencias 60

Lista de figuras

Pág.

Figura 2-1. Mecanismo de interacción entre los iones divalentes y el HPAM. Adaptado de [31] 25
Figura 2-2. Efecto cuántico de las nanopartículas. Adaptado de Johan Jarnestad/Real Academia Sueca de Ciencias [38] 26
Figura 2-3. Representación gráfica de los modelos de aprendizaje supervisado 27
Figura 3-1. Publicación de artículos por año para cada apartado. 32
Figura 4-1. Concentración de iones en las salmueras implementadas 40
Figura 5-1. Análisis FTIR para los CQDs utilizados 45
Figura 5-2. Análisis FTIR para el HPAM utilizado. 46
Figura 5-3. Viscosidad de nanofluidos poliméricos en el tiempo inicial y 7 s^{-1} para A) CQDred B) CQDblue y C) CQDgreen 48
Figura 5-4. Viscosidad de nanofluidos poliméricos en 7 s^{-1}, diferentes salinidades y tiempos de evaluación para A) Sin CQDs, B) CQDred, C) CQDblue y D) CQDgreen 49
Figura 5-5. Distribución del conjunto de datos generado 50
Figura 5-6. Matriz de correlación de Pearson para las variables continuas 51
Figura 5-7. Métricas de error obtenidas en la evaluación de los modelos 52
Figura 5-8. Comparación de valores reales y predichos para los modelos implementados 53
Figura 5-9. Valores residuales para los modelos implementados 54

Lista de tablas

Pág.

Tabla 3-1. Resultados de la revisión sistemática de literatura.....31
Tabla 3-2. Métricas de desempeño para predicción virtual de viscosidad. Tomado de [20], adaptado por el autor.....36
Tabla 3-3. Conclusiones de la revisión sistemática de literatura para la implementación de modelos de aprendizaje de máquinas en la predicción de propiedades de nanofluidos poliméricos.....38
Tabla 4-1. Parámetros y rangos de variación para el conjunto de datos generado. 41
Tabla 4-2. Hiperparámetros y rangos de variación para los métodos de predicción 43
Tabla 4-3. Métricas implementadas para los modelos de regresión.....44
Tabla 5-1. Tamaño y potencial Z para los CQDs utilizados.....46
Tabla 5-2. Análisis descriptivo de las viscosidades en el conjunto de datos 50
Tabla 5-3. Hiperparámetros resultantes de la sintonización de los modelos 56

Abreviaturas

Abreviatura	Término
HPAM	Poliacrilamida Parcialmente Hidrolizada
CQDs	Puntos Cuánticos de Carbono
EOR	Recobro mejorado de petróleo
RL	Regresión Lineal
RP	Regresión Polinomial
MLR	Regresión Lineal Múltiple
SVM	Máquinas de Soporte Vectorial
DTR	Arboles de Decisión
RF	Bosques Aleatorios

1.Introducción

1.1 Motivación

A pesar de contar con una matriz de generación eléctrica compuesta un 68% por renovables [1], Colombia presenta una matriz energética general donde el 76% se compone por hidrocarburos [2] con una demanda estimada por combustibles incrementando en todos los escenarios proyectados entre un 1 y 5% hasta el 2036 [3]. Esto ubica al sector energético en una posición clave para mantener el desarrollo de la nación desde el suministro constante y sostenible de energía. Sin embargo, el factor de recobro nacional se ubica cerca al 20.1%, muy por debajo del valor mundial de 38%. Dentro de las cifras de producción nacional se establece que el 36% de la producción corresponde a recuperación secundaria y recuperación mejorada de petróleo (EOR) concentrados en 33 campos de explotación [4]. De hecho, se estima que un incremento del 1% en el factor de recobro representaría el ingreso de 650 millones de barriles a la base de reservas probadas [5].

Las técnicas de recobro mejorado (EOR) comprenden alternativas ampliamente desarrolladas, sin embargo, su efectividad responde a parámetros propios de cada escenario de implementación [6]. Para escoger las técnicas de recobro mejorado se llevan metodologías de evaluación robustas (conocidas como *screening*) que permiten establecer las técnicas adecuadas en función de las características geofísicas del yacimiento, los fluidos que lo componen y las facilidades para la implementación [6]. La inyección de polímeros para recobro mejorado ha representado una alternativa eficiente para incrementar la cantidad de aceite en el país mostrando una eficiencia entre 0.02 y 1.51 barriles adicionales producidos por kilogramo de HPAM con 17 patrones de inyección entre 2011 y 2023 en Colombia [7].

Paralelo a estas alternativas se ejecutan procesos de inyección de trazadores con el fin de mejorar el entendimiento de las características del medio poroso, impacto de las tecnologías de recuperación y nuevos esquemas de tecnologías enfocadas a mejorar la recuperación de petróleo [8]. Los trazadores pasivos son un tipo de trazador utilizado ampliamente en la industria del petróleo y el gas para establecer los patrones de flujo al interior del yacimiento estableciendo la dirección del fluido trazador siendo este principalmente salmueras de inyección [8]. La inyección de CQDs sobresale como una alternativa eficiente de trazador pasivo en el yacimiento de menor costo para establecer los patrones de flujo en el medio poroso con altas capacidades para trazar [9].

A pesar de su relevancia, los efectos de los CQDs y su naturaleza en los nanofluidos poliméricos aún permanece bajo análisis exhaustivo dado que las propiedades resultan sensibles a cambios en salinidad, condiciones externas como nanopartículas y agentes químicos reduciendo su efectividad en el proceso [10-13]. Es por esto qué la predicción y seguimiento de la viscosidad de los nanofluidos poliméricos representa una herramienta importante para mantener los estándares de calidad en la inyección, sin embargo, dada su alta sensibilidad a parámetros externos, variaciones ligeras podrían generar grandes impactos en los valores de viscosidad fuera de las requeridas. Las técnicas de aprendizaje de máquinas nacen como una respuesta a problemas de alta complejidad dimensional y de modelamiento, entendimiento teórico limitado y la posibilidad de ser solucionadas a través de la información contenida en los datos medidos. En este trabajo se busca implementar modelos de aprendizaje de máquinas que permitan predecir la viscosidad de nanofluidos poliméricos ante diferentes escenarios de salinidad, tiempo, presencia y tipo de CQDs con un conjunto de datos generado experimentalmente como alternativa suplementaria para valorar esta técnica de recobro ante diferentes escenarios.

1.2 Trabajos previos

La revisión de literatura permitió identificar los trabajos de investigación realizados previamente donde se exponen algunos avances en EOR con polímeros, la inclusión de CQDs en la solución polimérica y la implementación de modelos de aprendizaje de máquinas para establecer las propiedades de flujo del polímero. Si bien los impactos de los CQDs en el HPAM comienzan a ser explicados, la fenomenología asociada al proceso aún se mantiene fuera del alcance del modelamiento bajo un escenario de aplicación real considerando que los impactos en la viscosidad dependen de la naturaleza del CQD y sus propiedades físicas.

La inclusión de CQDs en las soluciones poliméricas ha sido estudiada previamente por algunos autores donde se indica un impacto negativo en la viscosidad [14, 15], sin embargo, cuando se implementan puntos cuánticos de sílice, la viscosidad de las soluciones poliméricas incrementa cerca de un 39% con respecto al escenario base y se presenta un incremento en el recobro final de aceite del 28.3% [16]. Estos estudios confirman el efecto de la naturaleza del nanomaterial en la viscosidad del nanofluido polimérico, sin embargo, demuestra la necesidad de profundizar en la fenomenología asociada al proceso para cuantificar los impactos en implementaciones reales de proyectos de EOR. Estos trabajos permiten identificar que la naturaleza del CQD impacta la viscosidad de las soluciones poliméricas en diferentes medidas dificultando la predicción y modelamiento de las propiedades de flujo.

Yao *et al* [14] utilizaron simulación molecular para establecer la interacción entre el HPAM y CQDs en presencia de iones monovalentes de Na^+ indicando una irrupción de las cadenas poliméricas cuando el polímero se absorbe sobre la superficie del CQD por las zonas donde se distribuyen los grupos funcionales amida principalmente más que por los grupos funcionales carboxilados. Haruna *et al* [15] demostraron que el uso de composites

HPAM/CQDs tienen una viscosidad inferior al sistema base de referencia generando un sistema desordenado en la solución polimérica. Los impactos de los CQDs en las soluciones poliméricas muestran una alta afinidad de los nanomateriales por las cadenas poliméricas alterando la estabilidad de la solución reduciendo la viscosidad.

Por otro lado, la implementación de modelos de aprendizaje de máquinas para la predicción de viscosidad de fluidos poliméricos se viene desarrollando ampliamente para su aplicación en modelos asistidos [17-19] o la predicción directa de la viscosidad [20-23]. Shakeel *et al* [23] utilizaron modelos de regresión múltiple, máquinas de soporte vectorial, árboles de decisión y redes neuronales. Este trabajo permitió identificar que, para este tipo de aplicaciones, los modelos lineales no muestran un buen desempeño y requieren de sintonización específica para cada tipo de polímero.

Los trabajos revisados apuntan a la evaluación de parámetros en la solución y su respectivo impacto en la predicción de la viscosidad de nanofluidos poliméricos. Si bien la inclusión de CQDs en la solución polimérica comienza a ser relevante, ningún trabajo previo presenta la inclusión de este nanomaterial en los valores de viscosidad para predecir los respectivos valores utilizando modelos de aprendizaje de máquinas.

1.3 Planteamiento del problema

En procesos de EOR con polímero se implementa ampliamente el HPAM por su bajo costo y propiedades viscosificantes del fluido desplazante [24]. A pesar de sus propiedades y amplia implementación en campo el polímero es sensible a condiciones diversas de salinidad, temperatura y agentes externos afectando las propiedades físicas del polímero en solución. Paralelo a la inyección de polímeros las alternativas para trazar fluidos en el yacimiento se utilizan para profundizar el conocimiento de las características del medio poroso, así como el entendimiento de los efectos de los trazadores en los nanofluidos poliméricos también resulta clave para estimar el impacto de las alternativas combinadas en esquemas de inyección. Si bien las valoraciones experimentales son requeridas para estimar la efectividad de las estrategias de recobro para las condiciones específicas de despliegue, estas pueden ser costosas y su potencialidad para generalización se ven limitadas a escenarios cubiertos por los rangos de evaluación. Las técnicas de aprendizaje de máquinas resultan una herramienta de alto valor para predecir valores en escenarios de bajo relacionamiento donde los modelos típicos no alcanzan a incluir el efecto de las variaciones como la inclusión de CQDs o el tiempo de añejamiento evaluando modelos con capacidad de generalización que permitan establecer las propiedades de flujo de los polímeros ante cambios en la corriente de inyección.

1.4 Objetivos

1.4.1 Objetivo general

Implementar modelos de predicción para establecer la viscosidad de polímeros para EOR en presencia de CQDs y diferentes escenarios de salinidad, tiempo de añejamiento y orden de adición.

1.4.2 Objetivos específicos

- Valorar el efecto de la estrategia de conformación del polímero, salinidad y tasa de corte en las propiedades de flujo de la solución conformada.
- Construir un conjunto de datos representativo a partir de las variables de interés manteniendo la representatividad física del fenómeno y las condiciones de evaluación fijadas.
- Implementar y evaluar modelos de aprendizaje de máquinas supervisados ´por regresión, para la predicción de la viscosidad del polímero a partir de variables preestablecidas según el estado del arte y los resultados experimentales.
- Comparar los modelos de aprendizaje de máquinas supervisados por regresión implementados, para establecer la relación entre los métodos y la naturaleza física del problema.

1.5 Contribución

Los resultados de esta investigación pueden agruparse en dos (2) grupos principales asociados a la interacción y efecto de los CQDs en las propiedades de flujo del polímero y la predicción de la viscosidad utilizando métodos de aprendizaje de máquinas a partir de un conjunto de datos experimentales generado bajo diferentes condiciones fisicoquímicas.

- La evaluación experimental de los nanofluidos poliméricos bajo diferentes condiciones de conformación y parámetros fisicoquímicos permite establecer diferencias claras con el orden de adición, salinidad y tiempo de añejamiento de las soluciones poliméricas en presencia de CQDs para un proceso de recobro con polímero e inyección de trazadores al yacimiento.
- La implementación de los modelos de aprendizaje de máquinas supervisados por regresión permitió seleccionar los más eficientes para este tipo de aplicaciones donde se presenta una baja cantidad de datos, variaciones fenomenológicas en los

nanofluidos y un alto desbalanceo dadas las propiedades del fenómeno. Este estudio presenta un primer acercamiento como estrategia complementaria en los procesos de valoración de las técnicas de EOR y la posibilidad de extrapolar resultados experimentales utilizando estas técnicas automáticas.

Los resultados asociados a la etapa experimental inicial fueron presentados en dos eventos de difusión científica:

- X SUMMER SCHOOL: realizado en la Facultad de Minas los días 23 y 24 de abril del 2024 en modalidad de poster en tres (3) minutos obteniendo el primer lugar con el trabajo: **"Evaluation of addition of Carbon Quantum Dots (CQDs) on Polymeric solution for EOR processes: Effect of CQD type and Ion content on brine"** [25].
- XXI CONGRESO DEL PETROLEO, EL GAS Y LA ENERGÍA: realizado en la ciudad de Barranquilla del 28 al 30 de agosto en modalidad de poster con el trabajo titulado: **"EVALUACIÓN DE LA ADICIÓN DE PUNTOS CUANTICOS DE CARBONO (CQDs) EN SOLUCIONES POLIMERICAS PARA PROCESOS DE RECOBRO MEJORADO (EOR): EFECTO DEL TIPO DE CQD Y CONTENIDO DE IONES EN SALMUERA"** [26].

1.6 Estructura del documento

El documento cuenta con cinco (5) capitulos aparte de la introducción y las referencias. En el capítulo dos (2) se presenta el marco teórico de referencia relacionado con procesos de EOR con polímeros, la inclusión de trazadores y modelos de aprendizaje de máquinas utilizados en este trabajo. La revisión sistemática de literatura se presenta en el capítulo tres (3) mostrando el estado actual de los avances tecnológicos en materia de inclusión de CQDs en soluciones poliméricas a base de HPAM y la implementación de modelos de aprendizaje de máquinas para la predicción de viscosidad. En el capítulo cuatro (4) se desarrollan las técnicas experimentales para ejecutar el trabajo de investigación, el análisis descriptivo del conjunto de datos, los modelos y métricas de error implementadas. El capítulo cinco (5) expone los resultados obtenidos para finalmente presentar las conclusiones y trabajos futuros en el capítulo seis (6).

2. Marco teórico de referencia

En este capítulo se expone el marco teórico de referencia asociado a los mecanismos de interacción entre los CQDs y soluciones poliméricas integrando parámetros como la salinidad, temperatura, tiempo de añejamiento y tasa de corte. La descripción fenomenológica de este sistema pretende fundamentar los mecanismos representados a través de los datos manteniendo la relación física entre los resultados de los modelos de aprendizaje de máquinas. Después, se expone brevemente el funcionamiento interno de los modelos de aprendizaje de máquinas seleccionados implementados en este trabajo.

2.1 Fenómenos asociados a la viscosidad de nanofluidos poliméricos y la inclusión de CQDs para recobro mejorado con polímeros

En las aplicaciones generales de EOR con polímeros sobresale la implementación de la poliacrilamida parcialmente hidrolizada [10] (HPAM por sus siglas en inglés) por su bajo costo [13] y capacidades viscosificantes del fluido desplazante, sin embargo, es particularmente sensible a hidrolisis a alta temperatura, estrés mecánico y presencia de iones [12]. Algunos autores como Gallego *et al* [27] estudiaron el efecto de la presencia de iones y el grado de hidrolisis con la temperatura cuando se presenta un escenario de iones monovalentes, divalentes y mezcla compleja mostrando que, las cadenas aniónicas del HPAM presentan una alta afinidad por los iones divalentes interrumpiendo las cadenas poliméricas al unirse al extremo de la cadena donde se presenta la carga negativa [11, 27]. Cuando los iones divalentes entran en contacto con las cadenas poliméricas se cataliza el proceso de hidrolisis del HPAM como se presenta en la **Figura 2-1** [28] reduciendo la estabilidad y propiedades mecánicas de la solución e incrementando la cantidad de polímero adsorbido por el medio poroso [29]. Si bien las salmueras de formación presentan combinaciones de iones muy variadas [30] resulta importante evaluar el efecto del tipo de ion y a cantidad presente en solución [11, 12, 27].

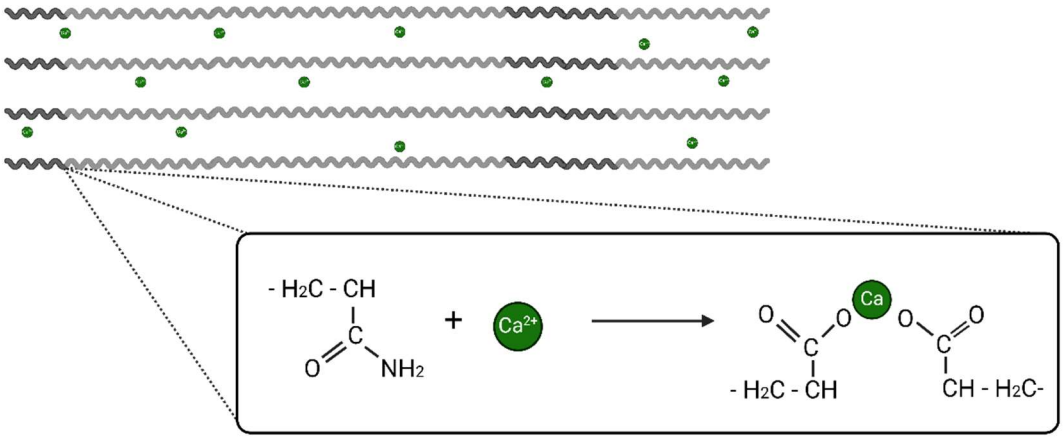

Figura 2-1. Mecanismo de interacción entre los iones divalentes y el HPAM. Adaptado de [31]

Adicional a la presencia y tipo de iones, la temperatura de prueba y añejamiento resulta importante para establecer el grado de hidrolisis del polímero a condiciones relevantes del yacimiento y así la viscosidad del polímero a diferentes regímenes de flujo, de hecho, el HPAM comienza a hidrolizarse desde los 25°C presentando un amplio rango de temperaturas que afectan las propiedades mecánicas del fluido [11]. Para mejorar la resistencia térmica se han implementado nanopartículas como agente estabilizante reduciendo la cantidad de polímero adsorbido por el medio poroso irrumpiendo en las cadenas poliméricas en zonas donde la interacción con los iones puede facilitarse [32]

Dentro de los procesos de recuperación de petróleo la inyección de trazadores es un mecanismo importante para monitorear el yacimiento, profundizar la caracterización del medio poroso, saturaciones, patrones de flujo e interconexión [33]. Los trazadores pueden ser clasificados según su funcionamiento (pasivos o activos) o características fisicoquímicas (radioactivos, sólidos, gaseosos o líquidos) [34]. Los trazadores activos o de partición son conocidos por tener interacciones con el medio poroso y los fluidos del yacimiento, tardan más tiempo en presentar ruptura en el pozo de producción y son herramientas de alta importancia para determinar la saturación de aceite en el medio poroso posterior o durante procesos de recobro secundario o terciario [33-35]. Por otro lado, los trazadores pasivos buscan establecer los patrones de flujo en el medio poroso para determinar tiempos de ruptura y zonas de contacto previo durante o posterior a la implementación de técnicas de recobro para vigilar el flujo en el yacimiento sin migración entre las fases o el medio poroso donde típicamente se inyectan en la salmuera [33, 34].

En procesos de EOR con polímeros la inyección de trazadores resulta clave para hacer seguimiento de la técnica y el barrido del polímero en el yacimiento. Los trazadores se inyectan para obtener información de la dirección, grado de retención y dilución del polímero [8]. En aplicaciones reales, los trazadores se inyectan como un pulso antes y

después de la inyección del polímero o bajo inyección continua por lo que es necesaria la implementación de trazadores pasivos que minimicen la interacción trazador-polímero manteniendo las propiedades técnicas necesarias para un proceso de recobro y trazado bajo los diferentes esquemas implementados [8].

Como alternativa novedosa surgen los puntos cuánticos de carbono (CQDs) por ser amigables con el medio ambiente, de bajo costo y efectivos como trazadores pasivos de la salmuera de inyección en el medio poroso [9]. Los puntos cuánticos de carbono son materiales carbonosos de tamaño nanométrico, alta biocompatibilidad y propiedades de fluorescencia óptimas para su implementación como trazadores de fluidos en el yacimiento [9, 36, 37] aprovechando el efecto cuántico del nanomaterial carbonoso como se muestra en la **Figura 2-2** [38]. Los estudios realizados previamente donde se despliegan los CQDs en aplicaciones relevantes y pilotos de inyección muestran que este tipo de materiales no afectan las propiedades del yacimiento y los fluidos, manteniendo sus propiedades ópticas [9, 36]. Si bien el entendimiento del fenómeno CQDs-Polímero apenas comienza a ser expuesto por algunos autores, no se presenta un consenso claro sobre sus efectos en las cadenas poliméricas y las propiedades de flujo [14].

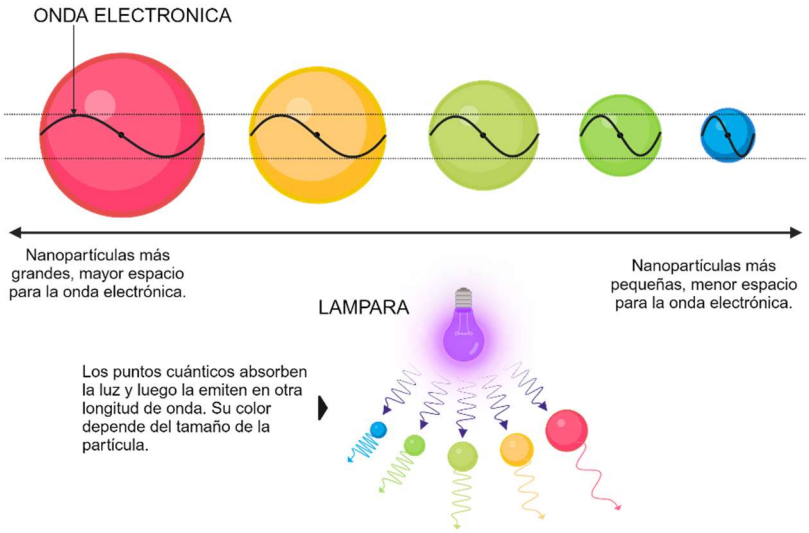

Figura 2-2. Efecto cuántico de las nanopartículas. Adaptado de Johan Jarnestad/Real Academia Sueca de Ciencias [38]

Dada su relevancia, los CQDs se muestran como una estrategia de alto valor para trazar los fluidos del yacimiento, sin embargo, es necesario establecer los impactos de su adición paralela en otras técnicas de recobro mejorado como polímeros, surfactantes y espumas [15, 39, 40]. La complejidad de las relaciones entre las variables permitió enfocar este estudio fijando la temperatura, concentración de CQDs, polímero y tipo de iones en la solución manteniendo como variables el tipo de CQD, orden de adición, tiempo de añejamiento, tasa de corte y salinidad para predecir la viscosidad de los nanofluidos poliméricos utilizando modelos de aprendizaje de máquinas.

2.2 Modelos de aprendizaje de máquinas seleccionados para la predicción de viscosidad de nanofluidos poliméricos

Considerando la naturaleza continua de la variable independiente (viscosidad) es necesario implementar modelos de aprendizaje de máquinas supervisados, específicamente, modelos de regresión. El aprendizaje de máquinas supervisado es una técnica que permite predecir la variable objetivo a partir de variables independientes también conocidas como variables explicativas, en este caso, las variables independientes del conjunto de datos corresponden a la salinidad, tiempo de añejamiento, tasa de corte, tipo de CQD y orden de adición a la solución polimérica cuyo objetivo de predicción es la viscosidad del nanofluido. Los modelos de aprendizaje de máquinas supervisados pueden dividirse en modelos de clasificación cuando se busca categorizar los datos o modelos de regresión cuando se busca la predicción de un valor continuo en función de las variables como se presenta en la **Figura 2-3**.

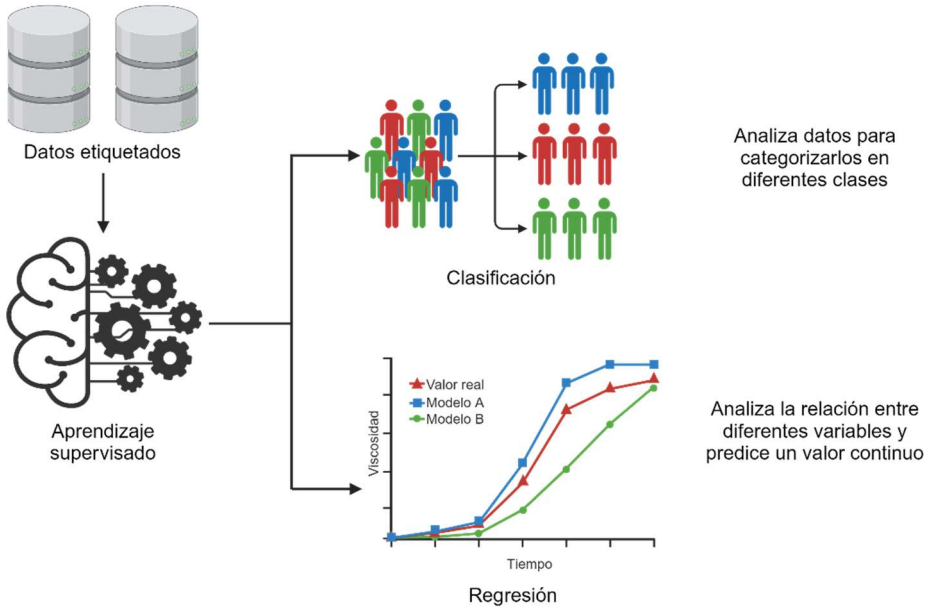

Figura 2-3. Representación gráfica de los modelos de aprendizaje supervisado

Fueron seleccionados los modelos de regresión lineal, máquinas de soporte vectorial, arboles de decisión y bosques aleatorios por su relevancia y aplicación exhibida en el capítulo 3. Si bien las redes neuronales corresponden a un modelo ampliamente utilizado, se pretende limitar la aplicación modelos de aprendizaje no profundos que no requieran de altos volúmenes de datos [41].

2.2.1 Regresión lineal

El modelo de regresión busca hallar dependencias entre las variables explicativas y la variable independiente utilizando funciones lineales. Para el escenario de múltiples variables explicativas se espera generar combinaciones lineales de los parámetros que permitan relacionar directamente cada parámetro como se presenta en la **Ecuación 2-1**.

Ecuación 2-1

$$y = \varepsilon_0 + \varepsilon_1 x_1 + \varepsilon_2 x_2 \dots . + \varepsilon_n x_n$$

Donde ε representa los coeficientes, x la variable explicativa y n el número de variables explicativas presentes en la regresión lineal. Si bien las regresiones lineales presentan un buen desempeño para algunos escenarios, es un modelo sensible a conjuntos altamente desbalanceados y distribuciones no normales [42].

2.2.2 Máquinas de soporte vectorial

Las máquinas de soporte vectorial surgen en los 90 como un desarrollo resultante de los avances en la teoría de aprendizaje estadístico de Vapnik [43, 44]. Bajo esta teoría, un de aprendizaje estadístico recibe observaciones de un sistema y entrega como salida una función que permite predecir valores futuros integrando dentro del modelo la capacidad de generalización a través de patrones regulares identificados del conjunto de datos con la habilidad para generalizar provenientes de patrones propios de los datos. [45].

La idea fundamental detrás de las máquinas de soporte vectorial se resume en incorporar los datos en un espacio de alta dimensionalidad donde es posible computar cálculos algebraicos para construir un hiperplano, maximizando los márgenes de separación y manteniendo la distancia de los puntos de entrenamiento con el hiperplano al mínimo [45]. La función de decisión implementada por las máquinas de soporte vectorial se presenta en la **Ecuación 2-2**.

Ecuación 2-2

$$f(x) = sign(\sum_{i}^{n} \alpha_i y_i k(x_i, x) + b)$$

Donde x_i, y_i, k son puntos de entrenamiento, las respectivas etiquetas y la función del kernel respectivamente con *x* como punto de evaluación arbitrario [45].

2.2.3 Árboles de decisión

Los árboles de decisión son modelos de aprendizaje de máquinas que generan reglas de decisión bajo estructuras similares a árboles generando nodos, ramas y particiones a partir de la distribución de los datos integrando criterios de parada o “podado” de los árboles que reducen la complejidad de la estructura interna [46]. El modelo genera regiones donde los

valores pueden distribuirse en función de los predictores de tal manera que se logre reducir la sumatoria de los cuadrados residuales descrito por la **Ecuación 2-3** [47].

Ecuación 2-3

$$RSS = \sum_{j=1}^{J} \sum_{i \in R} (y_i - \hat{y}_{R_j})^2$$

Los árboles de decisión pueden ser estructuras de alta complejidad y robustez, sin embargo, bajo criterios de parada se minimiza la complejidad del modelo entrenado minimizando el sobre entrenamiento y disminuye la capacidad de generalización [46, 47].

2.2.4 Bosques aleatorios

Bosques aleatorios es un modelo de ensamble propuesto por Breiman [48] que funciona construyendo arboles de manera aleatoria y con particiones del conjunto de datos para generar inicialmente, una predicción basada en arboles de decisión garantizando independencia entre los árboles y una baja correlación mutua [49]. Esta condición permite encontrar las características más importantes mientras se expanden los árboles incrementando la capacidad de generalización.

El modelo de bosques aleatorios combina secuencialmente modelos de remuestreo (bootstrapping), adición y combinación respectivamente [50]. En la primera etapa se generan particiones del conjunto de datos para mantener la independencia de cada árbol con las predicciones individuales. Posteriormente, se calculan los valores a partir de los conjuntos para promediar los valores resultantes en una predicción final. Dada las particiones aleatorias iniciales este método de aprendizaje de máquinas es eficaz disminuyendo el sobre entrenamiento y genera predicciones precisas incluso para conjuntos de datos con una alta cantidad de valores faltantes [49, 50].

3. Revisión sistemática de literatura

3.1 Introducción

La revisión sistemática de literatura fue llevada a cabo para establecer los avances y nuevas perspectivas enfocadas a la predicción de viscosidad de nanofluidos poliméricos en presencia de CQDs y diferentes parámetros fisicoquímicos de los nanofluidos. En este apartado se presentan los resultados de la revisión realizada en Scopus diferenciando la metodología de búsqueda, los resultados iniciales de la búsqueda y el desglose de los desarrollos relacionados con los temas abordados en este trabajo.

3.2 Metodología

Para el levantamiento de información relacionada con nanofluidos poliméricos, CQDs y machine learning fue utilizada la base de datos Scopus como principal fuente bibliográfica. Scopus es una base de datos multidisciplinaria de alto valor en el sector académico con herramientas de analítica de alta precisión. La base de datos permite realizar búsquedas utilizando operadores lógicos para establecer relaciones entre diferentes temas enriqueciendo así los resultados dentro de un trabajo multidisciplinario. En este caso, el uso de esta herramienta de analítica bibliográfica permite encontrar desarrollos en el área del recobro mejorado de hidrocarburos y el aprendizaje de máquinas con nanofluidos poliméricos.

Teniendo en cuenta el problema planteado se pueden establecer dos preguntas de investigación definidas como:

1. ¿Cuáles son los efectos de los parámetros fisicoquímicos en la conformación de nanofluidos poliméricos a base de HPAM y CQDs?
2. ¿Cuáles modelos de aprendizaje de máquinas se han implementado para predecir las propiedades de flujo de nanofluidos poliméricos y bajo que métricas fueron valorados?

Bajo estas preguntas de investigación, la revisión sistemática de literatura fue llevada a cabo en dos (2) apartados principales:

- A1: Revisión de avances tecnológicos de nanofluidos poliméricos con CQDs para recobro mejorado basados en HPAM.
- A2: Implementación de modelos de aprendizaje de máquinas para la predicción de propiedades físicas de soluciones poliméricas

Cada apartado fue desarrollado a través de ecuaciones de búsqueda manteniendo los periodos entre los años 2019 y 2025 para mantener el desarrollo de este trabajo dentro de los avances actuales. Posterior al levantamiento de información inicial, cada documento

fue revisado para establecer si representa alguna relevancia para la investigación o procede a ser descartado para la revisión.

3.3 Resultados y discusión

La revisión sistemática de literatura permitió identificar 45 artículos relevantes enfocados en las dos (2) áreas principales. En la **Tabla 3-1** se presentan las ecuaciones de búsqueda utilizadas para establecer los desarrollos recientes en EOR con nanofluidos poliméricos y aprendizaje de máquinas. Para cada apartado se presenta el respectivo grafico de barras con la cantidad de artículos publicados por año, mostrando la tendencia de desarrollo del conocimiento alrededor de estos temas.

Tabla 3-1. Resultados de la revisión sistemática de literatura

Apartado principal	Ecuación de búsqueda	Cantidad de artículos	Artículos seleccionados
A1	TITLE-ABS-KEY ({polymer flooding} AND {enhanced oil recovery} AND {HPAM}) AND PUBYEAR > 2018 AND PUBYEAR < 2025 AND (LIMIT-TO (EXACTKEYWORD , "Enhanced Oil Recovery") OR LIMIT-TO (EXACTKEYWORD , "Polymer Flooding") OR LIMIT-TO (EXACTKEYWORD , "Oil Well Flooding") OR LIMIT-TO (EXACTKEYWORD , "Enhanced Recovery") OR LIMIT-TO (EXACTKEYWORD , "Floods")) AND (LIMIT-TO (LANGUAGE , "English"))	113	20
A2	TITLE-ABS-KEY ({machine learning} AND {HPAM} AND {viscosity}) AND PUBYEAR > 2018 AND PUBYEAR < 2025	4	4
	TITLE-ABS-KEY ({machine learning} AND {polymer flooding}) AND PUBYEAR > 2018 AND PUBYEAR < 2025	14	5
	TITLE-ABS-KEY ({machine learning} AND {enhanced oil recovery} AND {polymer}) AND PUBYEAR > 2018 AND PUBYEAR < 2025 AND (LIMIT-TO (EXACTKEYWORD , "Enhanced Recovery") OR LIMIT-TO (EXACTKEYWORD , "Viscosity") OR LIMIT-TO (EXACTKEYWORD , "Machine Learning")) AND (LIMIT-TO (LANGUAGE , "English"))	12	9

Por otro lado, con las ecuaciones de búsqueda utilizadas se estableció la cantidad de artículos publicados por año desde el 2019. Los resultados se presentan en la **Figura 3-1** donde la cantidad de artículos sobre avances relacionados con recobro mejorado con polímeros se mantienen por encima de los quince (15) artículos por año con un promedio de publicación anual de diecinueve (19) artículos. Por otro lado, en el área de aprendizaje de máquinas aplicado al EOR y la inyección de polímeros se puede evidenciar una proporción menor en cuanto a la cantidad de artículos y estudios encontrados en la base de datos scopus. Referente al apartado dos (2), en el 2023 se reportó la cantidad máxima de artículos publicados con once (11) artículos, sin embargo, el promedio general de artículos para este apartado es cercano a cinco (5) artículos por año presentando un año sin publicaciones en el tema (2020).

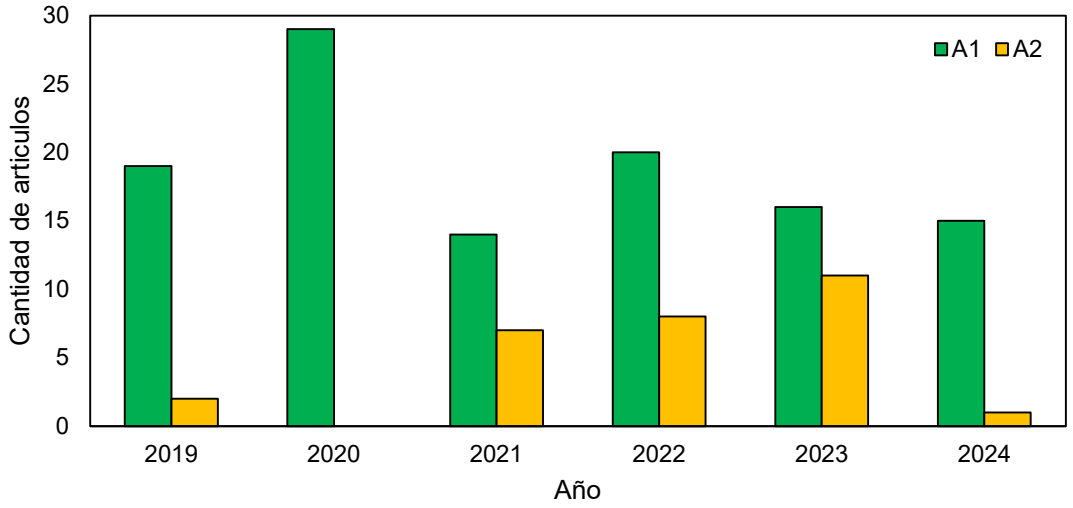

Figura 3-1. Publicación de artículos por año para cada apartado.

Con los artículos seleccionados se desarrollaron las siguientes secciones identificando los requerimientos para evaluar los parámetros fisicoquímicos en la conformación de nanofluidos poliméricos y la predicción de viscosidad utilizando modelos de aprendizaje de máquinas.

3.3.1 Efectos de los parámetros fisicoquímicos en la conformación de nanofluidos poliméricos y la inclusión de CQDs

La implementación de polímeros para EOR es, entre las técnicas de recobro químico, una estrategia eficiente, probada y masificada alrededor del mundo [7, 11, 51]. Los resultados de la revisión sistemática muestran un interés general por biopolímeros en procesos de recobro mejorado, mecanismos de interacción y potencialidades frente a alternativas típicas [11], sin embargo, los polímeros sintéticos como el HPAM representan hoy en día una alternativa fiable, económica y probada ampliamente [13, 51].

La presencia de iones en la solución polimérica es un factor clave en el entendimiento de las propiedades fisicoquímicas del fluido armado. Las cadenas poliméricas del HPAM muestran una alta afinidad por iones divalentes por encima de las interacciones en presencia de iones monovalentes. Gallego *et al* [27] estudiaron el efecto de nanopartículas de sílice funcionalizadas con aminas en soluciones poliméricas de diferentes composiciones de iones monovalentes (NaCl), divalentes ($CaCl_2{\cdot}2H_2O$) y una mezcla en relación 5:1 NaCl:$CaCl_2{\cdot}2H_2O$ discretizando la perdida de viscosidad en función de la concentración de polímero y nanopartículas. Para este estudio se presentaron perdidas de viscosidad del polímero de 35, 70 y 73% para las salmueras monovalentes, divalentes y mezcla de iones respectivamente a los 7 días de seguimiento. Los autores además mostraron que, cuando se adicionan nanopartícula con capacidad de adsorción de cationes la perdida de viscosidad puede verse reducida hasta 15, 50 y 0%

respectivamente. El efecto de las especies iónicas en soluciones poliméricas se conoce ampliamente, múltiples autores establecen que, adicional al tipo de ion presente en la salmuera, es necesario cuantificar la cantidad total dispersa en la solución polimérica [10, 11, 32, 52]. Salem *et al* [32] probaron el efecto de la cantidad de iones en la viscosidad mostrando que para cualquier escenario de valoración los incrementos en la salinidad reducen significativamente la viscosidad del polímero conformado.

Otro factor importante para evaluar los mecanismos de degradación de las soluciones poliméricas es la temperatura. Dentro de los estudios seleccionados como relevantes para este trabajo se establece un rango de temperatura fijo o variable que permite diferenciar los efectos de la temperatura y los mecanismos favorecidos cuando se incrementa la temperatura. Choi *et al* [53] explican que el proceso de hidrolisis se da por una distorsión y compresión de las cadenas poliméricas agravado por la presencia de iones divalentes que cierran las cadenas poliméricas por las zona de radicales libres del HPAM [53]. Si bien las variaciones de temperatura constituyen un mecanismo de degradación térmica esta favorece los procesos de degradación química producida por los radicales libres y los iones presentes en la solución [53].

Dugstad *et al* [8] muestran una breve revisión de la implementación de trazadores en proyectos de recobro mejorado. Para el apartado de recobro con polímeros exponen la importancia de los trazadores para establecer la dirección del polímero y en algunos casos el grado de retención en la roca. La inyección de trazadores para procesos de recobro con polímero puede darse como coinyección continua o inyecciones alternadas dependiendo de las facilidades presentes en el campo. Si bien los trazadores utilizados en este proceso son pasivos en su mayoría, es necesario comprender los mecanismos de interacción entre los trazadores y el polímero inyectado para ajustar parámetros que permitan mantener los niveles de producción dentro de los volúmenes necesarios [8].

Algunos avances enfocan los CQDs como materiales carbonosos con una alta afinidad por el HPAM. Yao *et al* [14] desarrollaron simulaciones moleculares con el fin de establecer el mecanismos fenomenológicos detrás de la reducción de viscosidad de soluciones poliméricas en presencia de CQDs. Las cadenas poliméricas del HPAM tienden a ser adsorbidas en la superficie del CQD cerrando las cadenas poliméricas de punto a punto sobre el CQD interrumpiendo la red polimérica. Esta interacción reduce el radio de giro de la molécula de HPAM impidiendo los procesos de hidratación del polímero y se asocia principalmente a los grupos amida presentes en la cadena polimérica más que a los grupos funcionales cargados [14]. Haruna *et al* [15] comprobaron a través de técnicas reológicas, térmicas y espectroscópicas el impacto de los CQDs en composites de HPAM encontrando una disminución en la viscosidad, mayor sensibilidad a la degradación térmica y la generación de grupos funcionales que facilitan la formación de cadenas cerradas alrededor del CQD [15]. Si bien las investigaciones tienden a apuntar que los CQDs reducen la viscosidad de soluciones poliméricas, Zhao *et al* [16] probaron puntos cuánticos de silicio sulfonados (S-SiQDs) soluciones poliméricas en presencia de iones monovalentes

mostrando que además de incrementar la viscosidad un 38% sobre el valor base con un recobro de aceite del 28.3% superior al 17.2% alcanzado cuando se evalúa la solución polimérica sin S-SiQDs [16]. Estas diferencias presentan una oportunidad para establecer claramente el efecto de los CQDs, su origen e impacto en las propiedades de flujo de las soluciones poliméricas.

Esta revisión sistemática permite establecer las variables relevantes para las valoraciones de los polímeros para EOR en presencia de CQDs. Las variables utilizadas para llevar a cabo este estudio son la salinidad, tiempo de añejamiento del nanofluido, tasa de corte, tipo de CQD y orden de adición siendo este último importante para conocer los cambios de viscosidad bajo diferentes esquemas de inyección. Con la viscosidad y las variables seleccionadas se implementarán modelos de aprendizaje de máquinas que permitan predecir la viscosidad de los nanofluidos poliméricos.

3.3.2 Implementación de modelos de aprendizaje de máquinas para la predicción de propiedades físicas de soluciones poliméricas

Los modelos de aprendizaje de máquinas se presentan como una estrategia eficiente y precisa para la predicción de viscosidad de soluciones poliméricas. Algunos autores destacan una oportunidad para integrar los parámetros fisicoquímicos de las soluciones poliméricas a la base de predicción incrementando la capacidad de generalización del modelo reduciendo los costos de evaluación experimental y los tiempos de implementación de las estrategias de recobro mejorado. Este mejoramiento permitiría extender los resultados experimentales a escenarios de implementación reales basados en parámetros de alta relevancia como la salinidad, concentración y temperatura entre otros parámetros.

Los desarrollos e implementaciones para los procesos de recobro con polímeros se dividen en dos categorías importantes: la predicción de viscosidad a partir de parámetros fisicoquímicos y la aplicación de modelos reológicos asistidos con modelos de aprendizaje de máquinas. Para cada caso es necesario generar un conjunto de datos representativo del comportamiento del polímero, sin embargo, los parámetros incluidos dentro del conjunto de datos dependen netamente de la capacidad experimental o enfoques abordados.

El efecto de los parámetros fisicoquímicos en las propiedades de flujo de las soluciones poliméricas es conocido ampliamente a nivel técnico y experimental, de hecho, se sabe que la relación entre los parámetros tiende a alejarse del comportamiento lineal [22, 23, 54] dificultando las tareas de predicción y modelamiento. Shakeel *et al* [23] implementaron modelos de regresión lineal múltiple, máquinas de soporte vectorial, arboles de decisión y redes neuronales para tres (3) polímeros diferentes variando la concentración de polímero, tasa de corte y temperatura de evaluación. Los resultados demostraron que los modelos de regresión lineal y máquinas de soporte vectorial no se ajustan a los datos contrario al desempeño exhibido por los modelos de árboles de decisión y redes neuronales evaluados utilizando las métricas RMSE y MAPE. El estudio permitió identificar, adicionalmente, que

para variaciones en el polímero las métricas de desempeño difieren considerablemente cuando se compara el mismo modelo, por ejemplo, para el modelo de árboles de decisión y los polímeros SAV-10 y DPTLB-2070 se presenta un RMSE aproximado de 30 y 130 respectivamente [23].

Maillón *et al* [22] probaron modelos de aprendizaje de máquinas para predecir la viscosidad de diferentes polímeros fabricados por la compañía SNF, en Francia. El conjunto de datos fue conformado por 74329 mediciones experimentales propias de la compañía con una amplia variedad de parámetros fisicoquímicos del polímero y las condiciones de armado como la salinidad, concentración del componente activo, temperatura entre otros. Inicialmente se demostró que la viscosidad del polímero presenta una fuerte correlación con la concentración, el peso molecular, los sólidos disueltos, la naturaleza química del polímero y la tasa de corte de medición. Los autores implementaron arboles de decisión obteniendo métricas de desempeño para entrenamiento y evaluación de 99.7 y 89.6% para R^2 y un MAPE de 3.8 y 35.8% respectivamente. Los autores concluyen que a pesar de este desempeño el modelo tiende a sobreajustar reduciendo su capacidad de generalización ante nuevas condiciones de evaluación resaltando la necesidad de incluir nuevas variables de entrada, la sintonización de hiperparámetros y mejorar la calidad de los datos.

Algunos de los avances en este apartado se orientan hacia el relacionamiento fenomenológico de la composición de la salmuera en la predicción de viscosidad. Por ejemplo, Shaik *et al* [55] implementaron el modelo XGboost para predecir la viscosidad de soluciones poliméricas en función de tasa de corte, concentración de polímero, temperatura, salinidad y concentración de iones en la solución obteniendo para entrenamiento y evaluación un MAE de 0.027 y 0.031 respectivamente. Lashari *et al* [21] incluyeron el efecto de composites de HPAM con oxido de grafeno y SiO_2 al conjunto de datos generando así predicciones a partir de la concentración de sales y nanopartículas, temperatura y tasa de corte integradas a once (11) modelos de aprendizaje de máquinas desde regresiones lineales y polinomiales hasta bosques aleatorios y XGboost siendo este último el del mejor desempeño con 0.22, 0.3 y 0.21 para las métricas MAE, RMSE y el criterio de información de Akaike.

La viscosidad representa el parámetro principal para los procesos de recobro mejorado con polímero enfocado al incremento de la eficiencia de barrido. La complejidad de la viscosidad de las soluciones poliméricas radica en su dependencia a diferentes parámetros fisicoquímicos, por ejemplo, para la misma solución se pueden presentar diferentes perfiles reológicos dependiendo del flujo a través del medio poroso, fracturas en el yacimiento, los perforados o inclusive en las tuberías de inyección [13]. Mantener la viscosidad dentro de los valores fijados para el proyecto de recobro es clave para garantizar los efectos de desplazamiento en el medio poroso y mantener los costos de implementación dentro de los rangos preestablecidos.

Karpan *et al* [20] desarrollaron mediciones virtuales de viscosidad utilizando modelos de aprendizaje de máquinas para predecir la viscosidad de polímeros para recobro mejorado en tiempo real utilizando regresión lineal, de segundo orden y bosques aleatorios. Los modelos tienen como objetivo predecir la viscosidad del polímero en función de parámetros como la presión en cabeza de pozo y la tasa de inyección al yacimiento utilizando como variable de entrenamiento mediciones de viscosidad obtenidas a partir de viscosímetros de línea. Las métricas de desempeño resultantes se presentan en la **Tabla 3-2** donde se puede evidenciar que el mejor desempeño corresponde al modelo de árboles aleatorios con un MAE y MSE de 0.22 y 0.18 respectivamente. Este modelo fue probado posteriormente con información de inyección de un campo diferente demostrando un claro ajuste inclusive cuando no se registraron datos experimentales o de los viscosímetros de línea. Los autores concluyen que a pesar del buen desempeño de los modelos la inclusión de parámetros fisicoquímicos y el incremento del tamaño del conjunto de datos tendría un efecto positivo en la calidad de la predicción.

Tabla 3-2. Métricas de desempeño para predicción virtual de viscosidad. Tomado de [20], adaptado por el autor

Modelo	MAE	MSE
Regresión lineal	0.68	0.78
Regresión polinomial	0.45	0.44
Bosques aleatorios	0.22	0.18

La implementación de modelos de aprendizaje de máquinas para mejorar la precisión de modelos de ajuste fenomenológico es un área ampliamente desarrollada. El concepto principal detrás de esta estrategia consiste en utilizar grandes conjuntos de datos para entrenar modelos y predecir los parámetros físicos a partir de mediciones rápidas, por ejemplo, viscosímetros de campo. Los modelos de predicción permiten acelerar los cálculos y reducir el costo computacional de simulaciones de base [56-61] , propiedades físicas [62-65] y procesamiento de imágenes [66] para el modelamiento de procesos en el medio poroso entre otras aplicaciones.

Para los modelos reológicos asistidos por inteligencia artificial representan una alternativa de alto valor para predecir el comportamiento de flujo ante diferentes escenarios de inyección. Dentro de los modelos relevantes se presenta el modelo Carreau-Yasuda utilizado especialmente para correlacionar la viscosidad de los fluidos poliméricos y establecer el comportamiento de flujo. El modelo Carreau-Yasuda se compone de parámetros que permiten describir el perfil reológico completo a partir de evaluaciones experimentales de baja complejidad mostrando una alta precisión cuando se extienden los resultados a otros rangos de tasa de corte [59]. A pesar de su relevancia, la implementación del modelo se ha limitado a la predicción de viscosidad mientras que el

entendimiento de los parámetros que componen la predicción se mantiene fuera de los análisis profundos del fluido [60].

Pérez *et al* [18] implementaron algoritmos genéticos destacando esta alternativa sobre las redes neuronales debido a conexiones más simples mejorando así el entendimiento de la predicción de los modelos. En este estudio se implementaron arboles de decisión para predecir los coeficientes del modelo Carreau-Yasuda a partir de un conjunto de datos previamente establecido que contiene el tipo de polímero, peso molecular, concentración y los coeficientes del ajuste. La implementación de los árboles de decisión permitió predecir la viscosidad a tasa de corte nula, tiempo de relajación del polímero y el índice de potencia con un valor de $R^2 > 0.947$ para un rango de temperatura y concentración fija. Los autores desarrollaron una metodología eficiente de bajo costo para predecir el comportamiento reológico completo de soluciones poliméricas utilizando mediciones sencillas provenientes de un viscosímetro brookfield [18].

Zeynalli *et al* [19] implementaron aprendizaje de máquinas para extender el modelo viscoelástico extendido unificado. El modelo permite predecir la degradación mecánica cerca de la zona de inyección del pozo donde se presentan altas tasas de corte a partir de propiedades de la roca y polímero implementando modelos de regresión simbólica. Para evaluar el desempeño, el modelo fue agregado en el software de simulación de yacimiento Matlab Simulation Toolbox (MRST). El modelo asistido por machine learning permite identificar los cuatro (4) regímenes de flujo durante el proceso de inyección, reduciendo la sobreestimación de aceite recuperado comparado con el modelo de referencia y Carreau-Yasuda con datos experimentales limitados.

A pesar de contar con resultados prometedores, los modelos asistidos por aprendizaje de máquinas cuentan con algunas desventajas. Por ejemplo, [17] compararon un modelo empírico desarrollado, el modelo Carreau-Yasuda y 9 métodos de machine learning desde regresiones lineales hasta bosques aleatorios y redes neuronales. El resultado demostró que, para las condiciones evaluadas, el modelo empírico presenta un mejor desempeño para estimar la viscosidad a tasa de corte nula mientras que los modelos de aprendizaje de máquinas predicen el índice de potencia y tiempo de relajación con tendencia a sobreajustar. Este acercamiento permitió identificar que, para conjuntos de datos pequeños, la tendencia al sobreajuste de los métodos es marcado y reduce la capacidad de generalización de los modelos ante variaciones en los parámetros fisicoquímicos.

En la **Tabla 3-3** se presentan los resultados generales de la revisión sistemática de literatura enfocada a la evaluación de modelos de aprendizaje de máquinas para la predicción de propiedades y predicción asistida del comportamiento reológico de polímeros. La revisión muestra que los modelos con mayor potencial de aplicación corresponden a las máquinas de soporte vectorial, arboles de decisión, bosques aleatorios y redes neuronales. Las regresiones lineales se implementan ampliamente como punto de comparación con los modelos complejos.

Tabla 3-3. Conclusiones de la revisión sistemática de literatura para la implementación de modelos de aprendizaje de máquinas en la predicción de propiedades de nanofluidos poliméricos

Estudio	Objetivos	Modelos implementados	Métricas de evaluación
[22]	Predicción de viscosidad con parámetros fisicoquímicos	Arboles de decisión Bosques aleatorios	R^2 Valores de referencia
[23]	Predicción de viscosidad con parámetros fisicoquímicos	Regresión lineal Máquinas de soporte vectorial Arboles de decisión Redes neuronales	RMSE MAPE
[55]	Predicción de viscosidad con parámetros fisicoquímicos	XGboost	MAE
[21]	Predicción de viscosidad con parámetros fisicoquímicos	RL-Lasso RL-Ridge Máquinas de soporte vectorial Arboles de decisión Bosques aleatorios Regresión de árboles extra Gradient boosting XGboost	MAE RMSE Criterio de información de Akaike
[20]	Predicción de viscosidad con parámetros de inyección	RP RL Bosques aleatorios	MAE RMSE
[18]	Modelos asistidos	Algoritmos genéticos	RMSE
[19]	Modelos asistidos	Regresión simbólica	Comparación con valores de referencia
[17]	Modelos asistidos	RP RL-Huber RP-Huber RL-Ridge Máquinas de Soporte Vectorial Arboles de Decisión Bosques aleatorios Redes Neuronales	RMSE

4. Metodología para la predicción de viscosidad de nanofluidos poliméricos utilizando modelos de aprendizaje de máquinas

En este capítulo se exponen los materiales y métodos utilizados para desarrollar y evaluar los modelos de aprendizaje de máquinas para la predicción de viscosidad de nanofluidos poliméricos. La primera sección experimental presenta los materiales, equipos y técnicas instrumentales para conformar, caracterizar y evaluar los componentes de los nanofluidos poliméricos. Posterior a esta sección se aborda la metodología para conformación del conjunto de datos con su respectivo preprocesamiento y análisis descriptivo para finalmente abordar los diferentes modelos de aprendizaje de máquinas previamente seleccionados y evaluados con métricas identificadas en la revisión sistemática de literatura.

4.1 Metodología experimental para la caracterización y medición de viscosidad de nanofluidos poliméricos

Considerando la relevancia técnica del HPAM para procesos de recobro este fue utilizado para el desarrollo de las pruebas estáticas y la conformación del conjunto de datos. Como base de comparación para los procesos de trazado de fluidos en el yacimiento fueron comparados dos CQD comerciales etiquetados como CQDred y CQDgreen por sus colores característicos de fluorescencia suministrados localmente. Adicionalmente, fue sintetizado un tipo de CQDblue a través de síntesis hidrotermal utilizando la ruta de síntesis expuesta previamente por [36].

Para establecer la naturaleza química de los materiales se realizaron medidas de espectroscopía infrarroja por transformada de Fourier (FTIR) con un espectrofotómetro IRAffinity 1-2 (Shimadzu Corporation, Japón) acompañado de mediciones de tamaño de partícula y potencial Z utilizando un nanoplus-3 (Micrometrics, USA) para los CQDs.

Para incluir los efectos de salinidad en los nanofluidos poliméricos fueron preparadas soluciones de 0, 1000, 3000 y 5000 $mg \cdot L^{-1}$ utilizando una formulación estudiada previamente por [27] para fijar la condición de degradación termoquímica del HPAM. Para la preparación de cada salmuera se utilizó una relación 5:1 de $NaCl:CaCl_2 \cdot 2H_2O$ para llevar a agitación durante 2 horas garantizando disolución total de cada componente en la fase acuosa. Para establecer el efecto del orden de adición en la conformación del nanofluido polimérico fue comparada la adición inicial de CQDs seguido de la adición del polímero mientras que, para el escenario contrario, el polímero fue conformado inicialmente para

finalizar con la adición del CQD una vez el polímero se encuentra conformado. Algunos estudios previos concluyen que la aplicación del HPAM se limita a cierta dureza presente en la salmuera en función de la temperatura, por ejemplo, para 75°C se recomiendan salmueras con dureza inferior a 2000 mg·L^{-1} [67]. En la **Figura 4-1** se presentan las concentraciones de iones utilizadas para cada valor de salinidad evaluado mostrando concentraciones máximas de 1639, 266 y 3060 mg·L^{-1} para los iones de Na^{+}, Ca^{2+} y Cl^{-} respectivamente.

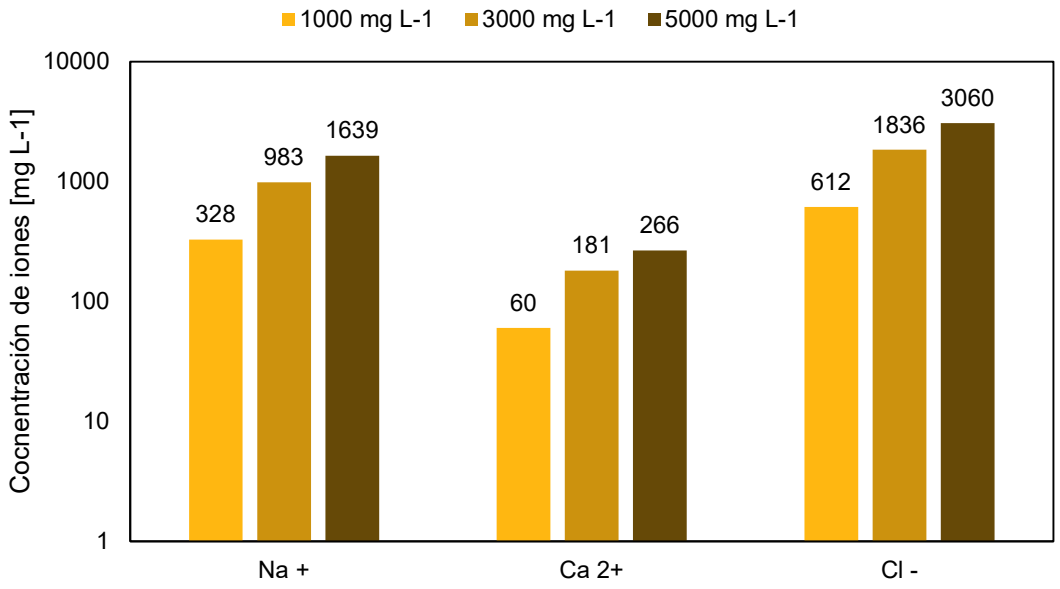

Figura 4-1. Concentración de iones en las salmueras implementadas

Cada solución polimérica fue armada en 100 mL de salmuera a una concentración fija de 2000 y 100 mg·L^{-1} para el HPAM y CQDs respectivamente establecidos previamente como parámetros típicos de inyección en procesos de EOR. Mantener las concentraciones de polímero y CQDs permite identificar la capacidad de predicción de los modelos cuando varia la salinidad y el tipo de CQD siendo estas variables identificadas con un alto potencial de variación en aplicaciones reales.

Cada solución fue preparada considerando el orden de adición y agitados durante 2 horas a 250 rpm utilizando un homogeneizador DLH (VELP Scientifica, Italia) para ser sometidos a temperatura de 70°C para inducir el proceso de añejamiento y evaluación reológica. La evaluación reológica fue llevada a cabo en un reómetro rotacional Kinexus Pro (Malvern Instruments, Reino Unido) con la geometría de cilindros concéntricos a 70°C y una tasa de corte variando entre 0.01 y 300 s^{-1} para cubrir el rango de interés de patrones de flujo en el yacimiento y transporte por tubería reportados [13]. Cada medición reológica fue llevada a cabo para el tiempo inicial una vez fue conformado el polímero, 24, 72 y 144 horas siguiendo el añejamiento y reducción de viscosidad asociado a la degradación térmica del nanofluido.

4.2 Conformación y análisis descriptivo del conjunto de datos experimentales.

Para la conformación del conjunto de datos fueron tomadas las curvas de flujo generadas en el reómetro y asociadas a un tiempo de añejamiento, orden de adición y tipo de CQD en un archivo almacenándolos en formato tabular y divididos en tasas de corte representativas de 7.1, 15, 50, 75, 100, 150, 200 y 300 s^{-1} para relacionar la información completa del comportamiento de flujo de los polímeros dentro de los modelos de aprendizaje de máquinas implementados. En la **Tabla 4-1** se presentan las variables y los rangos de variación para los experimentos realizados. Desde la conformación del conjunto de datos el orden de adición fue codificado como uno (1) para la adición inicial de CQDs al nanofluido, dos (2) la adición tardía y tres (3) como la conformación del polímero sin CQDs. La etapa experimental de conformación y medición inicial de cada polímero consta de cinco (5) horas para el posterior seguimiento durante 144 horas donde fueron generados 896 datos, sin embargo, posterior a una revisión inicial, 89 datos individuales fueron descartados por presentar inconsistencias físicas asociadas al proceso de medición. El conjunto de datos fue entonces cargado y preprocesado utilizando Python 3.12.2 utilizando las librerías de distribución libre pandas, numpy, sklearn y seaborn.

Tabla 4-1. Parámetros y rangos de variación para el conjunto de datos generado.

Parámetro	Rango de variación	Unidades
Concentración de HPAM	2000	$mg \cdot L^{-1}$
Concentración de CQDs	[0, 100]	$mg \cdot L^{-1}$
Tipo de CQD	[green, red, blue, none]	-
Salinidad	[0, 1000, 3000, 5000]	$mg \cdot L^{-1}$
Tiempo de añejamiento	[0, 24, 72, 144]	horas
Orden de adición	[1, 2, 3]	-
Tasa de corte	[7.1, 15, 50, 75, 100, 150, 200, 300]	s^{-1}

Una vez conformado el conjunto de datos fue preprocesado para definir las variables categóricas utilizando variables “dummy”. Las variables dummy son una codificación binaria que permite representar claramente la presencia o ausencia del valor categórico en la variable objetivo [68]. Para este caso, las variables categóricas son identificadas y codificadas para incluir el efecto de estos parámetros en el entrenamiento y predicción del modelo. Por otro lado, es importante destacar que los nanofluidos poliméricos cuentan con

una alta viscosidad en el tiempo inicial y presentan un descenso una vez se incrementa la salinidad y el tiempo de añejamiento inclusive para el tiempo de evaluación inicial de 24 horas. Algunos estudios presentados anteriormente evidencian que la inclusión de CQDs en los nanofluidos poliméricos reducen la viscosidad y afectan las propiedades de flujo de los nanofluidos. Esta condición de desbalanceo de datos afecta el desempeño del modelo dado al efecto de la magnitud donde el 75% de los datos se ubican por debajo de una viscosidad de 20 cP y el 25% restante presenta valores que oscilan entre 20 y 134 cP. A pesar de esta variación y condición de desbalanceo de los datos, los valores extremos se mantuvieron dentro del análisis sin modificación dando representatividad fenomenológica. El análisis de correlación de Pearson fue realizado para identificar el tipo de relación entre las variables numéricas con la viscosidad del polímero estableciendo la condición de linealidad en el conjunto de datos utilizando la **Ecuación 4-1** [69] . Esta ecuación calcula el coeficiente de correlación entre dos variables continuas de interés siendo 1 la correlación lineal directa, -1 correlación inversa y 0 la no correlación entre las variables.

Ecuación 4-1

$$corr(x,y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

4.3 Modelos de aprendizaje de máquinas para la predicción de viscosidad de nanofluidos poliméricos

Para predecir los valores de viscosidad en función de las condiciones propuestas se implementaron diferentes modelos de aprendizaje. Los modelos predictivos fueron regresión lineal múltiple (MLR), máquinas de soporte vectorial (SVM), árboles de decisión (DTR) y bosques aleatorios (RF) utilizados desde la librería sklearn. Cada uno de los modelos implementados fueron sometidos a una optimización de hiperparámetros utilizando el método GridSearchCV de la librería sklearn. Los parámetros y rangos de variación se presentan en la **Tabla 4-2.**

Tabla 4-2. Hiperparámetros y rangos de variación para los métodos de predicción

Modelo	Hiperparámetro	Valores
MLR	Ajuste al intercepto	{True, False}
	Positivo	{True, False}
SVM	Kernel	{linear, poly, rbf, sigmoid}
	Grado	{1-10}
	Gamma	{scale, auto}
	Coef0	{0, 1, 3, 5}
	C	{1, 3, 5, 10}
DTR	Criterio	{Friedman mse, absolute error, poisson}
	Divisor	{best, random}
	Profundidad máxima	{None – 30}
	Partición mínima	{1-20}
	Mínimo muestras por hoja	{1-30}
	Características máximas	{sqrt, log2}
RF	Cantidad de estimadores	{50-1000}
	Criterio	{Friedman mse, absolute error, poisson}
	Características Máximas	{1, sqrt, log2}
	Profundidad máxima	{3-30}
	Partición mínima	{1-20}
	Mínimo muestras por hoja	{1-30}
	Características máximas	{sqrt, log2, none}

Adicionalmente, para comparar los modelos de regresión implementados fueron utilizadas las métricas de error medio absoluto, raíz del error cuadrático medio y el coeficiente de determinación (MAE, RMSE y R^2 respectivamente) dado que representan la desviación de los valores predichos con los valores reales del conjunto de evaluación como se presenta en la **Tabla 4-3**. Para establecer visualmente el desempeño de los modelos fueron construidas curvas de desviación del valor real y residuales con la **Ecuación 4-2** como una medida para establecer la distancia con los valores reales.

Tabla 4-3. Métricas implementadas para los modelos de regresión

Métrica	Ecuación
MAE	$\frac{\sum_{i=1}^{n} \lvert y_{i,pred} - y_{i,true} \rvert}{n}$
RMSE	$\sqrt{\frac{\sum_{i=1}^{n} (y_{i,pred} - y_{i,true})^2}{n}}$
R^2	$1 - \frac{\sum_{i=1}^{n} (\hat{y}_{i,pred} - y_{i,true})^2}{\sum_{i=1}^{n} (y_{i,true} - \bar{y}_i)^2}$

Ecuación 4-2

$$Resd = y_{i,true} - y_{i,pred}$$

Las métricas seleccionadas corresponden a las utilizadas ampliamente identificadas en la revisión sistemática de literatura presentada en el capítulo 3.3.2 para determinar el ajuste del modelo. Tanto el MAE como el RMSE permiten identificar la diferencia entre el valor real y el valor predicho con diferencias en la estabilidad e interpretabilidad [70, 71]. Mientras que el MAE califica el desempeño del modelo como una distancia entre el valor real y el valor predicho el RMSE incluye el efecto de los valores atípicos magnificando la diferencia a través del exponente cuadrático para ser estandarizado mediante la raíz cuadrada. Por otro lado, el coeficiente de determinación R^2 permite evaluar con precisión el modelo brindando información sobre la calidad de la predicción con interpretabilidad superior a la de las métricas MAE o RMSE [71]. La interpretación conjunta de estas métricas permite obtener un análisis específico del desempeño de los modelos y su capacidad de predicción para las condiciones del conjunto de datos.

5. Resultados

Este capítulo busca exponer los resultados asociados a la valoración, exploración y predicción de la viscosidad en nanofluidos poliméricos. En el primer apartado de esta sección se presentan los resultados asociados a la caracterización y valoración estática de los nanofluidos poliméricos en presencia de distintos CQDs a través de curvas de flujo para una tasa de corte variando entre 0 y 300 s^{-1}. Seguido de esto, se muestra el análisis exploratorio del conjunto de datos donde se incluye la distribución de los datos y el análisis de correlación para establecer el tipo de relación entre las variables de las pruebas experimentales con la viscosidad de los nanofluidos poliméricos. Finalmente, en el apartado 3 se exponen los resultados asociados a las predicciones de los modelos de aprendizaje de máquinas incluyendo las métricas utilizadas, la comparación entre las predicciones y el valor real con su respectiva comparación de los valores residuales.

5.1 Caracterización y evaluaciones estáticas del nanofluido polimérico

Los materiales fueron caracterizados a través de técnicas instrumentales que permiten establecer la naturaleza física y química de cada CQD y el polímero implementado. Inicialmente, se presenta el análisis FTIR para los CQDs en la **Figura 5-1**. Este espectro revela bandas asociadas al estiramiento del grupo -OH y un pico de lectura C-H entre las longitudes de onda 3500 y 3000 cm^{-1}. Adicionalmente, se muestran picos muy marcados para cada CQD cerca de los 1650 cm^{-1} asociados posiblemente a la existencia del grupo funcional -COO [72, 73] . Los espectros presentados muestran naturalezas químicas similares, sin embargo, el CQDblue presenta leves diferencias en las longitudes entre 1650 y 1150 cm^{-1} asociadas a grupos carboxílicos e hidroxilos [36].

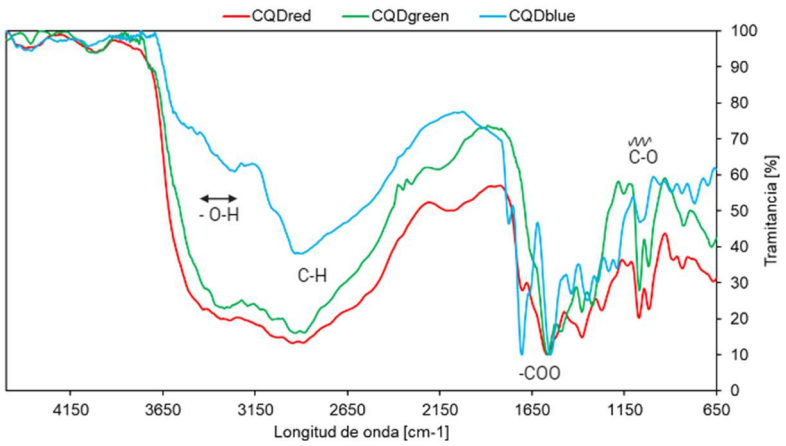

Figura 5-1. Análisis FTIR para los CQDs utilizados

Por otro lado, el análisis FTIR se presenta en la **Figura 5-2** donde se pueden evidenciar bandas características del HPAM. El espectro exhibe bandas típicas del HPAM donde se evidencian señales típicas de los grupos funcionales amidas presentes en la cadena polimérica asociados principalmente a las vibraciones y encogimiento de los grupos. Para la banda ubicada en 725 cm^{-1} se muestra un doblamiento y vibración del grupo N-H y vibraciones para los grupos N-C-O y N-H a 1106 y 3436 cm^{-1} respectivamente [74].

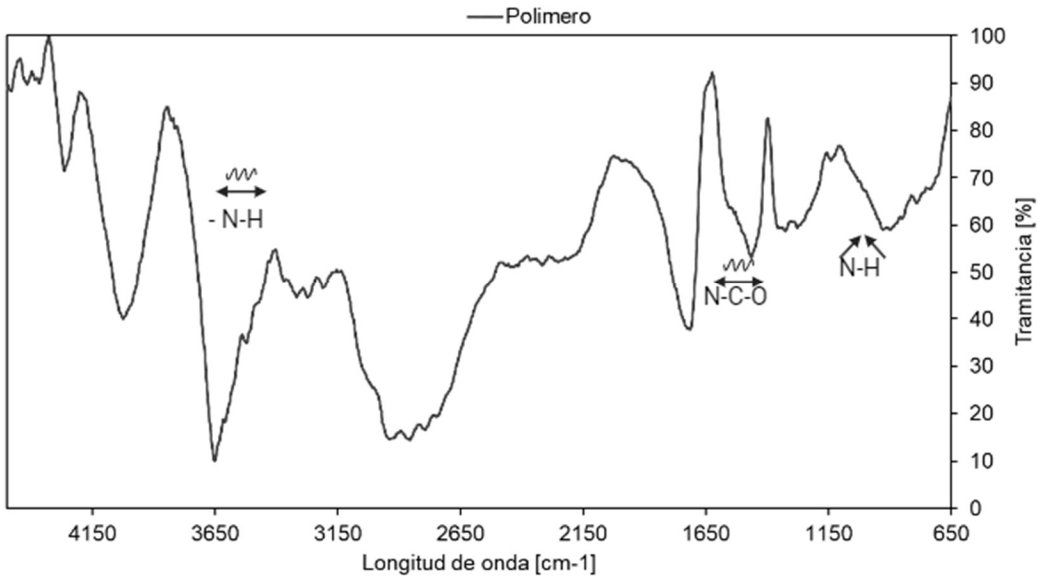

Figura 5-2. Análisis FTIR para el HPAM utilizado.

Por otro lado, en la **Tabla 5-1** se presentan los resultados para el tamaño de partícula y potencial Z de los CQDs implementados. Allí se puede evidenciar que el CQD sintetizado presentó el menor tamaño con 16.3 nm seguido por el CQDgreen comercial con 31.2 nm y finalmente el CQDred como el de mayor tamaño con 44.2 nm. Relacionado con el potencial Z, previamente se ha identificado que los CQDs presentan un potencial Z negativo asociado a la presencia de grupos funcionales oxigenados en la superficie del nanomaterial [36]. Los valores del potencial Z resultaron en -19.67, -47.2, -20.22 mV para el CQDblue, CQDgreen y CQDred respectivamente. Estos valores indican una alta estabilidad coloidal de cada uno de los CQDs en la solución siendo el CQDgreen el más estables de los CQDs evaluados.

Tabla 5-1. Tamaño y potencial Z para los CQDs utilizados

CQD	**Tamaño** [nm]	**Potencial Z** [mV]
Blue	16.3	-19.67
Green	31.2	- 47.2
Red	44.2	-20.22

Las evaluaciones estáticas permitieron identificar los efectos de la adición de CQDs en la solución polimérica. Los resultados permiten concluir el efecto descrito previamente por algunos autores donde se expone que el CQD reduce la viscosidad del polímero en solución debido a su interacción con los radicales libres y la adsorción sobre la superficie del nanomaterial [14, 15]. En **Figura 5-3** se presentan los resultados asociados a viscosidad para diferentes escenarios de salinidad para cada CQD evaluado. Las gráficas muestran un efecto de reducción de viscosidad del polímero cuando se tiene salmuera desionizada y la presencia de CQDs en solución asociado al rompimiento de las cadenas poliméricas por la adsorción del HPAM sobre la superficie del nanomaterial. Cuando la salinidad incrementa se evidencia una disminución significativa en la viscosidad del nanofluido polimérico asociado a la interacción de los iones divalentes con las cadenas del HPAM alcanzando una reducción del 88% cuando se evalúa el escenario de 5000 $mg \cdot L^{-1}$. Cuando se considera el orden de adición de los CQDs con el polímero es posible establecer que no sé presentan diferencias significativas, sin embargo, es posible diferenciar que las muestras donde se adicionan los CQDs inicialmente tienden a presentar valores de viscosidad por encima de los reportados para el escenario base o el escenario de adición de CQDs posterior a la conformación inicial del polímero. La implementación de métodos de aprendizaje de máquinas permitiría capturar esas ligeras diferencias en los mecanismos de conformación para predecir la viscosidad de nanofluidos poliméricos bajo diferentes esquemas de inyección.

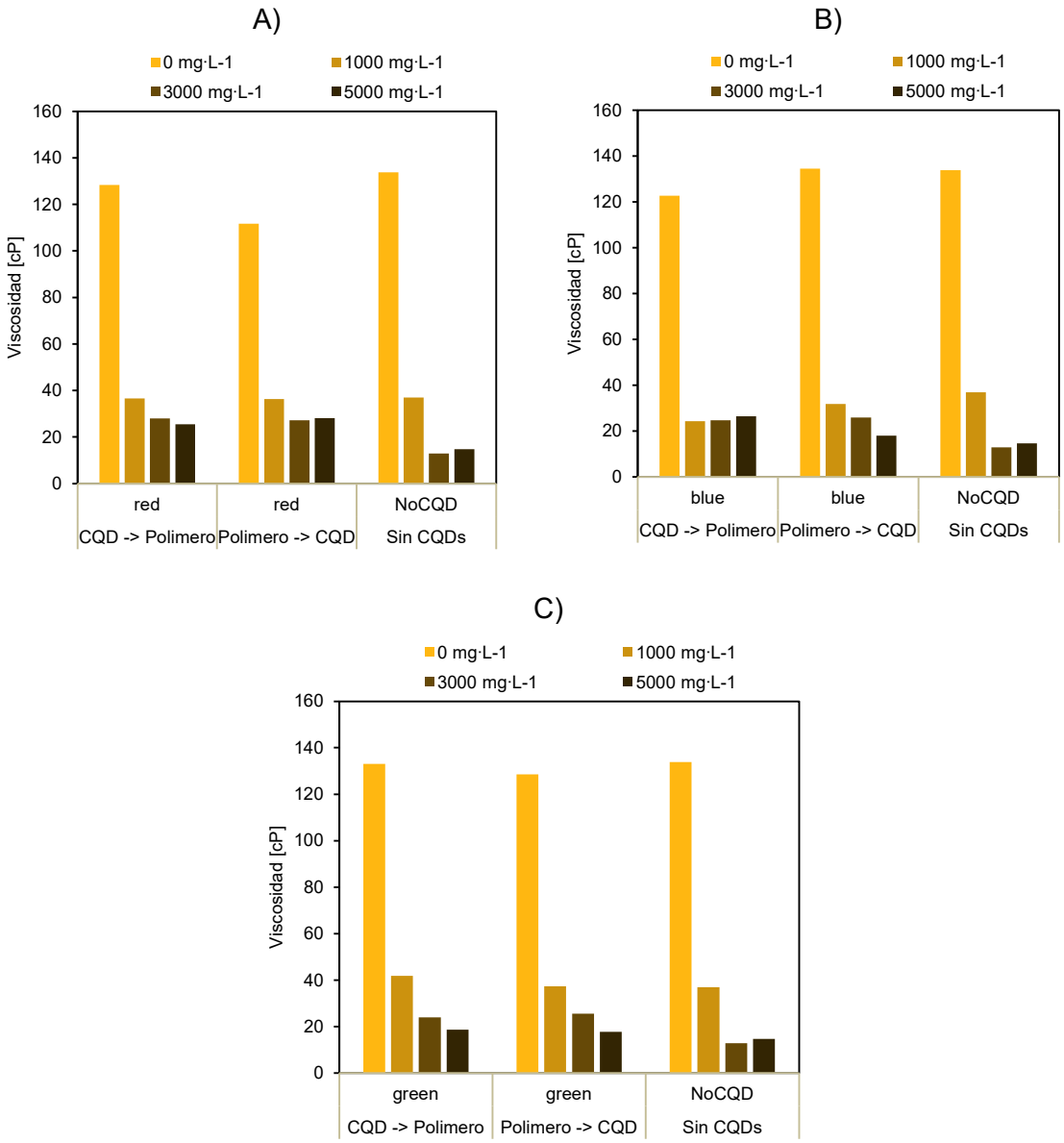

Figura 5-3. Viscosidad de nanofluidos poliméricos en el tiempo inicial y 7 s^{-1} para A) CQDred B) CQDblue y C) CQDgreen

El escenario de evaluación de viscosidad a diferentes salinidades y tiempo de añejamiento se presenta en la **Figura 5-4**. El polímero sin adición de CQDs muestra una disminución significativa de la viscosidad conforme incrementa el tiempo y la salinidad, por ejemplo, para el escenario libre de salinidad se reporta una disminución del 42% mientras que para las salinidades de 3000 y 5000 $mg{\cdot}L^{-1}$ las reducciones corresponden a 82 y 85% respectivamente. La adición de CQDs muestra una disminución importante en la viscosidad para cada CQD cuando se evalúa el escenario libre de presencia de iones con valores de 33.68, 32.47 y 29.03 cP para CQDred, CQDblue y CQDgreen respectivamente. Si bien cada CQD presenta reducciones de viscosidad diferentes, para el caso del CQDblue se presenta valores de viscosidad cercanos de 6 cP para condiciones de

salinidad de 3000 y 5000 mg·L^{-1} a 72 horas de evaluación, sin embargo, el punto de evaluación e 144 horas muestra viscosidades ligeramente superiores hasta 12 y 9.49 cP respectivamente indicando un posible reordenamiento de las cadenas poliméricas en el nanofluido polimérico.

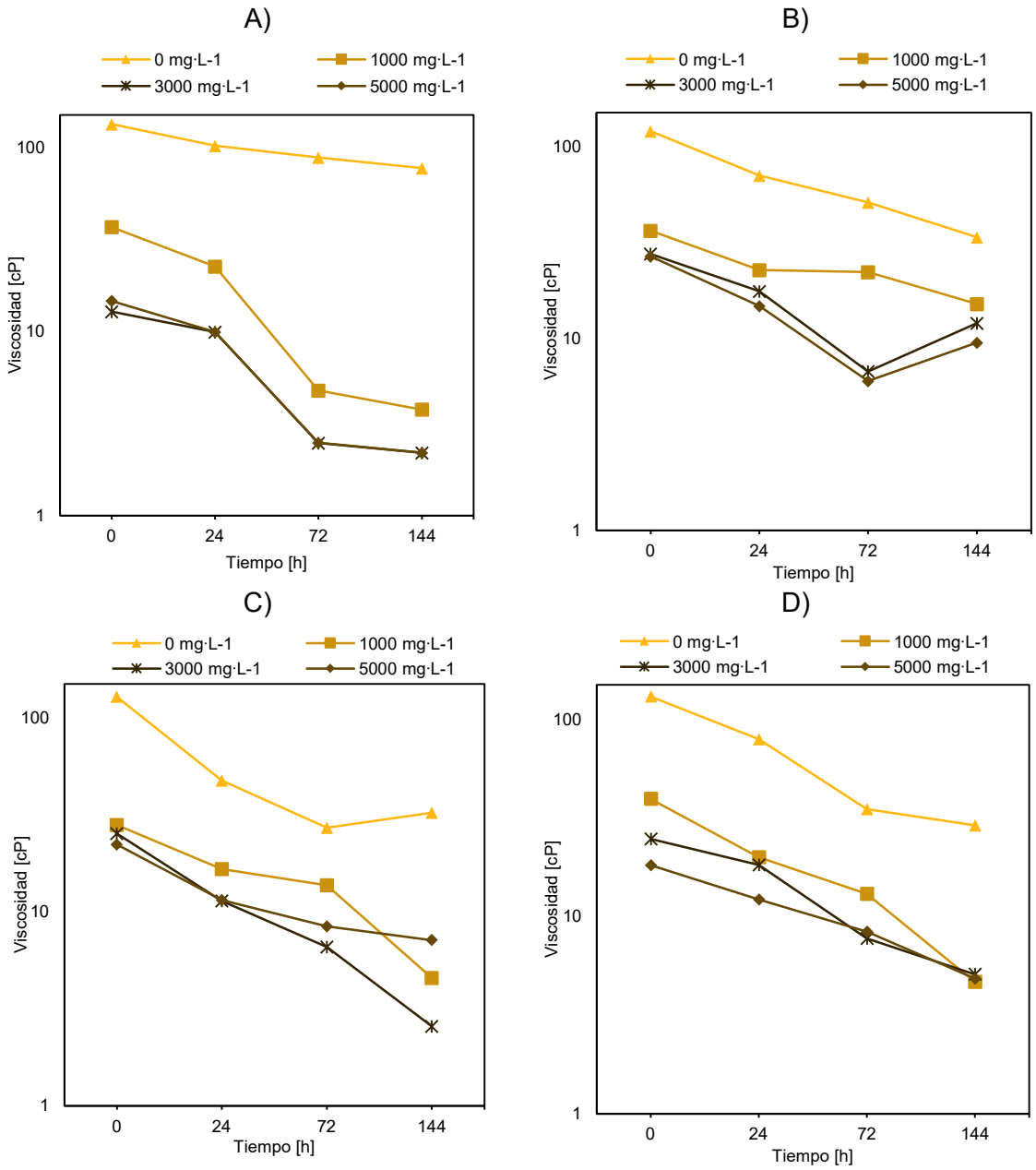

Figura 5-4. Viscosidad de nanofluidos poliméricos en 7 s^{-1}, diferentes salinidades y tiempos de evaluación para A) Sin CQDs, B) CQDred, C) CQDblue y D) CQDgreen

5.2 Análisis descriptivo del conjunto de datos generado

Con las valoraciones estáticas fue posible construir el conjunto de datos representativo contando con 890 filas y 8 columnas. Con el conjunto de datos se realizó el respectivo análisis descriptivo para establecer el tipo de distribución y posibles transformaciones. En la figura **Figura 5-5** se presenta la distribución de los datos donde se puede ver que la mayor concentración de los datos se ubica por debajo de 20 cP, específicamente, el 75% de los datos se ubica por debajo del valor de 18.80 cP. Esta distribución muestra un claro desbalanceo del conjunto de datos como se puede confirmar en la tabla **Tabla 5-2**, sin embargo, los datos ubicados por encima del 75% inicial representan valores de viscosidad importantes como las primeras etapas en un escenario de inyección donde el polímero no presenta ninguna condición de degradación.

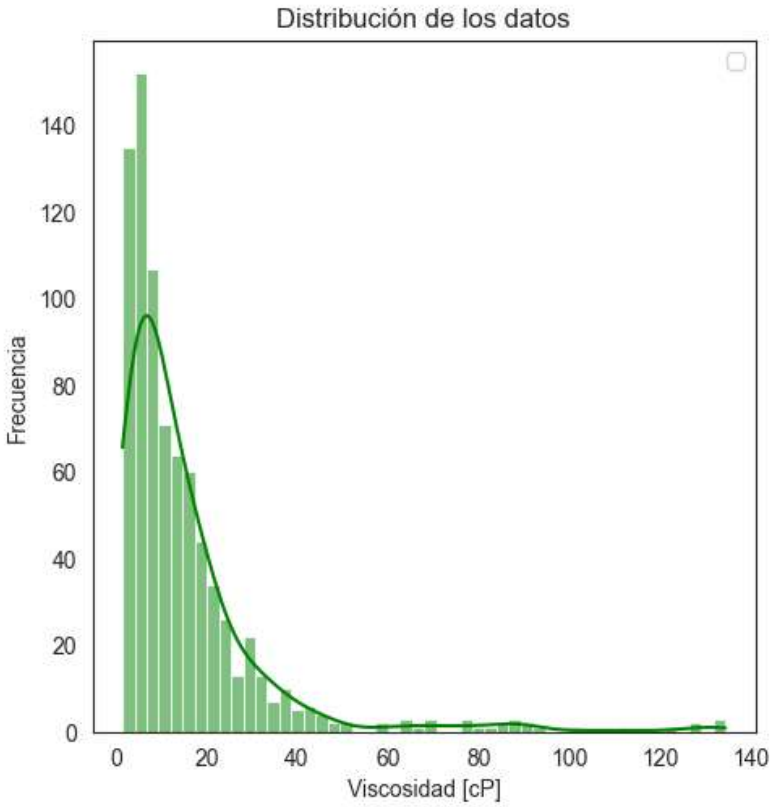

Figura 5-5. Distribución del conjunto de datos generado

Tabla 5-2. Análisis descriptivo de las viscosidades en el conjunto de datos

Variable	Promedio	STD	min	Q1	Q2	Q3	max
Viscosidad [cP]	15.24	18.03	1.40	4.84	9.63	18.20	134.5

Adicionalmente, para establecer el tipo de relación entre las variables numéricas y la viscosidad se calculó el coeficiente de correlación de Pearson presentado en la **Figura 5-6**. En esta matriz de correlación se puede evidenciar el tipo de relación que presentan las variables continuas de salinidad, tiempo, tasa de corte y viscosidad. Estos resultados muestran que estas variables no tienen ningún tipo de correlación entre sí, sin embargo, cuando se relaciona con la viscosidad se presentan valores de -0.47, -0.22, -0.35 para la relación entre salinidad, tiempo y tasa de corte respectivamente indicando una baja correlación inversa.

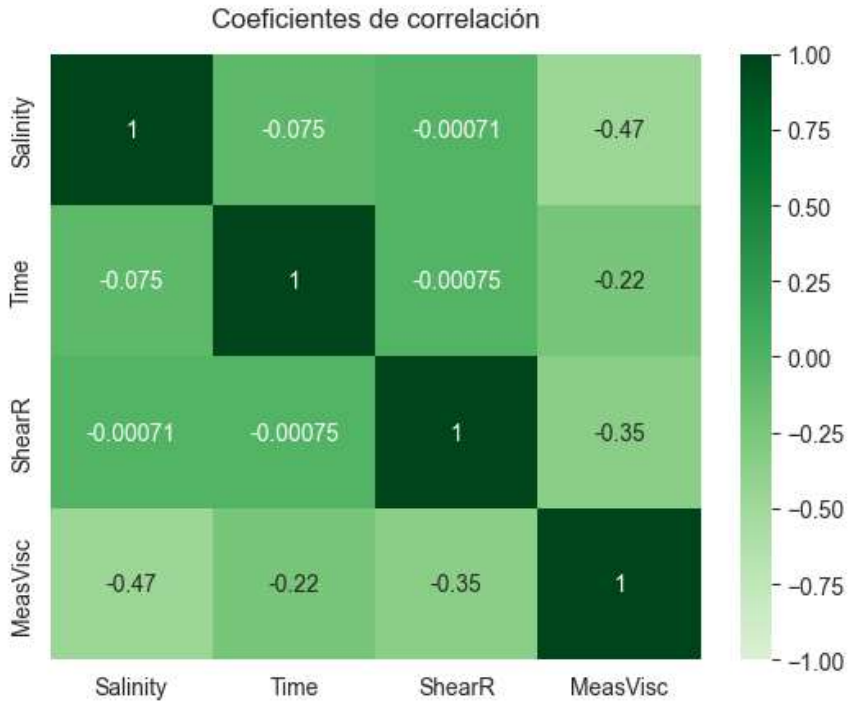

Figura 5-6. Matriz de correlación de Pearson para las variables continuas

Estos resultados permiten identificar un conjunto de datos altamente desbalanceado y con una baja correlación inversa entre las variables con la viscosidad como variable objetivo.

5.3 Desempeño de los modelos de aprendizaje de máquinas implementados

Considerando las métricas de error para los diferentes modelos de predicción de viscosidad, los valores más altos se alcanzaron para los modelos de regresión lineal multiple (MLR) y máquinas de soporte vectorial (SVM) respectivamente como se puede observar en la **Figura 5-7**. Estos resultados permiten identificar que estos modelos de aprendizaje de máquinas no se ajustan a las condiciones de distribución de los datos o correlación entre las variables. Por otro lado, los resultados exhibidos por los modelos de

árboles de decisión (DTR) y bosques aleatorios (RF) muestran que, a pesar de presentar valores altos, reducen significativamente el error y mejoran la predicción de viscosidad. Si bien los modelos DTR y RF presentan valores en las métricas menores, el desempeño del modelo RF sobresale por reducciones generales en el error y un acercamiento claro a los valores de viscosidad reales esperados con valores en el MAE, RMSE y R^2 de 2.7, 7.02 y 0.87 respectivamente.

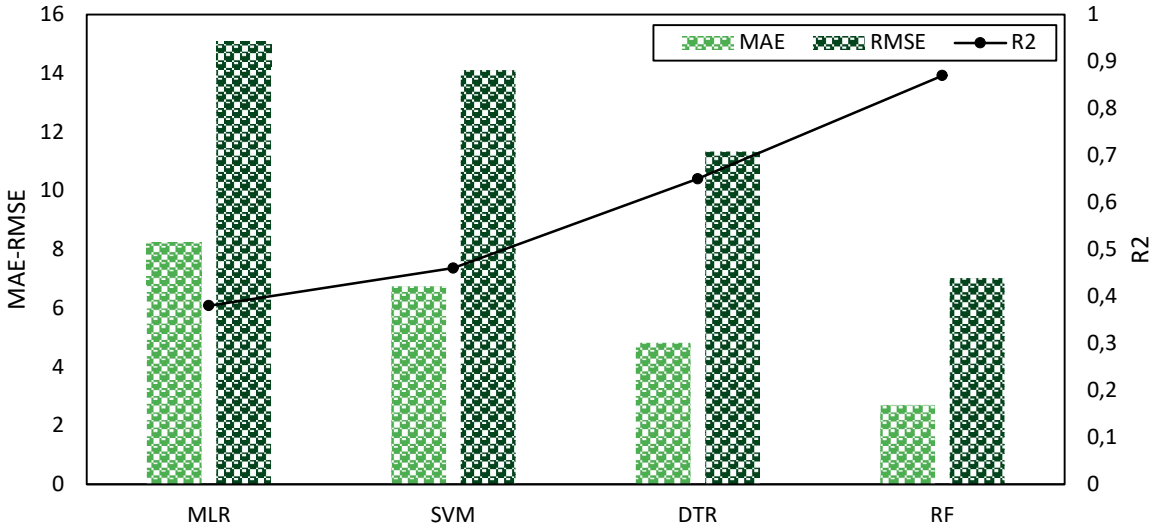

Figura 5-7. Métricas de error obtenidas en la evaluación de los modelos

La **Figura 5-8** presenta gráficos de dispersión donde se compara en el eje x los valores de viscosidad real mientras que en el eje y se ubican los valores predichos por cada modelo. Para cada gráfico se presenta la línea donde los valores reales coinciden con los valores predichos indicando una predicción exacta del valor. Para los modelos MLR y SVM se presentan dispersiones donde el valor predicho no supera 40 cP para cada modelo demostrando así el efecto del claro desbalanceo de los datos en el desempeño del modelo. Por otro lado, los modelos DTR y RF presentan gráficos de dispersión que pueden ubicarse en todo el eje de los valores reales, sin embargo, el modelo DTR tiende a ubicar los valores de alta viscosidad dispersos asociado posiblemente a la baja proporción del conjunto de datos. El modelo RF muestra que, si bien la precisión del modelo mejora, a medida que incrementa el valor de viscosidad presenta un comportamiento similar al mostrado por los modelos MLR y SVM donde las predicciones no superan al valor real.

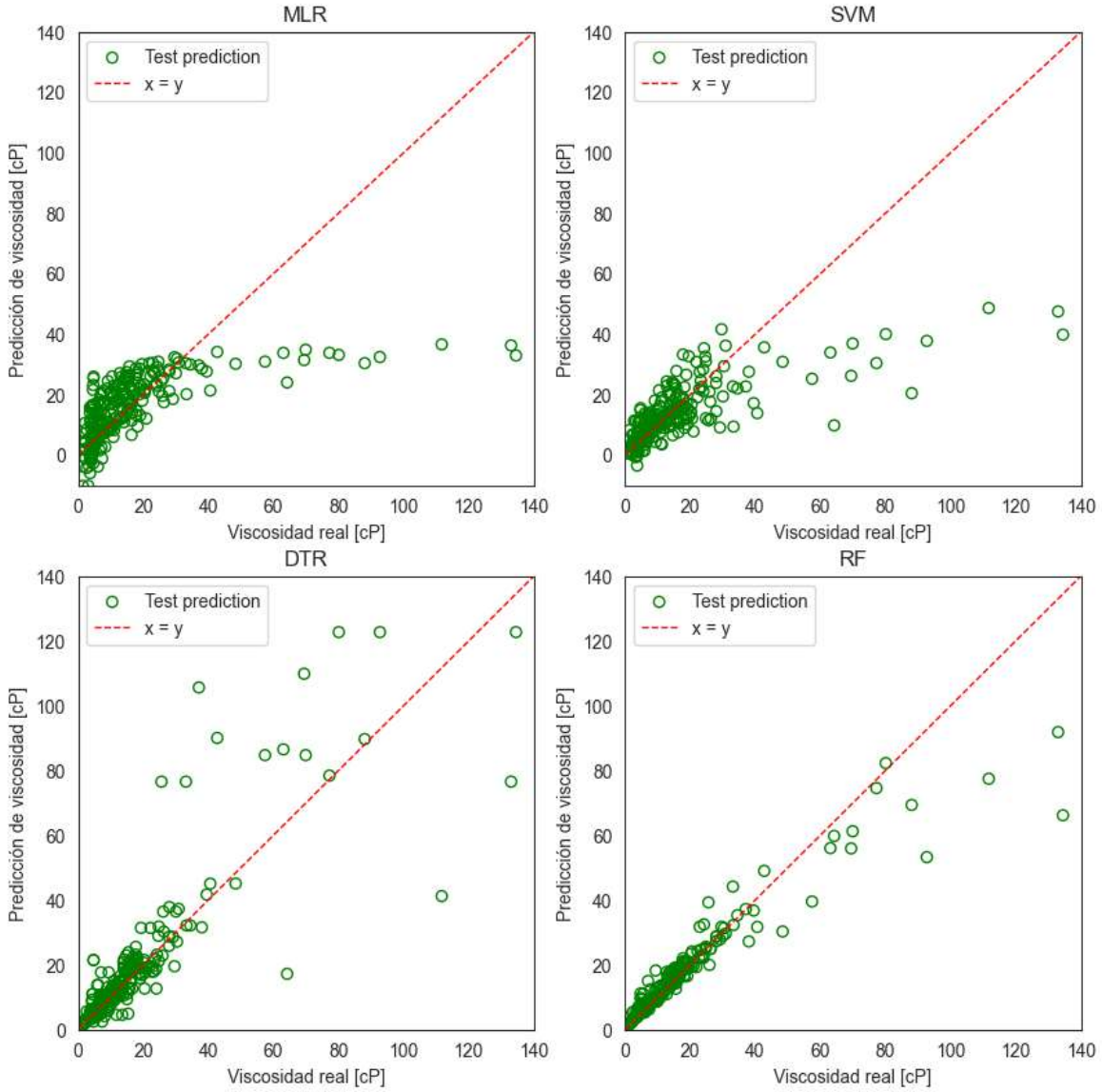

Figura 5-8. Comparación de valores reales y predichos para los modelos implementados

Por otro lado, en la **Figura 5-9** las curvas de valores residuales muestran la desviación de las predicciones en función de los valores de viscosidad real. Este gráfico permite demostrar con claridad las regiones donde se presentan mayores diferencias entre el valor predicho y el valor de viscosidad real. Para los modelos MLR y SVM se confirman los límites de predicción cercanos a 40 y 25 cP donde adicionalmente se evidencia una tendencia clara a subestimar las viscosidades comparado con el escenario real mientras que los modelos DTR y RF presentan residuales más bajos y tendencias a subestimar y sobreestimar respectivamente. Considerando el rango de bajas viscosidades los modelos DTR y RF presentan una menor dispersión confirmando el ajuste para el conjunto donde tiende a concentrase la mayor proporción de los datos.

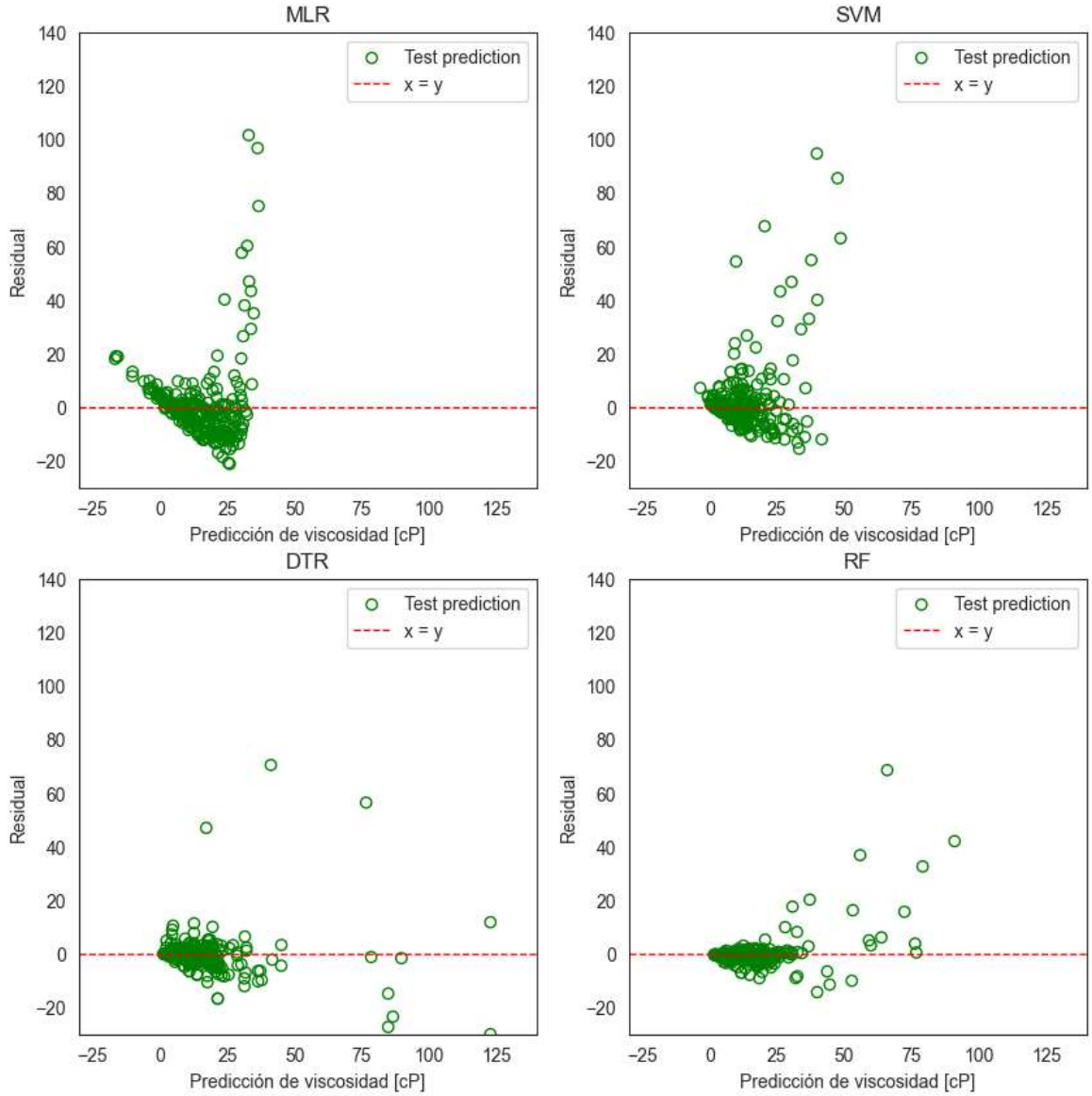

Figura 5-9. Valores residuales para los modelos implementados

Los resultados obtenidos permiten establecer el desempeño general de los modelos cuando se implementan con conjuntos de datos desbalanceados y donde típicamente se presenta una baja correlación entre las variables incluidas. El modelo DTR mejoró el desempeño de los modelos MLR y SVM incrementando el margen de predicción hasta 125 cP, sin embargo, los residuales calculados para este modelo permitieron identificar una tendencia a sobreestimar y a generar una predicción dispersa alrededor de los valores de viscosidad baja comparado con el modelo RF. Adicionalmente, este modelo presenta residuales que indican una tendencia a la subestimación reduciendo la dispersión de la predicción para bajas viscosidades. Si bien el modelo RF mejora la predicción hasta un R^2 de 0.87, las predicciones del modelo se ven limitadas hasta viscosidades cercanas a 80

cP asociadas a la baja proporción los datos de alta viscosidad en el conjunto de datos generado. Esta limitación del modelo puede ser abordada robusteciendo el conjunto de datos experimental a través de pruebas que permitan mejorar la capacidad de generalización del modelo.

Para establecer el desempeño de los modelos de aprendizaje bajo condiciones reales de entrenamiento, las curvas de aprendizaje se presentan en la **Figura 5-10** para los modelos DTR y RF con las métricas MAE y RMSE. Los resultados indican un sobreentrenamiento para el modelo DTR dado que se presenta un mejoramiento gradual en el error para la curva de entrenamiento, sin embargo, para la etapa de validación, el modelo presenta una alta oscilación y tiende a incrementar el error posterior al entrenamiento con 400 datos. El modelo RF por el contrario presenta un comportamiento similar en la curva de entrenamiento y una disminución estable en el error para la curva de validación indicando un aprendizaje estable con una disminución importante en la diferencia entre el error de entrenamiento y validación. Este resultado indica la posibilidad del modelo de bosques aleatorios para implementarse en aplicaciones reales donde el conjunto de datos presenta un alto desbalanceo y una proporción importante de valores fuera de la distribución media.

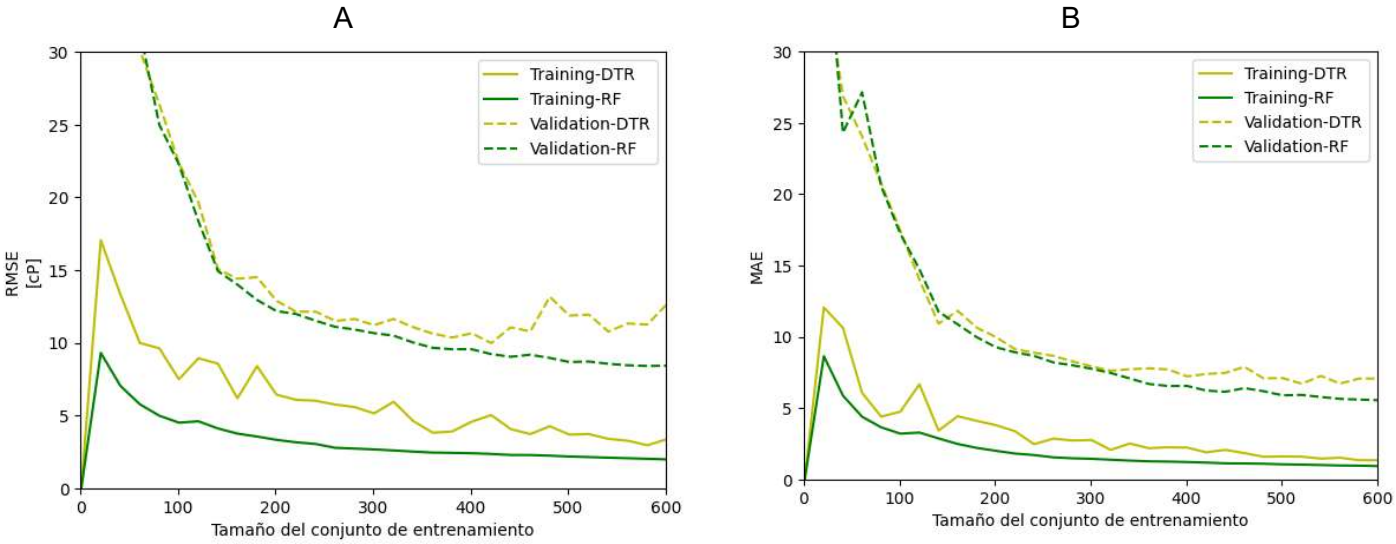

Figura 5-10. Curvas de aprendizaje para los modelos DTR y RF con las métricas **A)** RMSE y **B)** MAE

Finalmente, en la tabla **Tabla 5-3** se presentan los hiperparámetros resultantes del proceso de sintonización para cada modelo implementado. Resulta interesante la combinación final de hiperparámetros obtenida donde se establece que el modelo de bosques aleatorios presenta una mejora en el desempeño comparado con los árboles de decisión estableciendo la efectividad de los métodos de ensamble para este tipo de problemas de predicción con alto desbalanceo de datos. La efectividad de los métodos de ensamble por encima de los modelos convencionales puede asociarse a la etapa inicial de *bootstrapping* que permite exponer al modelo de aprendizaje a los datos atípicos presentes en el conjunto

de datos mejorando la respuesta ante los valores de alta viscosidad que, si bien son importantes para el fenómeno físico, su participación en el conjunto de datos es baja.

Tabla 5-3. Hiperparámetros resultantes de la sintonización de los modelos

Modelo	Hiperparámetro	Valores
MLR	Ajuste al intercepto	False
	Positivo	False
SVM	Kernel	poly
	Grado	7
	Gamma	scale
	Coef0	5
	C	3
DTR	Criterio	poisson
	Divisor	best
	Profundidad máxima	30
	Partición mínima	3
	Mínimo muestras por hoja	1
	Características máximas	sqrt
RF	Cantidad de estimadores	1000
	Criterio	friedman mse
	Profundidad máxima	30
	Partición mínima	2
	Mínimo muestras por hoja	1
	Características Máximas	sqrt

Los hiperparámetros presentados en la **Tabla 5-3** muestran que los árboles de decisión implementados tienen una configuración similar a los árboles cuando se implementan como método de ensamble en el modelo de bosques aleatorios con diferencias puntuales en la partición mínima y el criterio de evaluación de la partición. Para el modelo de bosques aleatorios se priorizó el criterio f*riedman mse* del tipo L2 donde el efecto de los valores atípicos se ve magnificado ajustando el modelo a la condición de desbalanceo presentada.

6. Conclusiones y trabajo futuro

6.1 Conclusiones

La predicción de viscosidad de nanofluidos poliméricos resulta de alta importancia para extender y garantizar la aplicabilidad de la alternativa en presencia de CQDs como trazador en el medio poroso. Este trabajo permitió caracterizar diferentes tipos de CQD y su efecto en las propiedades de flujo del nanofluido polimérico a base de HPAM. Como resultado de la evaluación de las propiedades de flujo del nanofluido fue conformado un conjunto de datos representativo con el fin de implementar modelos de aprendizaje de máquinas teniendo como variables principales parámetros de salinidad, orden de adición, tiempo de añejamiento, tipo de CQD y tasa de corte.

La evaluación inicial permitió establecer un efecto claro de la salinidad y la presencia de CQDs en la viscosidad del nanofluido polimérico confirmando que los escenarios de alta salinidad y presencia de CQDs afectan el polímero y facilitan el proceso de degradación en el tiempo. Este proceso se asocia principalmente a la interacción de los CQDs con las cadenas poliméricas donde se posicionan e irrumpen con la red conformada reduciendo su capacidad para mantener la viscosidad ante un escenario de alta temperatura durante un periodo de tiempo prolongado.

El conjunto de datos presentó una distribución desbalanceada con un 75% de los datos por debajo del valor de 20 cP de viscosidad y el 25% restante con viscosidades variando entre 20 y 134 cP. Esta condición de desbalanceo dificulta la tarea de predicción para los modelos, sin embargo, las condiciones de valoración experimental donde los datos no abundan son típicas y posicionan la alternativa para la extrapolación de resultados a otros escenarios no evaluados. El análisis de correlación de Pearson permitió establecer la baja correlación lineal entre las variables continúas brindando un indicio del desempeño de los modelos en función de la correlación de los datos.

Los modelos de aprendizaje de máquinas presentaron rendimientos variados siendo MLR y SVM los modelos con valores de error más altos asociados a la naturaleza no lineal de los parámetros, el desbalanceo y distribución del conjunto de datos. Contrario a este desempeño, los modelos DTR y RF se ajustaron a los datos bajo las condiciones de distribución presentados teniendo diferencias claras entre la distribución de las predicciones. Para el caso del modelo RF las predicciones tienden a estar subestimadas cuando se alcanzan valores superiores a 80 cP asociado a que en el conjunto de datos construido se presenta una baja proporción de datos de alta viscosidad. Por otro lado, el modelo DTR presenta una distribución de las predicciones cercanas a los valores reales y mantiene las predicciones de alta viscosidad cercanas a la referencia con una tendencia ligera a la sobreestimación. El mejor desempeño fue obtenido con el modelo de bosques aleatorios asociado posiblemente a la etapa inicial de bootstraping que permite diversificar

la muestra de datos alimentada a cada árbol de decisión. Estos resultados posicionan los métodos de ensamble como herramientas de alto potencial cuando el conjunto de datos presenta un alto desbalanceo y los valores atípicos registrados son relevantes para el fenómeno físico.

Si bien los modelos de aprendizaje tuvieron predicciones cercanas a los valores reales, resulta clara la necesidad de robustecer el conjunto de datos en función de parámetros fisicoquímicos para mejorar las métricas de error y la condición de desbalanceo de los datos. Adicionalmente, la variación en la concentración de CQDs y HPAM en la solución podría incrementar la capacidad de generalización de los modelos ampliando la capacidad de extrapolación para escenarios de evaluación de EOR con polímeros.

De acuerdo con la revisión sistemática de literatura desarrollada y los acercamientos de este trabajo de investigación resulta evidente la posibilidad de ampliar la capacidad de predicción de los modelos de aprendizaje de máquinas a través de evaluaciones experimentales. Este acercamiento permite incluso considerar el despliegue de estas herramientas bajo condiciones relevantes de campo manteniendo un desarrollo constante del conjunto de datos para mejorar la capacidad de generalización de los modelos desde las evaluaciones experimentales hasta la escala de producción real.

6.2 Trabajo futuro

Con los resultados obtenidos en este trabajo de investigación se pueden establecer algunas líneas de investigación para trabajos posteriores:

- Evaluación robusta de las condiciones de adición, tipo de CQD, tipos de iones y salinidad presente en las soluciones poliméricas variando la concentración de CQDs y HPAM. Este acercamiento permite definir con claridad el efecto puntual de los CQDs en presencia de iones bajo diferentes esquemas de adición.
- Ampliación y diversificación del conjunto de datos para mejorar la precisión de los modelos de aprendizaje de máquinas. Si bien el modelo de árboles aleatorios alcanzó un buen desempeño prediciendo la viscosidad de los nanofluidos, resulta necesario incrementar el tamaño del conjunto de datos para mejorar la base de entrenamiento para la predicción. La ampliación del conjunto de datos permitiría mejorar la capacidad de generalización del modelo.
- Evaluar el efecto de los CQDs en nanofluidos poliméricos a través de técnicas de valoración estáticas que permitan direccionar los estudios a aplicaciones de recobro real, por ejemplo, evaluando tensiones interfaciales para los nanofluidos armados bajo diferentes esquemas de conformación. Extender estos estudios

hacia pruebas dinámicas permitirían establecer las mejoras en la cantidad de aceite recuperado bajo condiciones representativas del medio poroso.

- La biodegradación de soluciones poliméricas, especialmente de HPAM, se ha posicionado como una herramienta relevante principalmente para el tratamiento de efluentes posterior a los procesos de recobro mejorado con polímeros. La aplicación de comunidades biológicas se estudia principalmente posterior a la producción como alternativa de recuperación del recurso hídrico. Algunos estudios muestran los CQDs como una alternativa promisoria para la eliminación de contaminantes que en presencia de comunidades biológicas puede potenciar los efectos de biorremediación. La biodegradación de soluciones poliméricas bajo condiciones de yacimiento en presencia de CQDs presenta una oportunidad de investigación relevante considerando este tipo de nanomateriales pueden participar inclusive como inhibidores del crecimiento microbiológico.

7. Referencias

[1] R. C. Paniagua, Duarte, A., "Generación eléctrica en Colombia y su transición hacia Fuentes Renovables No Convencionales " Corficolombiana, https://investigaciones.corfi.com/renta-variable/informe-especial/generacion-electrica-en-colombia-y-su-transicion-hacia-fuentes-renovables-no-convencionales/informe_7962602021.

[2] T. Castillo, Garcia, F., Mosquera, L., Rivadeneira, T., Segura, K., Yujato, M., "Panorama Energético de America Latina y El Caribe," Organización Latinoamericana de Energía, olade.org2021.

[3] UPME, "Proyección de la Demanda de Energía Electrica, Gas Natural y Combustibles Líquidos 2022-2036," UPME2021.

[4] L. Quiroga, "¿Qué es el recobro mejorado? La tecnología para 'raspar la olla' del petróleo," in *EL TIEMPO* ed. https://www.eltiempo.com/economia/sectores/petroleo-que-es-recobro-mejorado-la-tecnologia-mencionada-por-minminas-722449: EL TIEMPO 2022.

[5] A. López, "Así rasparán la olla para 100 mil barriles más de petróleo," in *El Colombiano*, ed. https://www.elcolombiano.com/negocios/asi-rasparan-la-olla-para-100-mil-barriles-mas-de-petroleo-EJ19543869: El Colombiano, 2022.

[6] B. A. Suleimanov, F. S. Ismayilov, O. A. Dyshin, and E. F. Veliyev, "Selection methodology for screening evaluation of EOR methods," *Petroleum Science and Technology,* vol. 34, pp. 961-970, 2016.

[7] M. Gutiérrez, R. Castro, L. Corredor, F. Fernández, J. Zapata, J. Jimenez*, et al.*, "Chemical Enhanced Oil Recovery Experiences in Colombia: Field Pilots Review," in *SPE Improved Oil Recovery Conference?*, 2024, p. D031S019R003.

[8] Ø. Dugstad, S. Viig, B. Krognes, R. Kleven, and O. Huseby, "Tracer monitoring of enhanced oil recovery projects," in *EPJ Web of Conferences*, 2013, p. 02002.

[9] C. A. Franco, C. H. Candela, J. Gallego, J. Marin, L. E. Patiño, N. Ospina*, et al.*, "Easy and rapid synthesis of carbon quantum dots from Mortino (Vaccinium Meridionale Swartz) extract for use as green tracers in the oil and gas industry: Lab-to-field trial development in Colombia," *Industrial & Engineering Chemistry Research,* vol. 59, pp. 11359-11369, 2020.

[10] M. Amiri, M. Fatemi, and E. B. Delijani, "Effect of brine salinity and hydrolyzed polyacrylamide concentration on the Oil/Brine and Brine/Rock Interactions: Implications on enhanced oil recovery by hybrid low salinity polymer flooding in sandstones," *Fuel,* vol. 324, p. 124630, 2022.

[11] S. Jouenne, "Polymer flooding in high temperature, high salinity conditions: Selection of polymer type and polymer chemistry, thermal stability," *Journal of Petroleum Science and Engineering,* vol. 195, p. 107545, 2020.

[12] S. Kumar, R. Tiwari, M. Husein, N. Kumar, and U. Yadav, "Enhancing the performance of HPAM polymer flooding using nano CuO/nanoclay blend," *Processes,* vol. 8, p. 907, 2020.

[13] R. Seright and D. Wang, "Polymer flooding: Current status and future directions," *Petroleum Science,* vol. 20, pp. 910-921, 2023.

[14] G. Yao, J. Zhao, M. A. Haruna, and D. Wen, "Molecular dynamics insight into viscosity reduction of hydrolysed polyacrylamide by using carbon quantum dots," *RSC advances,* vol. 11, pp. 26037-26048, 2021.

[15] M. A. Haruna, Z. Hu, H. Gao, J. Gardy, S. M. Magami, and D. Wen, "Influence of carbon quantum dots on the viscosity reduction of polyacrylamide solution," *Fuel,* vol. 248, pp. 205-214, 2019.

[16] M. Zhao, Y. Li, C. Dai, Y. Chen, Z. Yang, K. Liu*, et al.*, "Development of Novel Silicon Quantum Dots and Their Potential in Improving the Enhanced Oil Recovery of HPAM," *Langmuir,* vol. 40, pp. 3181-3189, 2024.

[17] M. Abdullah, M. Delshad, K. Sepehrnoori, M. Balhoff, J. Foster, and M. Al-Murayri, "Physics-Based and Data-Driven Polymer Rheology Model," *SPE Journal,* vol. 28, pp. 1857-1879, 2023.

[18] E. Pérez, D. Alviso, E. Manrique, and G. Artana, "Estimation of the rheological curve of HPAM solutions from measurements using the Brookfield viscometer," *Journal of Petroleum Science and Engineering,* vol. 216, p. 110793, 2022.

[19] M. Zeynalli, E. W. Al-Shalabi, and W. AlAmeri, "An Extended Viscoelastic Model for predicting polymer apparent viscosity at different shear rates," in *SPE Annual Technical Conference and Exhibition?*, 2021, p. D011S019R005.

[20] V. Karpan, S. Al Farsi, H. Al Sulaimani, D. Al Mahrouqi, R. Al Mjeni, and D. van Batenburg, "Introducing data-driven virtual viscosity measurements," in *Abu Dhabi International Petroleum Exhibition and Conference*, 2022, p. D031S071R001.

[21] N. Lashari, T. Ganat, D. Otchere, S. Kalam, and I. Ali, "Navigating viscosity of GO-SiO2/HPAM composite using response surface methodology and supervised machine learning models," *Journal of Petroleum Science and Engineering,* vol. 205, p. 108800, 2021.

[22] R. Maillon, P. Ghosh, and C. Wilson, "Predicting EOR Polymer Viscosity with Varying Molecular Weight and Chemistry Using Machine Learning," in *Abu Dhabi International Petroleum Exhibition and Conference*, 2023, p. D031S090R002.

[23] M. Shakeel, P. Pourafshary, M. R. Hashmet, and R. Muneer, "Application of machine learning techniques to predict viscosity of polymer solutions for enhanced oil recovery," *Energy Systems,* pp. 1-24, 2023.

[24] B. R. Cancela, L. C. M. Palermo, P. F. de Oliveira, and C. R. Mansur, "Rheological study of polymeric fluids based on HPAM and fillers for application in EOR," *Fuel,* vol. 330, p. 125647, 2022.

[25] A. A. R. Lady J. Giraldo, "Evaluation of addition of Carbon Quantum Dots (CQDs) on Polymeric solution for EOR processes: Effect of CQD type and Ion content on brine," presented at the X SUMMER SCHOOL: MEETING SOUTH-SOUTH-NORTH. SUBSURFACE RESOURCES IN THE CONTEXT OF THE ENERGY TRANSITION, Universidad Nacional de Colombia, Medellín, Facultad de Minas, 2024.

[26] A. A. R. Lady J. Giraldo, Esteban A. Taborda, Farid B. Cortés, Camilo A. Franco, "EVALUACIÓN DE LA ADICIÓN DE PUNTOS CUANTICOS DE CARBONO (CQDs) EN SOLUCIONES POLIMERICAS PARA PROCESOS DE RECOBRO MEJORADO (EOR): EFECTO DEL TIPO DE CQD Y CONTENIDO DE IONES EN SALMUERA," presented at the XXI CONGRESO COLOMBIANO DEL PETROLEO, GAS Y ENERGÍA: Impulsando el Progreso con Energía e Innovación, Barranquilla, Colombia, 2024.

[27] J. F. Gallego, L. J. Giraldo, H. I. Quintero, H. A. García, K. Zapata, S. Heidari*, et al.*, "Effect of monovalent/divalent ions and SiO2-based nanocomposite dosage on thermochemical stability of HPAM polymeric solutions," *Petroleum,* 2024.

[28] Q. Ma, P. J. Shuler, C. W. Aften, and Y. Tang, "Theoretical studies of hydrolysis and stability of polyacrylamide polymers," *Polymer degradation and stability,* vol. 121, pp. 69-77, 2015.

[29] A. M. Firozjaii and H. R. Saghafi, "Review on chemical enhanced oil recovery using polymer flooding: Fundamentals, experimental and numerical simulation," *Petroleum,* vol. 6, pp. 115-122, 2020.

[30] G. Rotko, E. Knapik, M. Piotrowski, and M. Marszałek, "Oilfield Brine as a Source of Water and Valuable Raw Materials—Proof of Concept on a Laboratory Scale," *Water,* vol. 16, p. 1461, 2024.

[31] W. Luo, S. Xu, and F. Torabi*, "Chemical degradation of HPAM by oxidization in produced water: experimental study," in *SPE Health, Safety, Security, Environment, & Social Responsibility Conference-North America*, 2013, pp. SPE-163751-MS.

[32] K. G. Salem, A. M. Salem, M. A. Tantawy, A. A. Gawish, S. Gomaa, and A. El-hoshoudy, "A Comprehensive Investigation of Nanocomposite Polymer Flooding at Reservoir Conditions: New Insights into Enhanced Oil Recovery," *Journal of Polymers and the Environment,* pp. 1-21, 2024.

[33] D. Tayyib, A. Al-Qasim, S. Kokal, and O. Huseby, "Overview of tracer applications in oil and gas industry," in *SPE Kuwait oil and gas show and conference*, 2019, p. D023S006R004.

[34] A. K. Patidar, D. Joshi, U. Dristant, and T. Choudhury, "A review of tracer testing techniques in porous media specially attributed to the oil and gas industry," *Journal of Petroleum Exploration and Production Technology,* vol. 12, pp. 3339-3356, 2022.

[35] C. Carpenter, "Use of partitioning tracers to estimate oil-saturation distribution," *Journal of Petroleum Technology,* vol. 68, pp. 72-73, 2016.

[36] S. Rosales, K. Zapata, F. B. Cortes, B. Rojano, C. Diaz, C. Cortes*, et al.*, "Simultaneous Detection of Carbon Quantum Dots as Tracers for Interwell Connectivity Evaluation in a Pattern with Two Injection Wells," *Nanomaterials,* vol. 14, p. 789, 2024.

[37] R. Das, R. Bandyopadhyay, and P. Pramanik, "Carbon quantum dots from natural resource: A review," *Materials today chemistry,* vol. 8, pp. 96-109, 2018.

[38] J. Durrani, "Explainer: why have quantum dots won the 2023 Nobel prize?," in *Chemistry World*, ed. https://www.chemistryworld.com/news/explainer-why-have-quantum-dots-won-the-2023-nobel-prize/4018168.article: Royal Society of Chemistry 2023.

[39] M. Razavifar, A. Khoshsima, M. Riazi, J. J. Sheng, and E. Esmaeilnezhad, "Recent developments, challenges, and prospects of carbon dots (CDs) for fluid flow investigation in porous media," *Petroleum Research,* 2024.

[40] X. Zhao, S. Chen, L. Zhang, H. Su, Y. Wu, and H. Ren, "Influence of carbon quantum dots on foam stability: a molecular dynamics study," in *Journal of Physics: Conference Series*, 2024, p. 012043.

[41] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma*, et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of big Data,* vol. 8, pp. 1-74, 2021.

[42] J. Jobson and J. Jobson, "Multiple linear regression," *Applied multivariate data analysis: regression and experimental design,* pp. 219-398, 1991.

[43] V. Vapnik, *The nature of statistical learning theory*: Springer science & business media, 2013.

[44] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and their applications,* vol. 13, pp. 18-28, 1998.

[45] G. Gordon, "Support vector machines and kernel methods," *Online]. Dosegljiv: https://www. cs. cmu. edu/~ ggordon/SVMs/new-svms-andkernels. pdf [Dostopano 28. 6. 2016],* 2004.

[46] Y.-Y. Song and L. Ying, "Decision tree methods: applications for classification and prediction," *Shanghai archives of psychiatry,* vol. 27, p. 130, 2015.

[47] O. C. Njoku, "Decision trees and their application for classification and regression problems," 2019.

[48] L. Breiman, "Random forests," *Machine learning,* vol. 45, pp. 5-32, 2001.

[49] G. Biau and E. Scornet, "A random forest guided tour," *Test,* vol. 25, pp. 197-227, 2016.

[50] S. J. Rigatti, "Random forest," *Journal of Insurance Medicine,* vol. 47, pp. 31-39, 2017.

[51] T. Erfando, G. Siregar, Y. F. Alli, and Y. B. Wangge, "Hydrolyzed Polyacrylamide Screening for EOR in an Indonesia Mature Oil Field," *Trends in Sciences,* vol. 19, pp. 3461-3461, 2022.

[52] H. A. Yousefvand and A. Jafari, "Stability and flooding analysis of nanosilica/NaCl/HPAM/SDS solution for enhanced heavy oil recovery," *Journal of Petroleum Science and Engineering,* vol. 162, pp. 283-291, 2018.

[53] B. Choi, M. S. Jeong, and K. S. Lee, "Temperature-dependent viscosity model of HPAM polymer through high-temperature reservoirs," *Polymer degradation and stability,* vol. 110, pp. 225-231, 2014.

[54] T. Duriez, S. L. Brunton, and B. R. Noack, *Machine learning control-taming nonlinear dynamics and turbulence* vol. 116: Springer, 2017.

[55] A. R. Shaik, W. AlAmeri, A. AlSumaiti, M. Muhammad, and N. C. Thomas, "Application of supervised machine learning technique to investigate the effect of brine hardness on polymer bulk rheology," in *Abu Dhabi International Petroleum Exhibition and Conference*, 2019, p. D022S149R001.

[56] N. Basha, R. Arcucci, P. Angeli, C. Anastasiou, T. Abadie, C. Q. Casas*, et al.*, "Machine learning and physics-driven modelling and simulation of multiphase systems," *International Journal of Multiphase Flow,* vol. 179, p. 104936, 2024.

[57] M. Lellep, J. Prexl, B. Eckhardt, and M. Linkmann, "Interpreted machine learning in fluid dynamics: explaining relaminarisation events in wall-bounded shear flows," *Journal of Fluid Mechanics,* vol. 942, p. A2, 2022.

[58] S. Rushd, U. Gazder, H. J. Qureshi, and M. Arifuzzaman, "Advanced machine learning applications to viscous oil-water multi-phase flow," *Applied Sciences,* vol. 12, p. 4871, 2022.

[59] B. D. Wood, X. He, and S. V. Apte, "Modeling turbulent flows in porous media," *Annual Review of Fluid Mechanics,* vol. 52, pp. 171-203, 2020.

[60] J. Zhang, U. Braga-Neto, and E. Gildin, "Physics-Informed Neural Networks for Multiphase Flow in Porous Media Considering Dual Shocks and Interphase Solubility," *Energy & Fuels,* 2024.

[61] L.-T. Zhu, X.-Z. Chen, B. Ouyang, W.-C. Yan, H. Lei, Z. Chen*, et al.*, "Review of machine learning for hydrodynamics, transport, and reactions in multiphase flows and reactors," *Industrial & Engineering Chemistry Research,* vol. 61, pp. 9901-9949, 2022.

[62] H. Khakurel, M. Taufique, A. Roy, G. Balasubramanian, G. Ouyang, J. Cui*, et al.*, "Machine learning assisted prediction of the Young's modulus of compositionally complex alloys," *Scientific reports,* vol. 11, p. 17149, 2021.

[63] K. M. Kim, P. Hurley, and J. P. Duarte, "Physics-informed machine learning-aided framework for prediction of minimum film boiling temperature," *International Journal of Heat and Mass Transfer,* vol. 191, p. 122839, 2022.

[64] J. Li, M. Zhou, H. H. Wu, L. Wang, J. Zhang, N. Wu*, et al.*, "Machine Learning-Assisted Property Prediction of Solid-State Electrolyte," *Advanced Energy Materials,* p. 2304480, 2024.

[65] H. A. Ritacco, "Foam-Assisted Oil Recovery: A Physics-Based Perspective," *Current Opinion in Colloid & Interface Science,* p. 101809, 2024.

[66] T. W. Thibodeaux, *Image-based Modeling of Flow through Porous Media: Development of Multiscale Techniques for the Pore Level*: Louisiana State University and Agricultural & Mechanical College, 2018.

[67] R. S. Seright, A. Campbell, P. Mozley, and P. Han, "Stability of partially hydrolyzed polyacrylamides at elevated temperatures in the absence of divalent cations," *SPE journal,* vol. 15, pp. 341-348, 2010.

[68] R. Nisbet, J. Elder, and G. D. Miner, *Handbook of statistical analysis and data mining applications*: Academic press, 2009.

[69] J. J. Berman, *Data simplification: taming information with open source tools*: Morgan Kaufmann, 2016.

[70] A. Jierula, S. Wang, T.-M. Oh, and P. Wang, "Study on accuracy metrics for evaluating the predictions of damage locations in deep piles using artificial neural networks with acoustic emission data," *Applied Sciences,* vol. 11, p. 2314, 2021.

[71] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *Peerj computer science,* vol. 7, p. e623, 2021.

[72] E. E. Ateia, O. Rabie, and A. T. Mohamed, "Assessment of the correlation between optical properties and CQD preparation approaches," *The European Physical Journal Plus,* vol. 139, p. 24, 2024.

[73] M. D. Malitha, M. T. H. Molla, M. A. Bashar, D. Chandra, and M. S. Ahsan, "Fabrication of a reusable carbon quantum dots (CQDs) modified nanocomposite with enhanced visible light photocatalytic activity," *Scientific Reports,* vol. 14, p. 17976, 2024.

[74] L. Cheng, Y. Qin, Y. Su, Y. Pan, Y. Wang, R. Liao*, et al.*, "Development of a high-strength and adhesive polyacrylamide gel for well plugging," *ACS omega,* vol. 7, pp. 6151-6159, 2022.