

Convergencia estratégica de Redes Neuronales de Grafos, REROSIM y Scaled Load Ratio para gemelos digitales de inyección de vapor

Strategic convergence of Graph Neural Networks, REROSIM and Scaled Load Ratio for steam-injection digital twins

DOI: <https://doi.org/10.5281/zenodo.21810662>

Recibido: 2026-05-17 Aceptado: 2026-06-25

Romero, Isbelis¹

Correo: isbelisromeroucv@gmail.com

Orcid: <https://orcid.org/0009-0001-5107-9535>

Resumen

Este artículo propone un modelo de gemelo digital para optimizar la gestión de yacimientos bajo inyección de vapor, integrando tres metodologías avanzadas de la literatura SPE: *Redes Neuronales de Grafos (GNN)*: Adaptadas del almacenamiento de CO₂ para actuar como sustitutos físicos rápidos, modelando la evolución de presión, temperatura, saturaciones y calidad de vapor. Método *REROSIM* (Reservoir Engineering Reservoir Simulation): Abstrae la conectividad inter-pozo para evaluar la dinámica del barrido sin depender de simulaciones de grilla fina. *Analítica de Scaled Load Ratio (SLR)*: Evalúa la integridad operacional del bombeo mecánico, previniendo fallas por fricción, fatiga o interferencia de gas. El enfoque unifica el subsuelo con la superficie. Al combinar física, velocidad computacional y vigilancia de integridad, la arquitectura permite ejecutar escenarios en minutos y aprender del dato histórico, garantizando una optimización que protege la confiabilidad mecánica del sistema.

Palabras clave: Redes Neuronales de Grafos, REROSIM, Conectividad inter-pozo, Scaled Load Ratio.

Abstract

This article proposes a digital twin model to optimize reservoir management under steam injection, integrating three advanced methodologies from SPE literature: Graph Neural Networks (GNNs)—adapted from CO₂ storage applications to serve as rapid physics-based surrogates that model the evolution of pressure, temperature, saturations, and steam quality; the REROSIM method—which abstracts inter-well connectivity to evaluate sweep dynamics without relying on fine-grid simulations; and Scaled Load Ratio (SLR) analytics—which assess the operational integrity of mechanical pumping systems, preventing failures caused by friction, fatigue, or gas interference. This approach unifies subsurface and surface operations. By combining physics, computational speed, and integrity monitoring, the architecture enables the execution of scenarios in minutes and facilitates learning from historical data, ensuring optimization that safeguards the system's mechanical reliability.

¹ Ingeniero de Petróleo. Petróleos & Renovables S.A. Venezuela.



Keywords: Graph Neural Networks, REROSIM, Interwell connectivity, Scaled Load Ratio, mechanical integrity.

Introducción

Los activos maduros de crudo pesado y extrapesado enfrentan una combinación de retos técnicos que, rara vez, se resuelve con una sola herramienta: heterogeneidad geológica, movilidad desfavorable, alta relación agua-petróleo, producción cíclica, restricciones de energía, riesgos de integridad y necesidad de justificar cada intervención con evidencias de valor. La inyección de vapor, por su capacidad para reducir la viscosidad del crudo y movilizar volúmenes remanentes, sigue siendo una de las alternativas de recuperación mejorada más relevantes en yacimientos térmicos. Sin embargo, su éxito depende de ubicar el calor donde realmente genera barrido, evitar canales preferenciales, controlar pérdidas térmicas y no sobrecargar los sistemas de producción artificial.

En los últimos años, la digitalización de yacimientos ha pasado de dashboards descriptivos hacia modelos predictivos que asimilan datos, simulan escenarios y proponen decisiones. La literatura técnica muestra tres direcciones complementarias. La primera es el uso de redes neuronales de grafos para representar yacimientos como grafos donde los nodos concentran propiedades de control de volumen y las aristas codifican adyacencia, transmisibilidad y capacidad de flujo. La segunda es REROSIM, que reduce la complejidad del yacimiento a una red de conexiones inyector-productor calibrada con datos de producción e inyección. La tercera es SLR, una métrica simple, normalizada y operativamente accesible para monitorear la condición de bombas de varillas desde cargas superficiales, sin reconstruir cartas de fondo mediante una ecuación de onda (Tariq, 2025).

La convergencia de estas líneas permite formular un gemelo digital térmico de nueva generación. En lugar de depender, únicamente, de simuladores numéricos de alta fidelidad, que continúan siendo indispensables para validación y diseño base, se propone una jerarquía híbrida: un GNN entrenado con casos físicos de inyección de vapor para estimar campos de presión, temperatura y saturación; un REROSIM térmico para traducir esas respuestas a lógica inter-pozo accionable; y un módulo SLR para asegurar que los cambios de vapor, producción o levantamiento artificial no comprometan la continuidad operacional. Esta integración no pretende eliminar la ingeniería convencional; busca organizarla en un flujo más ágil y útil para decisiones diarias.





Se desarrolla como artículo de síntesis aplicada; por tanto, no presenta un caso de campo propio ni resultados de simulación originales, sino una metodología estratégica que puede convertirse en programa piloto para campos térmicos de Venezuela, Latinoamérica o análogos internacionales.

1. Fundamentos teóricos

La inyección de vapor modifica simultáneamente movilidad del petróleo, balance de energía, presión local, saturación de agua caliente, estabilidad de los equipos de subsuelo y vida útil de los sistemas de levantamiento. En condiciones ideales, el vapor se distribuye hacia zonas con petróleo remanente móvil, transfiere calor de manera eficiente y retorna en forma de incremento de producción neta. En campos reales, el comportamiento suele ser más complejo: una fracción del vapor se canaliza por zonas de alta permeabilidad, otra se pierde hacia intervalos ya barridos, parte se condensa antes de llegar al frente útil y, algunos pozos productores, experimentan cambios abruptos de carga, emulsionamiento, arenamiento o inestabilidad del bombeo.

Los modelos tradicionales de yacimientos ofrecen una representación detallada de la física, pero pueden requerir semanas de preparación, calibración y corridas de sensibilidad. Su costo computacional se vuelve más relevante cuando la pregunta operacional cambia todos los días: qué inyector debe recibir más vapor, qué productor necesita reducción de velocidad, qué patrón muestra canalización térmica, qué pozo debe priorizarse por integridad y cuál escenario genera mayor valor con el menor consumo de energía. En un campo con decenas o cientos de pozos, la frecuencia de decisión excede la capacidad de ejecutar estudios completos para cada ajuste (Shaikh, 2025).

La propuesta de convergencia parte de una premisa: el yacimiento térmico debe ser leído como una red física y operativa. A escala de malla, el flujo y el calor obedecen ecuaciones de conservación que pueden aprenderse mediante GNN entrenadas con simulaciones de alta fidelidad. A escala de patrón, la comunicación efectiva entre inyectores y productores puede aproximarse con REROSIM, que sintetiza conectividad, volumen conectado y retardo. A escala de pozo, el desempeño mecánico puede vigilarse con SLR, transformando cargas superficiales en señales tempranas de falla. La estrategia se vuelve poderosa cuando estas capas no trabajan aisladas, sino dentro de un mismo ciclo de datos y decisión.





Esta integración también responde a una limitación frecuente de los proyectos de inteligencia artificial en petróleo: la desconexión entre modelos estadísticos y decisiones de ingeniería. Un algoritmo que predice producción sin explicar rutas de conectividad, sin respetar restricciones térmicas y sin evaluar la integridad del levantamiento puede ser rápido, pero no necesariamente confiable. Por eso, el marco propuesto no se basa en una caja negra única, sino en módulos especializados que conservan trazabilidad física y facilitan la revisión por equipos de geología, yacimientos, producción, operaciones, mantenimiento e integridad.

1.1. Redes neuronales de grafos para inyección de vapor

Las redes neuronales de grafos en la simulación de yacimientos resultan especialmente atractivas, para yacimientos, porque la roca y los fluidos no se comportan como píxeles independientes, sino como sistemas conectados por transmisibilidades, continuidad hidráulica y contrastes de propiedades. En un grafo de yacimiento, cada nodo puede representar un bloque de malla o una región agrupada con porosidad, permeabilidad, espesor, profundidad, facies, presión, temperatura y saturaciones. Cada arista representa vecindad, conductancia, distancia efectiva, contacto estratigráfico o conexión preferencial. El mensaje que viaja por las aristas resume cómo un cambio local se propaga hacia su entorno (Shaikh, 2025).

La aplicación de GNN al almacenamiento de CO₂ ha demostrado una idea transferible: los modelos pueden aprender la evolución espacio - temporal de presión y saturación sobre mallas regulares o irregulares si reciben suficientes casos de entrenamiento generados por simulación. En el caso de inyección de vapor, el aprendizaje debe ampliarse para incorporar energía. La red no sólo debe aproximar la distribución de fluido; también debe reconocer la transferencia de calor, condensación del vapor, pérdidas hacia capas adyacentes, disminución de viscosidad con temperatura y variación de movilidad relativa (Tariq, 2025).

La adaptación del enfoque GNN desde CO₂ a vapor requiere redefinir entradas y salidas. Las entradas estáticas incluyen propiedades de roca, límites estratigráficos, continuidad lateral, presencia de barreras y estado inicial. Las entradas dinámicas incorporan tasa de vapor, calidad de vapor, presión de inyección, entalpía, tiempo de remojo cuando aplique, tasa de producción, presión de fondo, temperatura registrada, eventos de cierre y restricciones de superficie. Las salidas relevantes serían campos de temperatura, presión, saturación de aceite, saturación de agua, condensado, calor acumulado, indicadores de barrido y probabilidad de canalización térmica.





La ganancia estratégica del GNN térmico no reside en reemplazar toda la simulación física, sino en acelerar las preguntas de alto volumen: sensibilidad a caudales, redistribución de vapor, comparación de patrones, estimación de impacto por cierre temporal o evaluación rápida de un nuevo ciclo. El simulador convencional seguiría generando casos de entrenamiento, validará escenarios críticos y resolvería decisiones de inversión. El GNN, en cambio, funcionaría como sustituto para el espacio de decisión diario y para optimización externa donde se evalúan miles de alternativas (Tariq y Shaikh, 2025).

Reformulación física: del CO₂ al vapor

La sustitución conceptual de CO₂ por vapor no es un cambio cosmético. El CO₂ en almacenamiento subterráneo, se analiza normalmente mediante flujo multifásico, presión, saturación de gas y salmuera, capilaridad y atrapamiento. El vapor introduce un problema termo-hidráulico: la variable temperatura modifica viscosidad del crudo, densidad de fases, movilidad, transferencia de masa por condensación y eficiencia de barrido. Además, la energía inyectada tiene un costo operativo directo, por lo que, un modelo útil debe pronosticar producción incremental y eficiencia térmica, no sólo movimiento de fluidos (Tissot y Welte, 1984).

El grafo térmico debe codificar aristas que representan tanto capacidad hidráulica como capacidad de transporte térmico. Dos bloques pueden tener buena comunicación de presión, pero mala transferencia de calor si existe una barrera vertical, alto enfriamiento o contacto limitado. Por esa razón, se propone una arista compuesta con atributos de transmisibilidad, distancia, contraste de permeabilidad, continuidad de facies, capacidad calorífica efectiva, conductividad térmica y pérdida estimada hacia rocas encajantes. Esta riqueza de atributos permite que el mensaje del GNN no sea una simple correlación espacial, sino una aproximación informada por la física.

Las funciones de pérdida del entrenamiento también deben cambiar. Para CO₂, el error en presión y saturación puede ser suficiente en muchas aplicaciones. Para vapor, conviene usar pérdidas ponderadas que penalicen errores de temperatura en zonas cercanas a productores, errores de presión en intervalos próximos a límites operativos y errores de saturación en regiones con petróleo remanente. También se puede añadir una penalización de balance de energía, obligando al modelo a respetar la relación entre calor inyectado, calor almacenado, calor producido y pérdidas.





Un elemento relevante es la estabilidad de largo horizonte. Las predicciones autoregresivas pueden acumular error cuando el GNN usa su propia salida como entrada para pasos posteriores. En inyección de vapor, ese error acumulado puede traducirse en frentes térmicos desplazados o sobreestimación de temperatura. Por ello, se recomienda entrenar con horizontes múltiples, incorporar datos de cierre y reinicio, usar aprendizaje curricular y contrastar las predicciones con puntos de control provenientes de registros de temperatura, pruebas de presión, trazadores térmicos y mediciones de producción (Tariq y Shaikh, 2025).

1.2. REROSIM como lógica inter-pozo y puente de interpretación

REROSIM ofrece una forma diferente de mirar el yacimiento: no como una malla con miles o millones de celdas, sino como una red de unidades de conexión entre inyectores y productores. Cada conexión contiene parámetros equivalentes de conductividad, volumen conectado, retardo y fracción de contribución. Este enfoque reduce la dimensionalidad y permite que los equipos operacionales visualicen la pregunta clave: qué inyector está afectando a qué productor, con qué intensidad y con qué tiempo de respuesta (Yan et al., 2025).

En campos sometidos a vapor, la conectividad inter-pozo tiene una expresión térmica. Un productor puede responder en presión sin recibir calor significativo; otro puede recibir agua caliente o vapor condensado por un canal de alta permeabilidad; un tercero puede mostrar incremento de temperatura antes que incremento de aceite, señal de barrido ineficiente o acercamiento de vapor vivo. REROSIM puede adaptarse a esta realidad incorporando no sólo tasas de inyección y producción, sino entalpía inyectada, temperatura de producción, corte de agua, respuesta de aceite, presión y eventos operacionales.

La matriz de conectividad térmica resultante permite clasificar patrones. Conexiones de alta fracción, bajo retardo y baja producción incremental sugieren reciclaje de calor o canalización. Conexiones de fracción media, retardo razonable y respuesta de aceite positiva indican barrido útil. Conexiones débiles pueden representar zonas no comunicadas, barreras o áreas con petróleo remanente que requieren rediseño de inyección. Esta lectura convierte el gemelo digital en una herramienta de priorización: no todos los inyectores merecen más vapor; algunos necesitan redistribución, control de perfil o reducción temporal (Yan et al., 2025).

La ventaja de REROSIM dentro de la convergencia es que traduce las predicciones de campo del GNN a relaciones inter-pozo comprensibles. El GNN puede estimar mapas de temperatura y presión, pero el operador necesita una recomendación: aumentar 10% el vapor en





V-12, reducir V-08 por canalización, monitorear P-23 por riesgo mecánico, ejecutar trazador en el patrón norte. REROSIM actúa como interfaz interpretativa entre alta resolución espacial y decisión por pozo.

REROSIM térmico: variables y ecuaciones de trabajo

Una formulación básica de REROSIM para vapor puede extender la relación de respuesta inter-pozo clásica. En lugar de usar únicamente tasas de inyección, se introduce el flujo de calor útil. Para un productor j , la respuesta de aceite, agua o temperatura puede expresarse como la suma de contribuciones de inyectores i ponderadas por una fracción de conectividad f_{ij} y un retardo τ_{ij} . En términos prácticos, la ecuación no pretende resolver toda la termodinámica, sino construir un modelo equivalente calibrado con datos históricos (Yan et al., 2025).

La forma conceptual sería:

$$R_j(t) = \sum f_{ij} \cdot U_i(t - \tau_{ij}) + D_j(t) + \epsilon$$

Donde:

R_j es una respuesta del productor,

U_i representa señal de vapor o energía inyectada,

D_j captura tendencia local de declinación o agotamiento y ϵ recoge ruido operacional.

Para inyección de vapor, U_i puede ser una combinación de tasa de vapor, calidad, presión, entalpía, volumen acumulado y duración del ciclo. La calibración permite estimar cuánto de esa energía afecta a cada productor.

El modelo puede incluir indicadores de eficiencia térmica por conexión: producción incremental de aceite por unidad de calor, incremento de temperatura por unidad de vapor, cambio en corte de agua y tiempo de residencia. Una conexión con alta temperatura y bajo aceite incremental no debe interpretarse como éxito; puede indicar que el vapor llegó al productor sin movilizar suficiente petróleo. En cambio, una respuesta diferida, con aumento gradual de aceite y control de agua, puede ser técnicamente más valiosa, aunque no sea la señal más intensa (Yan et al., 2025).

La calibración debe ser automática pero supervisada. REROSIM puede apoyarse en algoritmos no gradientes, filtros de ensamble o búsqueda bayesiana para ajustar conectividades, siempre con restricciones físicas: fracciones no negativas, suma acotada por inyector, retardos





consistentes con distancias y continuidad geológica, y penalización de conexiones incompatibles con barreras conocidas. Este balance entre dato y conocimiento de ingeniería evita mapas de conectividad matemáticamente correctos, pero geológicamente absurdos (Yan et al., 2025).

1.3. Scaled Load Ratio como capa de integridad mecánica

El Scaled Load Ratio (SLR) es una métrica diseñada para transformar cargas superficiales de bombas de varillas en una señal normalizada de condición. Su idea central es comparar la carga mínima y máxima observada con las cargas mínima y máxima esperadas en condición normal. Cuando el cociente escalado se mantiene cerca de la unidad, la bomba se encuentra en un régimen comparable al patrón sano. Cuando se aleja de la unidad, el sistema puede estar enfrentando llenado incompleto, fricción anormal, golpe de fluido, interferencia de gas, contacto de émbolo o cambios de carga que anticipan falla (Jung et al., 2026).

La importancia estratégica del SLR en un proyecto de vapor es que la mejora del yacimiento puede crear estrés adicional en el pozo. La inyección de vapor puede alterar temperatura de fluidos, viscosidad, producción de agua, gas liberado, emulsiones, sólidos, ciclos de arranque-parada y condiciones de bombeo. Una estrategia que aumenta producción, pero multiplica fallas de varillas, bombas o equipos de superficie no maximiza valor. Por tanto, el gemelo digital debe evaluar simultáneamente respuesta de yacimiento e integridad del sistema de levantamiento (Kokal y Sayegh, 1995).

La analítica SLR tiene una virtud operacional: usa datos de superficie y evita depender, para el primer nivel de alarma, del cálculo de cartas de fondo con especificaciones detalladas del tren de varillas. Esto no significa abandonar el diagnóstico tradicional; significa priorizar. El modelo puede generar una lista de pozos donde el comportamiento de carga se aleja del patrón sano con suficiente frecuencia, y los ingenieros pueden concentrar allí análisis más detallados, inspección de cartas, reducción de velocidad, cambio de carrera, tratamiento de gas, limpieza o intervención mecánica (Jung et al., 2026).

En la convergencia propuesta, SLR no es un módulo aislado de mantenimiento. Es una restricción dinámica para la optimización de vapor. Si un productor asociado a una conexión interpozo importante presenta SLR inestable, el algoritmo no debería recomendar aumentos agresivos de vapor sin evaluar la capacidad mecánica de recibir el nuevo régimen. De igual modo, si un inyector muestra alta oportunidad de barrido, pero los productores receptores tienen riesgo de





falla, la decisión óptima puede ser gradual, acompañada de ajustes de levantamiento artificial y vigilancia estrecha (Jung et al., 2026).

De la alarma al índice de integridad

El SLR puede convertirse en un índice de integridad mecánica cuando se combina con frecuencia de eventos, persistencia temporal, severidad de desviación, cambios de velocidad de bombeo y condiciones térmicas del pozo. Una desviación aislada puede ser ruido; una desviación sostenida después de un aumento de vapor puede señalar una relación causal. Por ello, se recomienda calcular SLR suavizado, frecuencia de anomalía en una ventana móvil, tendencia de variación y distancia al rango normal (Jung et al., 2026).

El índice de riesgo mecánico podría componerse de cuatro bloques: desviación SLR, variabilidad de carga, eventos operacionales y exposición térmica. La desviación SLR captura el alejamiento de la condición normal. La variabilidad de carga mide oscilaciones y posibles eventos intermitentes. Los eventos operacionales incorporan cambios de velocidad, paradas, arranques, tratamientos o modificaciones de inyección. La exposición térmica reconoce que la temperatura puede modificar viscosidad, expansión de materiales, elastómeros, corrosión y régimen de bombeo (Takacs, 2015).

La clasificación resultante permite ordenar los pozos en categorías de acción: *Verde*: condición estable, apta para optimización normal. *Amarilla*: vigilancia, ajustes conservadores y revisión de tendencia. *Naranja*: limitación temporal de cambios operacionales, diagnóstico de cartas y reducción de velocidad si aplica. *Roja*: alto riesgo de falla, intervención prioritaria o restricción automática en el optimizador. La utilidad no es sólo detectar fallas, sino impedir que el plan de yacimiento las acelere.

El aprendizaje automático puede mejorar la predicción de falla si se entrenan modelos con históricos de eventos. Sin embargo, en campos donde no existe una base etiquetada robusta, el SLR sigue siendo útil como métrica de arranque. Se recomienda iniciar con reglas transparentes y luego evolucionar hacia modelos supervisados o autoencoders cuando se acumulen datos suficientes de operación normal y fallas documentadas (Jung et al., 2026).

1.4. Arquitectura integrada del gemelo digital térmico

La arquitectura propuesta tiene cuatro capas: datos, modelos, decisión y aprendizaje: *La capa de datos* recibe información de producción, inyección, presión, temperatura, calidad de





vapor, cartas dinamométricas, eventos de mantenimiento, pruebas de pozo, propiedades de roca y fluidos, geología y límites operativos. *La capa de modelos* contiene el GNN térmico, REROSIM inter-pozo, SLR mecánico y reglas de negocio. *La capa de decisión* ejecuta escenarios, optimiza distribución de vapor, prioriza intervenciones y emite recomendaciones. *La capa de aprendizaje* compara pronóstico versus realidad y actualiza parámetros.

El flujo operativo inicia con depuración de datos. En yacimientos maduros, la calidad de la información es tan importante como el algoritmo. Un registro de vapor mal calibrado, una prueba de producción desfasada o una carta de bomba asociada a un pozo equivocado pueden inducir recomendaciones erróneas. Por ello, el gemelo debe contener reglas de validación: rangos físicos, consistencia de unidades, detección de outliers, reconciliación de eventos y trazabilidad de cambios.

La segunda etapa es la simulación base. Se generan escenarios con un simulador térmico de alta fidelidad para cubrir rangos de incertidumbre: permeabilidad, continuidad de facies, calidad de vapor, pérdidas de calor, presión, viscosidad, espaciamiento de pozos y estrategias de control. Estos escenarios entrenan el GNN. La calidad del sustituto depende de la amplitud y realismo del conjunto de entrenamiento; no se debe entrenar con casos que jamás podrían ocurrir ni excluir eventos críticos que sí aparecen en campo (Tariq y Shaikh, 2025).

La tercera etapa es la asimilación inter-pozo. REROSIM se calibra con datos reales de tasas, temperaturas, presiones y respuestas de productores. Así, la conectividad operativa puede diferir de la conectividad geológica esperada. Esta diferencia no se trata como error, sino como señal de aprendizaje: puede revelar una fractura, un canal de alta permeabilidad, una barrera no mapeada, un daño localizado o una comunicación vertical no anticipada (Yan et al., 2025).

La cuarta etapa es la vigilancia mecánica. El módulo SLR se actualiza con cargas superficiales y eventos de bombeo, calculando estado normal, desviaciones y alarmas. Cuando el optimizador prueba escenarios de redistribución de vapor, consulta este módulo para evitar recomendar cambios que superen la tolerancia mecánica de productores sensibles. De esta forma, el gemelo digital deja de ser un optimizador de barrido aislado y se convierte en un sistema de gestión de confiabilidad integral (Jung et al., 2026).





2. Metodología propuesta para implantación en campo

La implantación debe realizarse por fases, con entregables verificables y criterios de avance:

Fase cero, es una evaluación de madurez digital: disponibilidad de datos, confiabilidad de sensores, historial de fallas, estructura de bases de datos, estándares de nomenclatura, capacidad de cómputo y patrocinio operacional. Sin esta revisión, el proyecto corre el riesgo de construir un modelo sofisticado sobre datos débiles.

Fase uno, consiste en seleccionar un piloto de tamaño manejable: un patrón térmico, un bloque con historia representativa o un conjunto de inyectores y productores con alta prioridad de valor. El piloto debe tener suficiente variabilidad operacional para aprender, pero no tanta complejidad que impida cerrar el ciclo. Se recomienda incluir pozos con buena instrumentación, productores con levantamiento por varillas y al menos un historial de ajustes de vapor.

Fase dos, construye el modelo físico base. El equipo de yacimientos define casos de simulación térmica que cubren incertidumbres relevantes. El objetivo no es reproducir cada detalle del campo desde el inicio, sino crear una biblioteca de respuestas físicamente plausibles. Sobre esta biblioteca se entrena el GNN y se evalúan métricas de error en presión, temperatura, saturación y producción incremental. Los casos que generen mayor error se revisan para ampliar entrenamiento o ajustar arquitectura.

Fase tres, calibra REROSIM con datos de campo. Se calculan matrices de conectividad, retardos y eficiencia térmica. Los resultados se contrastan con conocimiento geológico, pruebas de trazadores, tendencias de temperatura, mapas de presión y experiencia de operación. Las conexiones fuertes y no explicadas se marcan como hipótesis para investigación; las conexiones esperadas pero ausentes se revisan por posible daño, barrera o falta de comunicación (Yan et al., 2025).

Fase cuatro, implementa SLR. Se define el periodo normal por pozo, se limpian outliers, se calculan cargas normalizadas y se entrenan reglas o modelos de alarma. El equipo de producción valida si las alarmas corresponden a eventos reales o a cambios operacionales benignos. Esta revisión es clave para evitar fatiga de alarmas. Una alarma útil debe ser suficientemente temprana para actuar y suficientemente precisa para merecer atención (Jung et al., 2026).





Fase cinco, integra los módulos en un tablero de decisiones. El usuario debe poder ver el mapa térmico estimado por GNN, matriz de conectividad REROSIM, estado SLR por productor, escenarios comparativos y recomendación final. La recomendación debe incluir explicación: ¿qué cambió?, ¿qué pozos se benefician?, ¿qué restricciones aplican?, ¿cuál es el riesgo? y ¿qué datos se usarán para verificar el resultado? La adopción depende de que los ingenieros puedan auditar el razonamiento del gemelo (Tariq y Shaikh, 2025; Yan et al., 2025; Jung et al., 2026).

3. Resultados

El éxito de la integración tecnológica en la industria petrolera depende de la sinergia entre la precisión de los datos y su ejecución en campo. En esta sección, se presentan y analizan los resultados clave obtenidos en la investigación, estructurados en cinco ejes fundamentales. En primer lugar, se evalúa la **gobernanza de datos** como pilar de la calidad de la información, seguido del **valor operativo** que aporta la convergencia tecnológica. Asimismo, se examinan los **escenarios de aplicación** específicos en yacimientos de crudo pesado y se realiza la **validación de las técnicas** mediante métricas de desempeño. Finalmente, se identifican los **riesgos técnicos** críticos y se proponen sus respectivas estrategias de mitigación para garantizar la viabilidad del proyecto.

3.1. Gobernanza de datos y calidad de la información

La gobernanza de datos no es un asunto administrativo; es una condición técnica. Los modelos integrados requieren que la tasa de vapor, calidad, presión de inyección, temperatura, producción de aceite y agua, velocidad de bombeo, cartas dinamométricas y los eventos de mantenimiento estén sincronizados en el tiempo. Una recomendación de redistribución de vapor puede fracasar si el modelo atribuye a conectividad lo que en realidad fue un cambio de bomba o una reparación de superficie.

Se recomienda diseñar un modelo de datos común por pozo y por evento. Cada registro debe tener identificador único, fecha y hora, fuente, unidad, calidad, estado de validación y relación con eventos. Los eventos son tan importantes como los valores continuos: cambios de velocidad, intervención con guaya, limpieza, falla eléctrica, cierre por mantenimiento, modificación de estrangulador, prueba de presión, ajuste de vapor, tratamiento químico o cambio de equipo. Sin eventos, el modelo puede interpretar discontinuidades como señales de yacimiento.





La calidad debe medirse mediante indicadores: porcentaje de datos faltantes, número de outliers removidos, retraso de actualización, discrepancia entre mediciones redundantes, trazabilidad de correcciones y estabilidad de sensores. Estos indicadores deben mostrarse junto con las predicciones. Si la confianza del dato baja, la confianza de la recomendación también debe bajar. Un gemelo digital responsable no oculta incertidumbre; la cuantifica y la comunica.

El aprendizaje continuo requiere versionamiento. Cada cambio de entrenamiento, calibración REROSIM, umbral SLR o regla de negocio debe quedar registrado. Cuando una recomendación produce buen resultado, se documenta la condición que la hizo exitosa. Cuando falla, se analiza si el problema fue dato, modelo, hipótesis física o ejecución. Esta disciplina convierte el piloto en una memoria técnica acumulativa y no en una aplicación aislada.

Valor operativo de la convergencia

El primer valor operativo es la velocidad. Un GNN entrenado puede evaluar escenarios de vapor en tiempos incompatiblemente menores que una simulación térmica completa. Esto permite explorar combinaciones de tasas, restricciones de presión, cierres temporales y redistribuciones sin esperar ciclos largos. La velocidad no debe confundirse con licencia para decidir sin revisión; debe usarse para ampliar el espacio de alternativas y seleccionar las mejores para análisis de ingeniería.

El segundo valor es la interpretabilidad inter-pozo. REROSIM ayuda a explicar por qué un escenario es recomendado. Si el GNN sugiere que aumentar vapor en un inyector mejora temperatura en una zona, REROSIM puede mostrar qué productores recibirían la energía y con qué retardo. Si la conexión se asocia a canalización, la estrategia puede cambiar de aumento de vapor a control de perfil, reducción de tasa o redistribución hacia otro patrón.

El tercer valor es la confiabilidad. SLR introduce una métrica que obliga a mirar el sistema de levantamiento artificial como parte del yacimiento extendido. En campos térmicos, el mejor barrido no genera valor si se pierde producción por fallas recurrentes. El módulo mecánico permite priorizar mantenimiento preventivo, ajustar velocidad, modificar bombas o planificar intervenciones antes de que la falla cause pérdida de producción (Jung et al., 2026; Takacs, 2015).

El cuarto valor es la integración multidisciplinaria. Geología aporta restricciones y barreras; yacimientos define física y escenarios; producción valida levantamiento y condiciones de pozo; mantenimiento interpreta fallas; operaciones ejecuta cambios; economía evalúa valor. El gemelo





digital se convierte en un espacio común donde cada disciplina ve cómo su información modifica la decisión final.

3.2. Escenarios de aplicación en campos de crudo pesado

En un patrón de inyección continua de vapor, el sistema puede usarse para balancear energía. El GNN predice cómo se moverá el frente térmico ante cambios de tasa. REROSIM identifica si la energía se dirige a productores con petróleo remanente o se recicla hacia pozos con alto corte de agua. SLR evalúa si los productores receptores pueden operar con el cambio previsto. La recomendación final podría ser aumentar vapor en un inyector, reducir otro, limitar velocidad de bombeo en un productor vulnerable y programar medición de temperatura para validar.

En un proceso cíclico tipo huff and puff, el gemelo puede aprender la relación entre volumen de vapor, tiempo de remojo, producción posterior, enfriamiento y fatiga operativa. REROSIM puede adaptarse para conexiones entre ciclos y vecinos afectados; el GNN estima distribución de calor alrededor del pozo; SLR vigila el retorno a producción, etapa en la que las cargas pueden cambiar por variación de viscosidad, gas, emulsiones o llenado.

En campos con canalización térmica, la convergencia puede identificar rutas de baja eficiencia. Un incremento rápido de temperatura en un productor, acompañado de poca producción incremental y conectividad REROSIM fuerte, sugiere que el vapor viaja por una vía preferencial. El plan podría incluir reducción de tasa, gel térmico, cambio de patrón, cierre temporal o redistribución a zonas menos drenadas. SLR añade una dimensión adicional: un productor con vapor cercano puede sufrir condiciones de bombeo distintas y requerir adaptación mecánica (Yan et al., 2025; Jung et al., 2026).

En campos con alta incertidumbre geológica, el sistema permite organizar campañas de datos. Las conexiones REROSIM no explicadas por el modelo estático pueden priorizar trazadores, registros de temperatura, pruebas de presión o actualización de facies. Las zonas donde el GNN muestra alta incertidumbre pueden recibir nuevos escenarios de simulación o datos de campo. Así, el gemelo no sólo responde preguntas; también indica qué datos reducen más incertidumbre (Tariq y Shaikh, 2025; Yan et al., 2025).





3.3. Validación de técnicas y métricas de desempeño

La validación debe cubrir tres niveles:

El primero es la validez del GNN térmico frente al simulador de referencia. Se deben medir errores en presión, temperatura, saturación, producción acumulada, tiempo de llegada térmica y balance de energía. No basta con un error global bajo; el modelo debe ser preciso en zonas de decisión, especialmente cerca de pozos, frentes térmicos, límites de presión y áreas de petróleo remanente (Tariq y Shaikh, 2025).

El segundo nivel es la validez de REROSIM frente al comportamiento de campo. Las métricas incluyen ajuste de tasas, temperatura, corte de agua, producción incremental, retardo de respuesta, consistencia de conectividad con trazadores y estabilidad temporal de la matriz. Un buen REROSIM no tiene que replicar cada fluctuación diaria; debe capturar relaciones útiles para decidir. La interpretación debe revisarse con geología y operaciones para evitar falsas conexiones inducidas por eventos no registrados.

El tercer nivel es la validez del SLR como predictor de riesgo. Las métricas incluyen precisión, recall, F1, anticipación promedio de fallas, tasa de falsas alarmas, reducción de tiempo fuera de servicio y aceptación por el equipo de producción. En etapas tempranas, la métrica más importante puede ser utilidad práctica: cuántas alarmas permitieron una acción preventiva razonable. A medida que se acumulan datos, se puede entrenar un modelo más robusto y estimar beneficios económicos (Jung et al., 2026).

La validación integrada debe medir valor de decisión. Se comparan escenarios recomendados contra línea base: producción incremental, consumo de vapor por barril incremental, reducción de agua producida, estabilidad de presión, número de fallas mecánicas, tiempo no productivo, costo de mantenimiento y variabilidad de operación. La convergencia sólo se justifica si mejora indicadores de yacimiento y confiabilidad al mismo tiempo.

3.4. Riesgos técnicos y mitigación

El primer riesgo es la extrapolación. Un GNN puede ser muy preciso dentro del rango de entrenamiento y fallar fuera de él. Si el campo opera con tasas, presiones o calidades de vapor no incluidas en los escenarios físicos, la predicción pierde confiabilidad. La mitigación es definir límites de validez, detectar condiciones fuera de dominio y redirigir casos extremos al simulador convencional antes de recomendar cambios (Tariq y Shaikh, 2025).





El segundo riesgo es la no unicidad de conectividad. Distintas matrices REROSIM pueden explicar datos similares si la información es insuficiente. La mitigación combina regularización física, edades geológicas, trazadores, pruebas de presión y revisión multidisciplinaria. La conectividad no debe interpretarse como verdad absoluta, sino como hipótesis dinámica con incertidumbre (Yan et al., 2025).

El tercer riesgo es la fatiga de alarmas SLR. Si los umbrales son demasiado sensibles, el personal deja de confiar en el sistema. Si son demasiado laxos, la alerta llega tarde. La mitigación es calibrar por pozo, usar ventanas móviles, distinguir alarmas informativas de críticas, registrar acciones tomadas y ajustar umbrales con retroalimentación operacional (Jung et al., 2026).

El cuarto riesgo es la fragmentación organizacional. Un gemelo digital integrado requiere acuerdos entre yacimientos, producción, mantenimiento y tecnología. Si cada área usa nomenclaturas, bases de datos y criterios distintos, la integración se vuelve frágil. La mitigación es nombrar responsables de datos, definir estándares de eventos y establecer reuniones periódicas donde las recomendaciones se revisen con evidencia.

3.5. Análisis y discusión de resultados

En esta sección se evalúan y discuten, de forma crítica, los hallazgos de la investigación, conectando los fundamentos teóricos con la viabilidad práctica en el entorno industrial. El análisis parte de las **discusiones estratégicas** que justifican el modelo, para luego profundizar en la **optimización operacional** a través de la definición de su función objetivo y restricciones. Con el fin de garantizar la aplicabilidad real, se aborda la **explicabilidad y aceptación** del sistema por parte de los equipos de campo, seguida de la propuesta estructural para el **diseño de un piloto**. Asimismo, se evalúan los **requerimientos computacionales y la escalabilidad** de la arquitectura tecnológica, complementados con un balance de los **indicadores económicos y de sostenibilidad operacional**. Finalmente, se examinan con transparencia las **limitaciones del enfoque** y se ofrece la **formulación mínima junto al pseudocódigo operativo** que sirve de guía para su implementación.

Discusiones estratégicas

La convergencia GNN-REROSIM-SLR representa un cambio de enfoque: de modelos que describen el pasado hacia sistemas que recomiendan acciones bajo incertidumbre. El valor técnico de esta propuesta radica en conectar inteligencia artificial con problemas concretos y





maximizar la recuperación de los campos maduros, reducir consumo de energía, evitar fallas mecánicas y transformar datos dispersos en decisiones auditables.

El enfoque también es coherente con una visión de integración energética. La eficiencia térmica no es sólo una métrica de yacimiento; tiene implicaciones de consumo de combustible, emisiones, costos y sostenibilidad. Un barril incremental obtenido con menor vapor por barril, menor tiempo fuera de servicio y menos intervenciones correctivas tiene una huella operacional más favorable. En ese sentido, el gemelo digital térmico puede ser una herramienta de productividad y de gestión responsable de energía.

Desde el punto de vista de investigación aplicada, el marco abre líneas de trabajo. La primera es construir datasets térmicos sintéticos para entrenamiento GNN en geometrías representativas de crudos pesados. La segunda es desarrollar REROSIM térmico con variables de entalpía y temperatura. La tercera es adaptar SLR a pozos térmicos con levantamiento por varillas, considerando efectos de viscosidad, temperatura y ciclos. La cuarta es diseñar tableros con explicabilidad, donde cada recomendación incluya por qué técnico (Yan et al., 2025).

La propuesta no busca presentar la inteligencia artificial como sustituto del criterio experto. Al contrario, ubica la experiencia de ingeniería como condición de diseño, calibración, validación y adopción. La IA acelera, ordena y detecta patrones; la ingeniería define límites, interpreta causas y decide acciones de campo. El resultado más deseable es una alianza: algoritmos rápidos con juicio técnico maduro.

Optimización operacional: función objetivo y restricciones

La convergencia propuesta permite formular una función objetivo más realista que la maximización simple de producción. En un campo térmico, el objetivo debe considerar petróleo incremental, consumo de vapor, costo energético, manejo de agua, estabilidad de presión, riesgo mecánico y capacidad de facilidades. Una formulación práctica puede maximizar valor neto esperado/periodo, restando costo de generación y distribución de vapor, costo de mantenimiento asociado al riesgo SLR y penalización por operar cerca de límites de presión o temperatura.

El GNN aporta velocidad para evaluar miles de combinaciones de tasas de vapor. REROSIM reduce el problema al mostrar qué conexiones son más sensibles y cuáles responden con retardo. SLR agrega una penalización por productor cuando el escenario exige condiciones que pueden deteriorar la bomba. El optimizador no busca el mayor calentamiento, sino el mayor





valor sostenible: calor aplicado donde produce petróleo, con equipos capaces de sostener el nuevo régimen y con incertidumbre dentro de límites aceptables.

Las restricciones deben definirse con participación de operaciones. Entre ellas se incluyen presión máxima de inyección, capacidad de generación de vapor, límites de calidad, disponibilidad de líneas, restricciones ambientales, límites de producción de agua, presión de fondo de productores, capacidad de separación y categoría de integridad mecánica. Un pozo en categoría roja por SLR puede permanecer en el modelo, pero con tasa de cambio limitada o con bloqueo temporal para recomendaciones agresivas.

La optimización debe operar en dos horizontes. El horizonte táctico, diario o semanal, ajusta tasas dentro de rangos seguros. El horizonte estratégico, mensual o trimestral, evalúa cambios de patrón, nuevas conversiones, trabajos de control de perfil, perforación de infill o rediseño del esquema térmico. En ambos casos, el valor del sistema es la trazabilidad: cada recomendación debe mostrar qué predijo el GNN, qué conexión justificó REROSIM, qué restricción impuso SLR y qué indicador económico motivó la decisión.

Explicabilidad y aceptación por los equipos de campo

La aceptación de un gemelo digital no depende sólo de la exactitud estadística; depende de que los usuarios reconozcan una lógica de ingeniería. Un tablero que muestre únicamente una tasa recomendada sin explicación genera resistencia. En cambio, una recomendación acompañada por conectividad, sensibilidad térmica, estado mecánico y rango de incertidumbre facilita la discusión técnica. La explicabilidad debe diseñarse desde el inicio y no agregarse al final como elemento decorativo.

Para el GNN, la explicabilidad puede apoyarse en mapas de sensibilidad, importancia de atributos y comparación contra escenarios físicos. El usuario debe poder identificar si la predicción se debe a permeabilidad, temperatura inicial, presión, cambio de vapor o cercanía a una conexión dominante. Para REROSIM, la explicación es más directa: fracciones de contribución, retardos, conexiones fuertes y eficiencia por unidad de calor. Para SLR, la señal debe mostrar evolución temporal, rango normal y eventos que explican la desviación.

Una práctica recomendada es usar fichas de decisión por recomendación. Cada ficha incluye descripción del cambio propuesto, beneficio esperado, pozos impactados, evidencia GNN, evidencia REROSIM, estado SLR, restricciones, incertidumbre, plan de seguimiento y fecha de revisión. Estas fichas convierten la IA en un proceso auditable y sirven para el





aprendizaje organizacional. Si la recomendación se ejecuta, se compara el resultado contra pronóstico; si no se ejecuta, se documenta la causa.

La adopción también mejora cuando el sistema reconoce conocimiento experto. Los ingenieros deben poder bloquear una conexión por evidencia geológica, ajustar un rango de presión por experiencia de campo o marcar un evento no capturado automáticamente. El gemelo digital debe ser interactivo: aprende de datos, pero también de observaciones técnicas verificables. Esta interacción evita el falso dilema entre modelo y experiencia.

Diseño de un piloto

Un piloto venezolano debería enfocarse en un área donde exista historia de inyección de vapor, mediciones de producción confiables y equipos de levantamiento con datos de cargas superficiales. La selección del patrón no debe basarse únicamente en producción alta; conviene escoger un bloque donde existan problemas visibles de eficiencia térmica, variabilidad inter-pozo y fallas mecánicas recurrentes. Allí la convergencia puede mostrar beneficios en tres dimensiones: más producción, menos vapor desperdiciado y menos paradas.

El primer entregable del piloto sería una línea base integrada. Esta línea base debe documentar producción por pozo, vapor acumulado, relación vapor-petróleo, temperatura de producción, corte de agua, eventos de falla, frecuencia de intervención, disponibilidad mecánica y costos asociados. Sin línea base, no es posible demostrar valor. La línea base también debe separar producción incremental atribuible a vapor de fluctuaciones por mantenimiento, cambios de bomba o condiciones de superficie.

El segundo entregable sería un mapa de conectividad térmica. Este mapa puede revelar patrones intuitivos y no intuitivos. Las conexiones intuitivas validan el modelo y generan confianza; las no intuitivas abren preguntas técnicas. Por ejemplo, una conexión fuerte entre inyector y productor alejados puede sugerir canal de alta permeabilidad, comunicación vertical o evento operacional mal registrado. La decisión no es aceptar ciegamente la matriz, sino convertirla en hipótesis de trabajo para confirmar con datos.

El tercer entregable sería un ranking de pozos por oportunidad y riesgo. La oportunidad se calcula con potencial de respuesta térmica y eficiencia de vapor; el riesgo se calcula con SLR, historial de fallas, exposición térmica y criticidad operacional. Este ranking ayuda a decidir dónde ejecutar ajustes de vapor, dónde instalar mejores sensores, dónde hacer mantenimiento preventivo y dónde no conviene forzar producción hasta resolver integridad.





Requerimientos computacionales y escalabilidad

El entrenamiento de GNN sobre datos espacio - temporales de yacimiento puede ser intensivo en memoria y tiempo, especialmente, si se usan mallas de alta resolución y horizontes largos. La literatura base muestra que el entrenamiento multi-GPU, el paralelismo de datos, el almacenamiento local de caché y el procesamiento por lotes conscientes del dominio reducen cuellos de botella. Para un piloto, no siempre se requiere un clúster grande, pero sí una arquitectura escalable que pueda crecer desde un bloque hasta un activo completo (Tariq y Shaikh, 2025).

La estrategia recomendada es iniciar con subgrafos. En lugar de entrenar sobre todo el campo, se agrupan patrones o regiones con continuidad operativa. Esto reduce tamaño, acelera iteración y permite detectar errores de datos más rápido. Cuando el piloto madura, se conectan subgrafos mediante aristas de borde que representan comunicación entre patrones. Este enfoque modular también facilita que diferentes equipos trabajen por áreas y luego integren resultados (Tariq y Shaikh, 2025).

El sistema debe separar entrenamiento e inferencia. El entrenamiento puede ejecutarse en ventanas programadas, usando simulaciones físicas y datos históricos revisados. La inferencia debe ser suficientemente rápida para correr escenarios operativos y actualizar tableros. REROSIM y SLR pueden actualizarse con mayor frecuencia que el GNN si llegan nuevos datos de producción y cargas. Esta asincronía es útil: no todos los módulos necesitan reentrenarse al mismo ritmo (Tariq y Shaikh, 2025; Yan et al., 2025; Jung et al., 2026).

También es importante definir almacenamiento reproducible. Los datasets de entrenamiento, modelos, parámetros REROSIM, umbrales SLR y resultados de escenarios deben versionar. Si una recomendación tomada en marzo debe revisarse en junio, el equipo debe poder reconstruir con qué modelo, qué datos y qué supuestos se generó. La trazabilidad no es burocracia; es gestión de riesgo técnico.

Indicadores económicos y de sostenibilidad operacional

El indicador clásico de proyectos térmicos es la relación vapor-petróleo, pero la convergencia permite ampliar el tablero. Se recomienda medir barriles incrementales / tonelada equivalente de vapor, energía consumida/barril incremental, reducción de vapor reciclado, producción diferida evitada por mantenimiento preventivo, reducción de fallas/ mil días de operación, costo de intervención/barril incremental y disponibilidad de equipos críticos.





La sostenibilidad operacional se refleja en usar menos energía para obtener el mismo o mayor volumen de petróleo. Un patrón donde el vapor circula por canales preferenciales consume combustible, agua tratada y capacidad de generación sin generar barrido proporcional. REROSIM ayuda a detectar esa ineficiencia; el GNN estima alternativas de redistribución; SLR evita que el ajuste se transforme en fallas mecánicas. El resultado puede ser una reducción simultánea de costo energético y tiempo no productivo.

La evaluación económica debe considerar incertidumbre. No se recomienda reportar una única cifra de producción incremental, sino rangos asociados a escenarios de conectividad, desempeño mecánico y calidad de datos. Un caso base, un caso conservador y un caso optimista permiten comunicar mejor la decisión. Cuando la incertidumbre sea alta, la recomendación puede ser adquirir datos adicionales antes de ejecutar cambios irreversibles.

En términos de gobernanza corporativa, el sistema puede contribuir a justificar inversiones en sensores, cómputo y mantenimiento predictivo. Si el piloto demuestra que mejores datos reducen vapor desperdiciado o evitan fallas, la infraestructura digital deja de verse como gasto tecnológico y se convierte en habilitador de producción rentable.

Limitaciones del enfoque propuesto

La primera limitación es que el artículo propone un marco conceptual aplicado, no una validación con datos de campo propios. Por lo tanto, las ecuaciones, flujos de trabajo y métricas deben interpretarse como diseño metodológico. La implantación real requiere simulaciones térmicas, datos históricos, sensores, validación con expertos y análisis económico específico del activo.

La segunda limitación es que la transferencia desde CO₂ a vapor exige modificaciones sustantivas. Un GNN entrenado para CO₂ no puede usarse directamente para vapor. Deben cambiarse variables, pérdidas, escenarios, datos y controles. La contribución es la estructura de aprendizaje en grafos, no los pesos de un modelo entrenado en otra física (Tariq y Shaikh, 2025).

La tercera limitación se relaciona con REROSIM. Al ser una simplificación inter-pozo, puede perder detalle local y confundir efectos simultáneos si los datos de eventos son pobres. Su fortaleza es la velocidad y la interpretación, no la sustitución de modelos térmicos detallados en decisiones de diseño mayor (Yan et al., 2025).

La cuarta limitación corresponde a SLR. La métrica se desarrolló para bombas de varillas y debe adaptarse con cuidado a condiciones térmicas, crudos viscosos, cambios de régimen y





particularidades de completación. Para equipos de levantamiento distintos, se requerirían métricas equivalentes de integridad mecánica (Jung et al., 2026; Takacs, 2015).

Formulación mínima y pseudocódigo operativo

El flujo operativo puede resumirse en una secuencia reproducible. Primero, se actualizan datos de campo y se valida la calidad. Segundo, se calcula el estado mecánico por pozo mediante SLR y se asigna categoría de integridad. Tercero, se actualiza REROSIM con tasas, temperaturas y eventos recientes para recalcular conectividad y retardos. Cuarto, se ejecutan escenarios de vapor en el GNN térmico dentro de límites de validez. Quinto, se comparan escenarios con una función objetivo que incorpora producción incremental, eficiencia térmica, restricciones de presión y penalización mecánica.

En pseudocódigo de ingeniería: para cada patrón, leer datos validados; si calidad de datos es insuficiente, emitir recomendación de adquisición antes de optimizar; calcular SLR por productor; actualizar conectividad REROSIM; generar conjunto de escenarios de vapor; estimar respuesta térmica con GNN; descartar escenarios que violen límites de presión, vapor o integridad; ordenar escenarios por valor esperado; entregar recomendación con evidencia y plan de verificación.

La fórmula conceptual de SLR usada como señal primaria puede escribirse como $SLR = (L_{min}/L_{min_normal}) / (L_{max}/L_{max_normal})$. Para el índice de integridad se puede usar una combinación ponderada: $IM = w_1 \cdot |SLR - 1| + w_2 \cdot AFR + w_3 \cdot \sigma_{carga} + w_4 \cdot E_{eventos} + w_5 \cdot E_{termica}$, donde AFR es la frecuencia de anomalía en ventana móvil, σ_{carga} representa variabilidad y $E_{eventos}/E_{termica}$ incorporan contexto operacional. Los pesos deben calibrarse por campo y por tipo de equipo (Jung et al., 2026).

La respuesta inter-pozo térmica puede expresarse como $R_j(t) = \sum f_{ij} \cdot U_i(t - \tau_{ij}) + D_j(t)$, donde U_i puede ser tasa de vapor, entalpía o calor útil acumulado. Esta ecuación no reemplaza el simulador térmico; es una representación reducida para interpretar conectividad y asignar responsabilidades de respuesta entre inyectores y productores. Cuando REROSIM y GNN discrepan, la divergencia debe tratarse como oportunidad de diagnóstico: puede indicar datos faltantes, evento no registrado, conectividad no modelada o escenario fuera del dominio de entrenamiento (Yan et al., 2025).

La disciplina de cierre del ciclo es esencial. Cada recomendación debe tener una hipótesis medible: incremento esperado, tiempo de respuesta, pozos impactados y señal mecánica



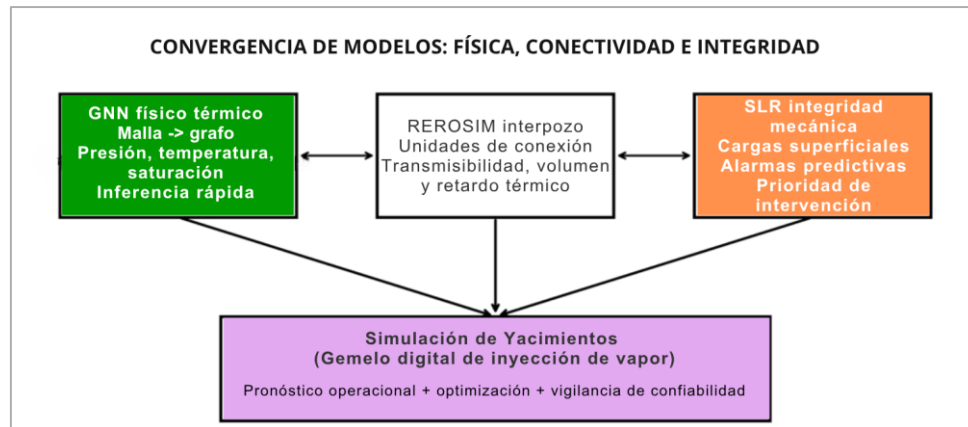


esperada. Después de la ejecución, el equipo compara el resultado real, actualiza los módulos y decide si la política queda aceptada, ajustada o descartada. Así el gemelo digital evoluciona como sistema de aprendizaje operacional y no como reporte estático.

Conclusiones

La integración de GNN, REROSIM y SLR permite diseñar un gemelo digital de inyección de vapor con tres capacidades complementarias: modelado físico rápido, interpretación inter-pozo y vigilancia de integridad mecánica. Esta combinación responde mejor a la realidad de campos maduros que un modelo único enfocado sólo en producción (Figura 1).

Figura 1: Arquitectura de convergencia GNN–REROSIM–SLR para un gemelo digital térmico



Fuente: elaboración propia con base conceptual en Tariq y Shaikh (2025), Yan et al. (2025) y Jung et al. (2026).

La transferencia de GNN desde aplicaciones de CO₂ hacia vapor es técnicamente viable como línea de investigación aplicada, siempre que se incorporen variables térmicas, balance de energía, calidad de vapor, condensación y dependencia de viscosidad con temperatura. El sustituto no reemplaza el simulador físico; lo convierte en motor de entrenamiento y validación para decisiones rápidas (Tariq y Shaikh, 2025).

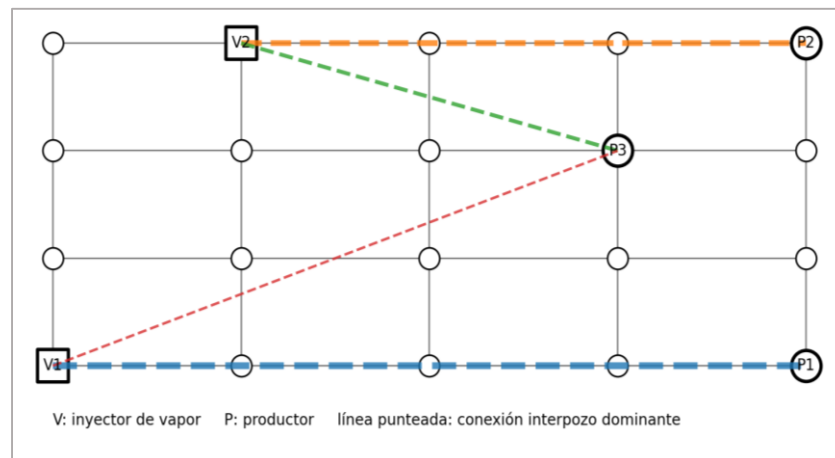
REROSIM aporta una capa de conectividad dinámica que traduce mapas y predicciones a relaciones inyector-productor accionables. Su extensión térmica puede identificar reciclaje de calor, canales preferenciales, retardos útiles y oportunidades de redistribución de vapor (Yan et al., 2025).



SLR convierte la integridad mecánica en restricción de optimización y no en análisis posterior a la falla. La producción incremental debe evaluarse junto con el riesgo de bombas, varillas y continuidad operacional (Jung et al., 2026).

La metodología propuesta debe implementarse por fases: calidad de datos, piloto, simulación base, entrenamiento GNN, calibración REROSIM, despliegue SLR, integración en tablero, validación y aprendizaje continuo. La gobernanza de eventos y la revisión multidisciplinaria son condiciones críticas de éxito. (Figura 2).

Figura 2. Esquema de grafo térmico: la malla de yacimiento se interpreta como nodos conectados y como vínculos inter-pozo dominantes.



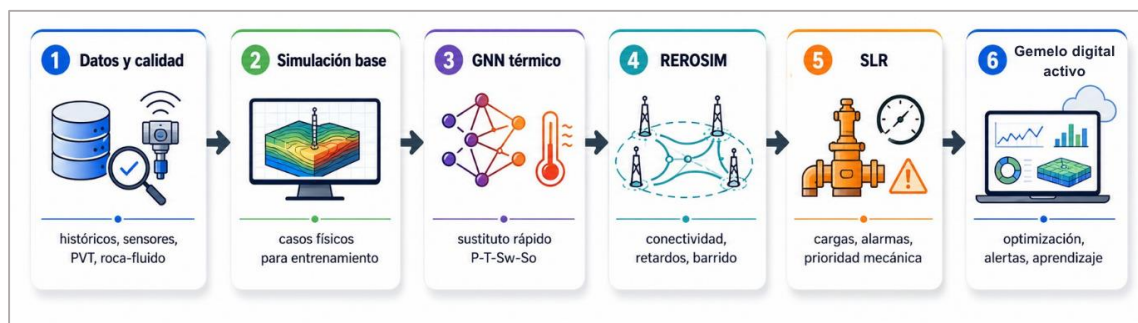
Fuente: elaboración propia.

El valor estratégico de la convergencia está en decidir mejor y más rápido: menos vapor desperdiciado, mayor barrido efectivo, menor tiempo no productivo, intervenciones priorizadas y una cultura técnica donde los datos se convierten en acciones trazables.

La figura 3 muestra la ruta de implantación por etapas para un piloto de gemelo digital de inyección de vapor.



Figura 3. Ruta de implantación por etapas para un piloto de gemelo digital de inyección de vapor.



Fuente: elaboración propia.

En las tablas del 1 al 3 se presentan:

- Esquemas mínimos de datos para el gemelo digital térmico (Fuente: elaboración propia)
- Métricas de validación recomendadas por capa del modelo
- Fases de ejecución sugeridas para un piloto PetroRenova de gemelo digital térmico (Fuente: elaboración propia)

Tabla 1. Esquema mínimo de datos para el gemelo digital térmico

Dominio	Variables mínimas	Uso en GNN	Uso en REROSIM	Uso en SLR / integridad
Yacimiento	Porosidad, permeabilidad, espesor, facies, profundidad, presión, saturaciones, viscosidad.	Atributos de nodos y condiciones iniciales.	Restricciones geológicas para conectividad.	Contexto de riesgo por fluido y régimen.
Vapor	Tasa, calidad, presión, temperatura, entalpía, tiempo de ciclo, pérdidas estimadas.	Control dinámico y balance de energía.	Señal de entrada por inyector y retardo térmico.	Evento operacional que puede modificar cargas.
Producción	Aceite, agua, gas, temperatura, presión de fondo/cabeza, corte de agua, pruebas.	Salidas y validación por pozo.	Respuesta por productor y eficiencia de conexión.	Condiciones que explican desviaciones de carga.
Levantamiento	Cargas mínimas/máxima, velocidad, carrera, torque, estado de bomba, eventos.	Variable auxiliar para escenarios operativos.	Restricción de producción por productor.	Cálculo SLR, alarmas, clasificación de riesgo.
Eventos	Cierres, intervenciones, cambios de bomba, tratamientos, ajustes de vapor, fallas.	Etiquetas y discontinuidades controladas.	Evita atribuir eventos a conectividad falsa.	Etiquetas para entrenamiento de falla.

Fuente: elaboración propia



Tabla 2. Métricas de validación recomendadas por capa del modelo

Capa	Métricas técnicas	Criterio de aceptación inicial	Decisión habilitada
GNN térmico	RMSE/MAE de presión, temperatura y saturación; error de llegada térmica; balance de energía.	Error estable en zonas de pozo y sin deriva relevante en horizonte de pronóstico.	Evaluar escenarios de vapor y sensibilidad operacional.
REROSIM	Ajuste de producción, corte de agua, temperatura, retardo, estabilidad de conectividad.	Conexiones coherentes con datos y geología; incertidumbre documentada.	Redistribuir vapor, detectar canalización, priorizar trazadores.
SLR	F1, recall, falsas alarmas, anticipación de falla, tiempo no productivo evitado.	Alarmas accionables y validadas por producción/mantenimiento.	Planificar mantenimiento, ajustar velocidad y limitar escenarios riesgosos.
Integración	NP incremental, vapor por barril, disponibilidad mecánica, costo por barril incremental.	Mejora frente a línea base y revisión multidisciplinaria.	Emitir recomendación operativa trazable.

Fuente: elaboración propia

Tabla 3. Fases de ejecución sugeridas para un piloto PetroRenova de gemelo digital térmico

Fase	Actividad	Entregable	Riesgo de no ejecutar
0	Evaluación de madurez digital y calidad de datos.	Inventario de datos, brechas, diccionario y plan de limpieza.	Entrenar modelos con información inconsistente.
1	Selección de patrón piloto y objetivos de valor.	Alcance, KPIs, línea base y criterios de éxito.	Piloto demasiado amplio o sin impacto medible.
2	Biblioteca de simulación térmica.	Casos físicos para entrenamiento y validación GNN.	Sustituto rápido sin fundamento físico suficiente.
3	Entrenamiento y prueba del GNN.	Modelo térmico con límites de validez y métricas.	Predicciones rápidas, pero no confiables fuera del rango.
4	Calibración REROSIM térmica.	Matriz de conectividad, retardos y eficiencia térmica.	Recomendaciones sin explicación interpozo.
5	Implementación SLR.	Umbrales por pozo, alarmas y categorías de integridad.	Optimizar vapor ignorando riesgo mecánico.
6	Tablero integrado y cierre de ciclo.	Recomendaciones, seguimiento y aprendizaje continuo.	No convertir análisis en acciones de campo.

Fuente: elaboración propia





Referencias

- Jung, Y., Kim, Y., Oh, B., Jun, J., Sun, W. y Jeong, H. (2026). Prediction of Sucker Rod Pump Failures Using Scaled Load Ratios and Machine Learning. *SPE Journal*. SPE-233386-PA.
- Tariq, Z. y Shaikh, M. (2025). Fast, Scalable, and Memory-Efficient High-Performance Graph Neural Networks for Subsurface in Porous Media Simulations using Multi-GPU Parallelism. SPE-229244-MS. *Society of Petroleum Engineers*. <https://doi.org/10.2118/229244-MS>
- Yan, L., Zhou, Z., Li, B., Fang, Q., Yuankai y Shi, Z. (2025). A Reservoir Dynamic Prediction Model Based on the REROSIM Method: Digital Twin and Simulation Studies Driven by Digital Transformation. SPE-227319-MS. *Society of Petroleum Engineers*. <https://doi.org/10.2118/227319-MS>
- Takacs, G. (2015). Sucker-Rod Pumping Handbook: Production Engineering Fundamentals and Long-Stroke Rod Pumping. *Elsevier*.
- Kokal, S. y Sayegh, S. (1995). Asphaltenes: The Cholesterol of Petroleum. *SPE* 29787.
- Shaikh, [Iniciales]. (2025). Automated well optimization. *Journal of Petroleum Science and Engineering* (ahora Geoenergy Science and Engineering) presentado en una conferencia de la SPE (Society of Petroleum Engineers).
- Tissot, B. P. y Welte, D. H. (1984). *Petroleum Formation and Occurrence*. Springer-Verlag.

Declaración de conflicto de interés y originalidad

Conforme a lo estipulado en el *Código de ética y buenas prácticas* publicado en **PetroRenova Indexed, Revista Científica de la Energía**, la autora **Romero, Isabelis**, declara al Comité Editorial que no tiene situaciones que representen conflicto de interés real, potencial o evidente, de carácter académico, financiero, intelectual o con derechos de propiedad intelectual relacionados con el contenido del artículo: **Convergencia estratégica de Redes Neuronales de Grafos, REROSIM y Scaled Load Ratio para gemelos digitales de inyección de vapor**, en relación con su publicación. De igual manera, declara que el trabajo es original, no ha sido publicado parcial ni totalmente en otro medio de difusión, no se utilizaron ideas, formulaciones, citas o ilustraciones diversas, extraídas de distintas fuentes, sin mencionar de forma clara y estricta su origen y sin ser referenciadas debidamente en la bibliografía correspondiente. Consiente que el Comité Editorial aplique cualquier sistema de detección de plagio para verificar su originalidad.

La autora declara que, en la preparación de este manuscrito, no utilizó herramientas de inteligencia artificial generativa para la redacción de textos o interpretación de datos.

Para citar este artículo:

Romero, I. (2026). Convergencia estratégica de Redes Neuronales de Grafos, REROSIM y Scaled Load Ratio para gemelos digitales de inyección de vapor. *PetroRenova Indexed, Revista Científica de la Energía*. Vol. 2, núm. 3(70-96), julio-septiembre, 2026. <https://doi.org/10.5281/zenodo.21810662>

