Research Club Scientia Vol. 4, No 4
Frecuencia: Octubre–Diciembre 2026
ISSN: 3103 - 1315
DOI: https://doi.org/10.65415/xs708305

Aprendizaje por refuerzo frente a métodos experimentales tradicionales en la determinación de ciclos de sinterización cerámicos: una revisión sistemática de literatura

Reinforcement learning versus traditional experimental methods for determining ceramic sintering cycles: a systematic literature review

Apprendimento per rinforzo rispetto ai metodi sperimentali tradizionali nella determinazione dei cicli di sinterizzazione ceramici: una revisione sistematica della letteratura

Jesús Andres Mendoza Fonseca[I]
jmendozaf3@unemi.edu.ec
ORCID: https://orcid.org/0009-0000-2965-1956
Facultad de Ciencias de la Ingeniería, Carrera Ingeniería Industrial, Universidad Estatal de Milagro, Ecuador

Christian Rolando Díaz Cruz[I]
cdiazc6@unemi.edu.ec
ORCID: https://orcid.org/0009-0008-3879-8683
Facultad de Ciencias de la Ingeniería, Carrera Ingeniería Industrial, Universidad Estatal de Milagro, Ecuador

Liggia Brigitte Cotto Peñafiel[I]
lcottop@unemi.edu.ec
ORCID: https://orcid.org/0009-0009-4537-2311
Facultad de Ciencias de la Ingeniería, Carrera Ingeniería Industrial, Universidad Estatal de Milagro, Ecuador

George Xavier Toral Espinoza[I]
gtorale@unemi.edu.ec
ORCID: https://orcid.org/0009-0001-5724-8758
Facultad de Ciencias de la Ingeniería, Carrera Ingeniería Industrial, Universidad Estatal de Milagro, Ecuador

Jorge Tomas Holguín Anzules[I]
jholguina7@unemi.edu.ec
ORCID: https://orcid.org/0000-0003-4454-4184
Facultad de Ciencias de la Ingeniería, Carrera Ingeniería Industrial, Universidad Estatal de Milagro, Ecuador

.

Aprendizaje por refuerzo frente a métodos experimentales tradicionales en la determinación de ciclos de sinterización cerámico una revisión sistemática de literatura

Correspondencia: jmendozaf3@unemi.edu.ec

Artículo de Investigación

Aceptado: 27/Noviembre/2026 **Publicado:** 28/Noviembre/2026

I. Facultad de Ciencias de la Ingeniería, Carrera Ingeniería Industrial, Universidad Estatal de Milagro, Ecuador

.

Resumen

La determinación de los ciclos de sinterización de cerámicos avanzados se basa en métodos experimentales de prueba y error, con ciclos estáticos que se corrigen manualmente y defectos que solo se detectan al finalizar el proceso. Esté presente artículo se desarrolla una revisión sistemática de literatura (RSL) de documentos publicados entre 2024 y 2026 orientada por la pregunta: ¿En qué medida el aprendizaje por refuerzo (RL) supera a los métodos experimentales tradicionales en la determinación de ciclos de sinterización cerámicos? Se analizaron 10 estudios, 6 de ellos revisados por pares, y se compararon ambos enfoques en seis criterios: tiempo y número de ensayos, consumo energético, calidad del sinterizado, adaptación en tiempo real, optimización de la temperatura y nivel de validación. Los resultados indican que la superioridad del RL es todavía parcial y principalmente potencial: un agente de Q-learning profundo acoplado a un gemelo digital redujo entre 24 y 27 % el consumo energético predicho de un horno industrial de inducción frente a su patrón de operación habitual, y una red Q profunda redujo en 111 °C la temperatura de sinterización en el diseño de composiciones. En cambio, la ventaja medible frente a la prueba y error procede de técnicas de apoyo: modelos sustitutos que predicen la densidad final con errores de 0.17 – 0.42 % y optimización bayesiana que redujo el esfuerzo de desarrollo en 72 %. Ningún estudio compara directamente el RL con el método tradicional en un horno de sinterización cerámica, por lo que esta última conserva la ventaja en validación industrial.

Palabras clave: Aprendizaje por refuerzo; Q-learning; sinterización; ciclos térmicos; cerámicos avanzados.

Abstract

Determining sintering cycles for advanced ceramics still relies on trial-and-error experimental methods, with static cycles that are corrected manually and defects that are detected only at the end of the process. This article presents a systematic literature review (SLR) of documents published between 2024 and 2026 guided by the question: to what extent does reinforcement learning (RL) outperform traditional experimental methods in determining ceramic sintering cycles? Ten studies were analyzed, six of them peer-reviewed, and both approaches were compared on six criteria: time and number of trials, energy consumption, sintered quality, real-time adaptation, temperature optimization and level of validation. The results indicate that the superiority of RL is still partial and mainly potential: a deep Q-learning agent coupled with a digital twin reduced the predicted energy consumption of an industrial induction furnace by 24–27 % compared with its usual operating pattern, and a deep Q-network lowered the sintering temperature by 111 °C in composition design. By contrast, the measurable advantage over trial and error currently comes from supporting techniques: surrogate models that predict final density with errors of 0.17–0.42 % and Bayesian optimization that reduced development effort by 72 %. No study

directly compares RL with the traditional method in a ceramic sintering furnace, so the latter retains the advantage in industrial validation.

Keywords: Reinforcement learning; Q-learning; sintering; thermal cycles; advanced ceramics.

Riassunto

La determinazione dei cicli di sinterizzazione dei ceramici avanzati si basa ancora su metodi sperimentali per tentativi ed errori, con cicli statici corretti manualmente e difetti rilevati solo al termine del processo. Il presente articolo sviluppa una revisione sistematica della letteratura (RSL) di documenti pubblicati tra il 2024 e il 2026 guidata dalla domanda: in che misura l'apprendimento per rinforzo (RL) supera i metodi sperimentali tradizionali nella determinazione dei cicli di sinterizzazione ceramici? Sono stati analizzati 10 studi, 6 dei quali con revisione paritaria, e i due approcci sono stati confrontati su sei criteri: tempo e numero di prove, consumo energetico, qualità del sinterizzato, adattamento in tempo reale, ottimizzazione della temperatura e livello di convalida. I risultati indicano che la superiorità dell'IRL è ancora parziale e principalmente potenziale: un agente di Q-learning profondo accoppiato a un gemello digitale ha ridotto del 24–27 % il consumo energetico previsto di un forno industriale a induzione rispetto al suo schema operativo abituale, e una rete Q profonda ha abbassato di 111 °C la temperatura di sinterizzazione nella progettazione delle composizioni. Al contrario, il vantaggio misurabile rispetto al metodo per tentativi ed errori proviene oggi da tecniche di supporto: modelli surrogati che prevedono la densità finale con errori dello 0,17–0,42 % e ottimizzazione bayesiana che ha ridotto lo sforzo di sviluppo del 72 %. Nessuno studio confronta direttamente l'RL con il metodo tradizionale in un forno di sinterizzazione ceramica, per cui quest'ultimo mantiene il vantaggio nella convalida industriale.

Parole chiave: Apprendimento per rinforzo; Q-learning; sinterizzazione; cicli termici; ceramici avanzati.

1. Introducción

En cuanto a procesamiento de materiales cerámicos se refiere los polvos preparados por vía coloidal, liofilización o sol-gel, terminan por tornarse un compacto en verde que alcanza sus propiedades físicas durante la sinterización, razón cual, las propiedades de la pieza dependen directamente de cómo se controle la etapa térmica (Monteiro et al., 2026). En cerámicas avanzadas como la alúmina, la zirconia o el carburo de silicio, los ciclos térmicos son extremos, con periodos de tiempo donde la temperatura llega hasta 1600 °C, y cualquier desviación produce gradientes de densidad, porosidad residual o agrietamiento por choque térmico (Gutiérrez Calderón, 2025).

No obstante, los ciclos se siguen definiendo por prueba y error: los periodos se ajustan manualmente para evitar la sobre sinterización(Gómez Domínguez, 2024), cosa que consume

recursos de producción y solo permite detectar defectos durante el transcurso de la misma (Gunupudi et al., 2026). El aprendizaje autónomo ya se aplica para relacionar el historial térmico con la densidad y la porosidad, aunque la escasez de datos estandarizados limita su paso a la producción (Monteiro et al., 2026). El aprendizaje por refuerzo va un paso más allá: un agente aprende a controlar el proceso observando un estado, ejecutando una acción y recibiendo una recompensa, dentro del marco de los procesos de decisión de Markov, y en su variante Q-learning lo hace sin necesitar un modelo explícito del sistema(Portillo-Ramírez et al., 2025).

Este estudio se orienta por la pregunta: ¿En qué medida el aprendizaje por refuerzo supera a los métodos experimentales tradicionales en la determinación de ciclos de sinterización cerámicos? Dicha pregunta se dividió en 3 incógnitas más específicas:

- ¿Qué tipo de limitaciones presentan los métodos tradicionales?
- ¿Existe evidencia cuantitativa acerca del aprendizaje por refuerzo y las técnicas de IA que lo apoyan?
- ¿existen vacíos cuales limitan su superioridad?

Con este fin se realizó una revisión sistemática de la literatura publicada entre 2024 y 2026 referente al tema de estudio.

2. Metodología

Se ejecutó una revisión sistemática de literatura (RSL) basada en documentos publicados entre enero de 2024 y noviembre de 2026. La búsqueda se realizó en noviembre de 2026 en repositorios institucionales: Séneca (Universidad de los Andes), GREDOS (Universidad de Salamanca) y el portal AJEA de la Universidad Tecnológica Nacional, y en las plataformas de editoriales académicas SpringerLink, MDPI, Wiley Online Library, Nature Portfolio, Frontiers e IntechOpen. Se combinaron términos en español e inglés mediante operadores booleanos: ("aprendizaje por refuerzo" OR "reinforcement learning" OR "Q-learning" OR "machine learning" OR "Bayesian optimization") AND ("sinterización" OR "sintering" OR "furnace temperature control") AND ("cerámicos avanzados" OR "advanced ceramics" OR "ceramics").

Se incluyeron documentos académicos (artículos, capítulos, tesis o actas) publicados entre 2024 y 2026 que abordaron la determinación del ciclo de sinterización cerámica o el control de procesos térmicos industriales análogos, ya sea con métodos experimentales tradicionales o con aprendizaje por refuerzo (incluso otras técnicas basadas en IA), y que reportan métricas comparables (tiempo, número de ensayos, consumo energético, densidad, porosidad, temperatura o errores de predicción). Se excluyeron textos divulgativos o comerciales y trabajos centrados solo en la síntesis química. La tesis de: (Gómez Domínguez, 2024); se incluyó, a pesar de tratar la cerámica

de residuos, debido que documenta con detalle el método experimental de referencia; también se admitieron estudios de RL en hornos no cerámicos como evidencia análoga. Las obras teóricas se emplearon únicamente como marco conceptual.

La evidencia se contrastó en seis criterios: tiempo y número de ensayos, consumo energético, calidad del sinterizado, adaptación en tiempo real, optimización de la temperatura y nivel de validación. En cada uno, la superioridad del RL se clasificó como:

- demostrada (comparación directa con datos experimentales o de planta)
- indicativa (evidencia simulada o de sistemas análogos)
- no evaluada.

La calidad de cada estudio se valoró con tres criterios: revisión por pares, métricas cuantitativas verificables y validación experimental o con datos de planta; la puntuación se clasificó como alta (3), media (2) o baja (0–1).

3. Resultados

Selección y caracterización general de los estudios incluidos

La muestra se distribuye en 2024, 2025 y 2026 . Seis son artículos publicados en revistas revisadas por pares y cuatro corresponden a una tesis doctoral, un trabajo de grado, un libro de actas y un capítulo de revisión. Predominan los enfoques cuantitativos de tipo experimental o computacional. Según la valoración de calidad, tres estudios obtuvieron calidad alta(Chen et al., 2025; Cheng & Luo, 2026; Saleem et al., 2026), cuatro calidad media (Bolívar et al., 2026; Gómez Domínguez, 2024; Kakanuru et al., 2024; Karpovich et al., 2024)y tres calidad baja (Gunupudi et al., 2026; Gutiérrez Calderón, 2025; Rectorado - Utn, 2024).

Los hallazgos se agruparon en cuatro categorías temáticas:

- control térmico convencional
- modelos sustitutos supervisados
- optimización bayesiana
- aprendizaje por refuerzo aplicado a materiales y al control térmico.

Síntesis de la evidencia por enfoque

El enfoque convencional, (Gómez Domínguez, 2024)documenta los ajustes manuales de los tiempos para evitar la sobre sinterización; (Gutiérrez Calderón, 2025) muestra que en cerámicas

avanzadas, con ciclos de hasta 1600 °C, esos ajustes no logran evitar gradientes de densidad, porosidad residual ni choque térmico.

Entre los modelos sustitutos, (Kakanuru et al., 2024) entrenaron SVR, CNN y un modelo de mezcla gaussiana con 10.000 simulaciones; con el modelo de Skorohod–Olevsky: la densidad final se predijo con MAPE de 0.17–0.42 % y la temperatura de mantenimiento con 1.1–2.2 %, pero el tiempo de mantenimiento llegó a 22.7–39.9 %; además, la calibración del material pasó de cerca de dos semanas a un aproximado de 90 s. Con datos experimentales, (Chen et al., 2025) predijeron la porosidad aparente de cerámicas de mullita-corindón, con $R^2 = 0.94$ y errores de validación inferiores a 5.1 %, aunque el módulo de rotura se desvió más de 20 % a altas temperaturas. (Saleem et al., 2026) entrenaron una red neuronal con 322 000 datos teóricos, la ajustaron con dilatometría y la integraron en ABAQUS, y (Gunupudi et al., 2026) reportan errores de densidad de 1–2 % en sistemas Cu–Al.

En cuanto a la optimización bayesiana, (Cheng & Luo, 2026) se obtuvieron cerámicas piezoeléctricas de 692 pC/N con solo 28 series de experimentos, mejoraron en 72 % la eficiencia de desarrollo y limitaron a menos de 8.3 % la degradación ante variaciones de ±20 °C.

En aprendizaje por refuerzo, (Bolívar et al., 2026) entrenaron un agente de Q-learning profundo con 200 acciones discretas sobre un sustituto de un horno industrial de inducción para acero (RMSE de 9.85 °C) y estimaron ahorros energéticos de 24–27 % frente al patrón de operación habitual, aún sin validar en planta. (Karpovich et al., 2024) usaron una DQN para diseñar composiciones con una temperatura de sinterización 111 °C menor, mientras que(Universidad Tecnológica Nacional, 2024) y (Gunupudi et al., 2026) presentan el RL como herramienta de control térmico multivariable y como línea emergente.

Comparación por criterio frente a los métodos experimentales tradicionales

En los seis criterios se muestra superioridad indicativa en consumo energético, adaptación en tiempo real y optimización de la temperatura; en tiempo, número de ensayos y en calidad del sinterizado no ha sido evaluado; además la ventaja medible corresponde a las técnicas de apoyo. En validación, el método tradicional conserva la ventaja, y ningún criterio alcanza superioridad demostrada (Tabla 1).

Research Club Scientia Vol. 4, No 4, Octubre - Diciembre, pp. 1821-1831, ISSN: 3103 – 1315

Tabla 1. *Grado de superioridad del aprendizaje por refuerzo frente a los métodos experimentales tradicionales, por criterio*

Criterio	**Métodos experimentales tradicionales**	**Aprendizaje por refuerzo**	**Técnicas de IA de apoyo**	**Grado de superioridad del RL**
Tiempo y número de ensayos	Prueba y error con monitoreo humano constante.	Sin estudios que lo midan en sinterización cerámica.	Calibración de ~2 semanas a ~90 s; 28 series de experimentos y 72 % más eficiencia.	No evaluado
Consumo energético	Ciclos estáticos de alto consumo.	Ahorro predicho de 24–27 % frente al patrón de operación habitual en un horno de inducción para acero.	—	Indicativo (sistema análogo y simulado)
Calidad del sinterizado	Defectos detectados al final del ciclo.	Sin estudios que lo midan.	Densidad final con MAPE de 0.17–0.42 %; error < 5.1 % con validación experimental.	No evaluado
Adaptación en tiempo real	Ajuste manual de mesetas a posteriori.	La política ajusta la potencia según el estado del proceso; control multivariable.	Degradación < 8.3 % ante ±20 °C.	Indicativo
Optimización de la temperatura	Definida por ensayos sucesivos.	DQN reduce 111 °C la temperatura de sinterización por diseño de composición.	—	Indicativo (diseño de material, no del ciclo)

Criterio	Métodos experimentales tradicionales	Aprendizaje por refuerzo	Técnicas de IA de apoyo	Grado de superioridad del RL
Nivel de validación	Método aplicado de forma rutinaria en producción.	Sin validación en planta ni en hornos cerámicos.	Validación experimental en laboratorio.	Ventaja del método tradicional

Nota. Elaboración propia a partir de los estudios incluidos en la revisión.

Vacíos identificados en la literatura

Se identificaron cuatro vacíos: ningún estudio aplica Q-learning a un horno de sinterización cerámica; varios modelos precisos se entrenan con datos sintéticos y se validan fuera de su rango de entrenamiento; faltan conjuntos de datos experimentales estandarizados que relacionen el ciclo térmico con la densidad, la porosidad y las tensiones residuales; ningún trabajo reporta la validación industrial de un controlador inteligente del ciclo.

4. Discusión

Los resultados de esta revisión permiten responder a la pregunta orientadora: el aprendizaje por refuerzo supera a los métodos experimentales tradicionales en la determinación de ciclos de sinterización cerámicos solo en una medida parcial. La superioridad es indicativa en cuanto consumo energético se refiere; con ahorros predichos de 24–27 % frente al patrón de operación habitual de un horno industrial (Bolívar et al., 2026) , en capacidad de adaptación en tiempo real y reducción de temperatura en proceso de sinterización por diseño de composición (Karpovich et al., 2024); sin embargo, ninguno de estos resultados procede de un horno de sinterización cerámica ni ha sido verificado en planta. Por otra parte, la ventaja medible frente a la prueba y error ya está demostrada para las técnicas que acompañan al RL: los modelos sustitutos y la optimización bayesiana reducen de semanas a segundos el cálculo del ciclo y disminuyen drásticamente el número de ensayos(Cheng & Luo, 2026; Kakanuru et al., 2024).

En primer lugar, se observó que la principal limitación del enfoque convencional no es la falta de conocimiento físico del proceso, sino la rigidez del control. Los estudios de: (Gutiérrez Calderón, 2025; Kakanuru et al., 2024); coinciden en que las variables de calidad (densidad, porosidad, contracción) están bien definidas, pero se corrigen a posteriori y de forma manual. Esta observación es especialmente crítica en cerámicas avanzadas, donde los márgenes de tolerancia son estrechos y un error en la rampa o en loa periodos de temperatura estable, se traduce en piezas defectuosas y en pérdidas energéticas difíciles de recuperar.

La dificultad para identificar el tiempo de mantenimiento en el trabajo de: (Kakanuru et al., 2024); puede explicarse porque la densidad tiende a estabilizarse antes de que finalice el periodo de temperatura estable, lo que vuelve poco sensible la salida del modelo a ese parámetro; esto sugiere que el tiempo de mantenimiento es precisamente la variable en la que un controlador adaptativo podría aportar mayor ahorro. A su vez, el contraste entre los errores inferiores a 0.5 % obtenidos con datos simulados y los cercanos a 5 % reportados por: (Chen et al., 2025);con datos experimentales muestra el costo real de pasar del simulador al laboratorio, y respalda estrategias híbridas como la de(Saleem et al., 2026), que entrenan primero con datos teóricos y después ajusta el modelo con mediciones de dilatometría.

La revisión permite formular cómo se integraría un agente de Q-learning en un horno de sinterización. El estado estaría definido por la temperatura actual, la densidad o contracción estimada y la etapa del ciclo; las acciones corresponden al ajuste de la potencia, la velocidad de rampa o la duración de los periodos de temperatura estable; siendo recompensado la aproximación a la densidad máxima, penalizando la porosidad residual, las tensiones residuales y el consumo energético. Las métricas definidas por (Gómez Domínguez, 2024) se traducen así en componentes de la función de recompensa. Esta arquitectura ya no es solo teórica:(Bolívar et al., 2026) implementaron está en un horno de inducción, con un gemelo digital, un modelo sustituto como entorno de entrenamiento y una recompensa escalonada que penaliza las desviaciones de temperatura antes de premiar el ahorro energético. Su trasposición a la sinterización cerámica, empleando entornos sustitutos como los de(Kakanuru et al., 2024), es una propuesta derivada del presente análisis.

La comparación entre la optimización bayesiana y el aprendizaje por refuerzo ayuda a precisar la respuesta a la pregunta orientadora. La optimización bayesiana busca el mejor ciclo estático con el menor número de ensayos, de modo que sustituye directamente al diseño de experimentos por prueba y error incluso puede incorporarse de inmediato a la práctica de laboratorio, como muestran (Cheng & Luo, 2026) con 28 series de experimentos. El aprendizaje por refuerzo, en cambio, no entrega un ciclo fijo, en su lugar muestra una política que decide en cada instante cómo ajustar la potencia o la meseta según el estado del material. Esta diferencia explica por qué el RL promete una ventaja mayor “un ciclo que se corrige mientras ocurre”, a la vez, por qué exige más: necesita miles de episodios de entrenamiento que no pueden realizarse en un horno real, sino sobre un entorno simulado. Ambas técnicas son, por tanto, complementarias y no competidoras.

Esta complementariedad también se observa en el manejo de los datos.(Chen et al., 2025) demostraron que es posible construir un modelo útil reuniendo 495 muestras de 22 publicaciones, lo que indica que la literatura existente ya puede servir para entrenar los primeros sustitutos sin realizar nuevas campañas experimentales. Sin embargo, su validación muestra que propiedades como la resistencia a flexión se predicen ineficientemente a temperaturas elevadas, justamente la

zona más crítica del ciclo de sinterización. Para un agente de RL, un entorno impreciso en esa zona se traducirá en políticas que parecen óptimas en la simulación, pero que fallan en un entorno real; por ello, la calidad del sustituto condiciona directamente cuánto puede superar el RL al método tradicional.

No por ende, la aplicación de Q-learning a la sinterización enfrenta tres desafíos técnicos que la literatura revisada permite precisar. El primero es la representación del estado: el Q-learning clásico opera con tablas de estados y acciones discretos, mientras que la temperatura y la densidad son variables continuas; categorizar las finamente hace crecer la tabla de forma exponencial, por lo que los trabajos recientes recurren a redes Q profundas que aproximan la función de valor, manteniendo discretas solo las acciones, como las 200 acciones de ajuste de potencia de (Bolívar et al., 2026) o la DQN de (Karpovich et al., 2024). El segundo es la transferencia de la simulación al horno real: una política entrenada sobre un sustituto hereda sus errores, y por ello(Bolívar et al., 2026) advierten que sus ahorros requieren reevaluación por elementos finitos y validación en planta. El tercero es la seguridad durante el aprendizaje: un agente que explora acciones al azar puede provocar choque térmico o afectar a las piezas y el equipo, de modo que conviene entrenar fuera de línea sobre el sustituto, acotar físicamente las acciones: como el límite de 0 a 1000 kW impuesto por(Bolívar et al., 2026); y penalizar con fuerza en la recompensa las desviaciones peligrosas.

Aunque, la presente revisión debe interpretarse considerando algunas limitaciones metodológicas. Entre ellas destacan el tamaño de la muestra, la inclusión de literatura gris (tesis, actas y un capítulo) junto con publicaciones revisadas por pares, la ausencia de métricas homogéneas entre estudios y el hecho de que buena parte de la evidencia cuantitativa provenga de datos simulados. Así mismo, la evidencia sobre RL procede de hornos de acero y de diseño computacional de materiales, lo que exige cautela al extrapolar su desempeño a hornos cerámicos de alta temperatura.

5. Conclusiones

En respuesta a la pregunta orientadora, el aprendizaje por refuerzo supera a los métodos experimentales tradicionales en la determinación de ciclos de sinterización cerámicos solo de forma parcial y potencial: su ventaja es indicativa en consumo energético, adaptación en tiempo real y optimización de la temperatura, pero no ha sido aplicado en hornos ceramicos, y el método tradicional conserva la ventaja en validación industrial.

Los métodos experimentales de prueba y error, basados en ciclos estáticos y ajustes manuales, resultan insuficientes para cerámicos avanzados que exigen perfiles extremos y tolerancias estrechas, ya que solo permiten corregir los defectos una vez concluido el ciclo; frente a ellos, los modelos sustitutos y la optimización bayesiana ya ofrecen una superioridad medible en tiempo,

Research Club Scientia Vol. 4, No 4, Octubre - Diciembre, pp. 1821-1831, ISSN: 3103 – 1315

número de ensayos y precisión, y constituyen el entorno de entrenamiento natural para futuros agentes de control.

Para que el aprendizaje por refuerzo alcance una superioridad demostrada se requiere resolver la representación continua del estado, la transferencia de la simulación al horno real y la seguridad durante el aprendizaje, así como realizar comparaciones directas frente a ciclos experimentales de referencia con mediciones de densidad, porosidad y consumo energético; estas condiciones definen la principal línea de investigación futura.

En el corto plazo, la vía más realista para superar a los métodos de prueba y error no es sustituirlos de inmediato por un agente de aprendizaje por refuerzo, sino combinar modelos sustitutos entrenados con datos de la literatura y ensayos propios con optimización bayesiana, que ya ha demostrado reducir el número de experimentos y el tiempo de desarrollo; esta etapa genera, además, los datos y el entorno simulado que el RL necesita.

Como aporte, esta revisión ofrece un marco de comparación en seis criterios que permite valorar de forma objetiva cuándo una técnica inteligente supera al método tradicional. Para futuras investigaciones se recomienda diseñar estudios comparativos con un mismo material y un mismo horno, contrastando un ciclo experimental de referencia con la política aprendida por el agente y midiendo densidad, porosidad, tensiones residuales, consumo energético y tiempo total del ciclo.

Referencias

Bolívar, S., Lloret, L., Sáinz-Aja, J. A., Ferreño, D., Oria, C., Ruiz, E., & Cobo, M. (2026). Optimization of the energy consumption of an induction furnace for steel billets using a reinforcement learning framework applied to a finite element digital twin. *Journal of Intelligent Manufacturing*. https://doi.org/10.1007/s10845-026-02933-0

Chen, Q., Zhang, W., Liang, X., Feng, H., Xu, W., Wang, P., Pan, J., & Cheng, B. (2025). Machine Learning-Assisted Multi-Property Prediction and Sintering Mechanism Exploration of Mullite–Corundum Ceramics. *Materials*, *18*(6), 1384. https://doi.org/10.3390/ma18061384

Cheng, T., & Luo, J. (2026). Bayesian optimization driven co-regulation of composition design and sintering process parameters for piezoelectric ceramics. *Frontiers in Materials*, *13*, 1906244. https://doi.org/10.3389/fmats.2026.1906244

Gómez Domínguez, J. I. (2024). *Estudio experimental y numérico de materiales cerámicos procedentes de residuos de construcción*. https://doi.org/10.14201/gredos.159705

Gunupudi, C., Faruque, O., Mondal, S., Reddy Rachepalli, J., & Reddy Annapareddy and Kalyana Krishna Kondapalli, N. (2026). The Application of Artificial Intelligence (AI) in Sintering

Technology. En M. Simona Bălțatu & P. Vizureanu (Eds.), *Sintering Technologies in Materials Design and Processing* (Vol. 26). IntechOpen. https://doi.org/10.5772/intechopen.1015341

Gutiérrez Calderón, J. A. (2025). *Desarrollo de una propuesta de solución innovadora para la manufactura aditiva de cerámicos avanzados mediante la integración de microinyección de polvos y extrusión 3D*. https://hdl.handle.net/1992/76611

Kakanuru, P., Terway, P., Jha, N., & Pochiraju, K. (2024). Process–Material–Performance Trade-off Exploration of Materials Sintering with Machine Learning Models. *Integrating Materials and Manufacturing Innovation*, *13*(4), 927-941. https://doi.org/10.1007/s40192-024-00380-4

Karpovich, C., Pan, E., & Olivetti, E. A. (2024). Deep reinforcement learning for inverse inorganic materials design. *Npj Computational Materials*, *10*(1), 287. https://doi.org/10.1038/s41524-024-01474-5

Monteiro, S. T., Monteiro, I. A. T., Fernandez-Gamboa, J. R., & Ruiz, Y. L. (2026). Emerging Machine Learning Strategies for Digital Processing and Additive Manufacturing of Advanced Ceramics. *International Journal of Applied Ceramic Technology*, *23*(3), e70195. https://doi.org/10.1111/ijac.70195

Portillo-Ramírez, G., Cruz-Suárez, H., & Velasco-Luna, F. (2025). Un panorama general del aprendizaje por refuerzo en líneas de espera controladas a tiempo discreto. *CIENCIA ergo-sum*, *32*, e290. https://doi.org/10.30878/ces.v32n0a29

Rectorado - Utn, S. D. P. A. Y. P. (2024). 7as Jornadas de Estudiantes de Doctorado UTN. *AJEA*, (AJEA 37). https://doi.org/10.33414/ajea.1741.2024

Saleem, B., Polak, P., He, R., Patel, K., Phillips, J., Savva, S., & Pan, J. (2026). A machine learning approach to finite element modeling of sintering deformation using densification data. *International Journal of Applied Ceramic Technology*, *23*(1), e70105. https://doi.org/10.1111/ijac.70105

Universidad Tecnológica Nacional. (2024). 7as Jornadas de Estudiantes de Doctorado UTN. *AJEA (Actas de Jornadas y Eventos Académicos de UTN)*, (37). https://doi.org/10.33414/ajea.1741.2024