

# **Spinlock SRL: Diseño de Sistema de Transferencia y Mantenimiento de Temperatura en Muestras**

Fotinós, Jeronimo <sup>\*1</sup>, y Stauber López, Daniela <sup>\*\*1</sup>

<sup>1</sup>Instituto de Física Enrique Gaviola (IFEG) - CONICET

Abril 2025

## **Resumen**

El presente trabajo aborda posibles mejoras al muestreador automático SLK AS 30 de la empresa Spinlock SRL. El mismo es un equipo que utiliza resonancia magnética nuclear para medir la presencia de distintos compuestos en semillas y otras muestras similares. Para lograr mediciones más precisas, se requiere homogeneizar la temperatura de las muestras en el equipo. A tal fin, proponemos tres enfoques. El primero se trata de mejoras en el diseño y en la elección de materiales para distintas partes del equipo. El segundo consta de simular el funcionamiento del equipo mediante elementos finitos para evaluar distintas configuraciones del mismo. Finalmente, se plantea un esquema de aprendizaje por refuerzos para lograr un control dinámico y robusto de las temperaturas. Si bien no brindamos una solución cerrada al problema, pretendemos que estos lineamientos provean un marco en el cual lograr la optimización del equipo.

## **1. Introducción y Motivación**

El presente trabajo se desarrolla en el marco del Sexto Taller de Matemática Industrial, VI TAMI 2025, llevado a cabo en la FaMAF, UNC, entre el 3 y el 7 de febrero de 2025, de manera presencial. En el mismo, las empresas participantes presentaban problemáticas que limitaban sus procesos productivos, y luego se asignaba a cada una un grupo de estudiantes para intentar encontrar soluciones a las mismas. En el presente trabajo, nos proponemos dar lineamientos para solucionar la problemática planteada desde la empresa Spinlock SRL.

---

<sup>\*</sup>[jerofti@gmail.com](mailto:jerofti@gmail.com)

<sup>\*\*</sup>[daniela.stauber@unc.edu.ar](mailto:daniela.stauber@unc.edu.ar)

Spinlock es una empresa tecnológica argentina especializada en el desarrollo y fabricación de equipos de Resonancia Magnética Nuclear (RMN) y Resonancia Cuadripolar Nuclear (RCN) orientados a aplicaciones industriales. Su tecnología permite a distintos sectores controlar la calidad de sus productos mediante analizadores como el de contenido de aceite polimérico, entre otros. Actualmente, Spinlock brinda soluciones al sector agroalimentario, textil, de biocombustibles y de Oil & Gas [6].

Uno de sus dispositivos destacados es el muestreador automático SLK AS 30 (ver Fig. 1). Este equipo extrae una muestra desde un módulo con tubos de ensayo, la pesa y la introduce en un analizador Spinlock para su medición. Durante el proceso se determina el contenido de aceites, grasas y humedad en semillas oleaginosas, granos y productos alimenticios como girasol, canola, maní, soja, maíz, olivas, lácteos, entre otros. Finalizado el análisis, la muestra puede ser devuelta a su posición original o reubicada según un criterio definido.

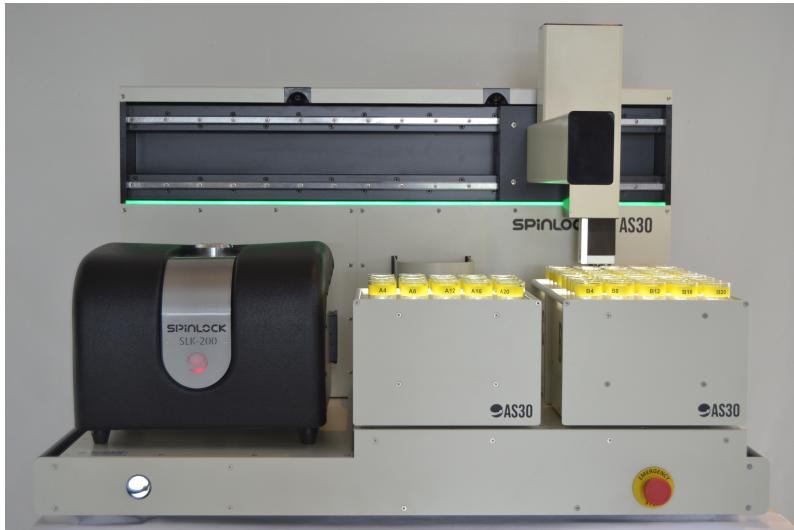


Figura 1: Muestreador automático SLK AS 30

El muestreador cuenta con tres racks, ilustrados en la figura 2, con capacidad para 20 muestras cada uno. Están fabricados en aluminio, aislante térmico y polietileno de alta densidad. Las muestras se sitúan dentro de tubos, localizados en la zona enmarcada en la imagen 2(c).

Con vistas a acceder a nuevos sectores del mercado, desde la empresa quieren aumentar la precisión en las mediciones de sus equipos. El principal obstáculo a tal fin, es la incertidumbre y la variabilidad en la temperatura de las muestras al entrar al medidor luego de ser pesadas. Para reducir la misma, habría que lograr un mayor control en la temperatura dentro de los racks, de modo que las muestras partan de la temperatura objetivo. Idealmente, la empresa quiere homogeneizar y estabilizar la temperatura en la zona de las muestras dentro de cada rack del robot completo. Este opera en salas a una temperatura ambiente de aproximadamente 25°C, mientras que se requiere mantener las muestras a 27°C, con una tolerancia de  $\pm 0.5^\circ\text{C}$ . La temperatura en la zona de interés esta controlada por resistencias ubicadas por debajo de las muestras. La razón de querer mantener una temperatura ligeramente superior a la ambiente es evitar la necesidad de enfriar, cosa que aumentaría signifi-

ficativamente los costos. Manteniendo la temperatura en torno a valores más altos que la temperatura ambiente, alcanza con tener resistencias disipando calor para lograr el control requerido.

La empresa planteó múltiples grados de libertad para abordar el problema, incluyendo el tipo de materiales utilizados en los racks, sus dimensiones — dentro de un margen de 1cm, así como el diseño, número y ubicación de las fuentes de calor. Asimismo, también se incluían restricciones de diversa índole: limitaciones en el peso total de los racks; limitaciones en el costo total, lo que a su vez restringe el número de sensores y actuadores; restricciones sobre el tipo de materiales visibles desde el exterior para no dañar el atractivo visual del equipo; etc.

Considerando que esta es la primera aproximación al problema, y que la empresa todavía no contaba con un framework previo bajo el cual abordar el mismo, nuestro objetivo será brindar lineamientos a seguir para lograr la precisión necesaria. Estos serán dados en orden de prioridad, que a su vez se condice con un orden de complejidad creciente. Como se indicará en la sección de Métodos, comenzaremos dando algunas sugerencias respecto al diseño del equipo. Luego, indicaremos como podrían simularse diferentes configuraciones del equipo para evaluar su rendimiento. Finalmente, expondremos cómo implementar un sistema de aprendizaje por refuerzos para lograr un control dinámico, robusto y en tiempo real de la temperatura dentro de los racks.

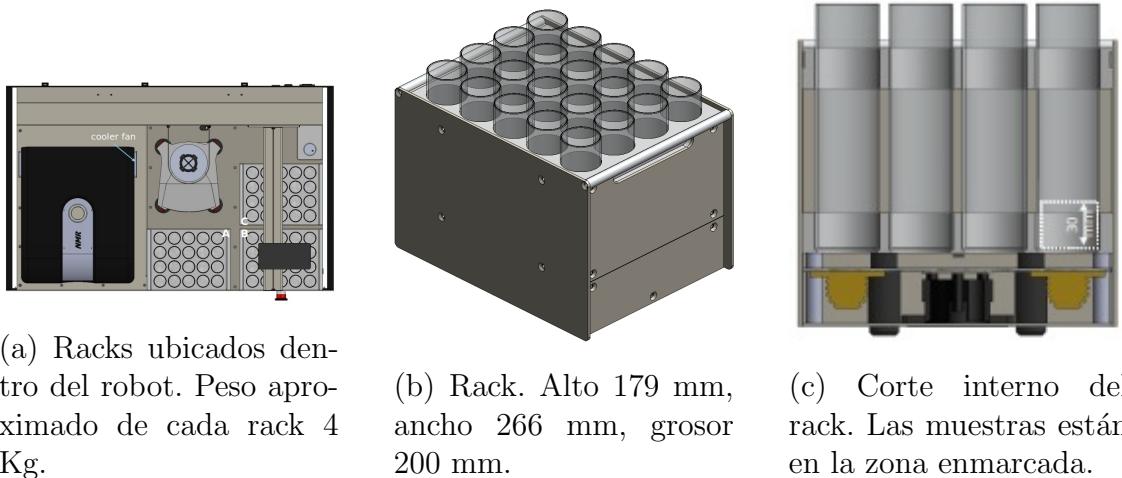


Figura 2: Vistas del rack y su integración en el sistema.

## 2. Preliminares

### 2.1. Ecuación de calor

La ecuación de calor es una ecuación en derivadas parciales que describe cómo varía la temperatura  $u(\mathbf{x}, t)$  en un medio con el tiempo  $t$  y el espacio  $\mathbf{x}$ . En su forma más general, en un medio tridimensional, se expresa como:

$$\rho c_p \frac{\partial u}{\partial t} = \nabla \cdot (k \nabla u) + Q, \quad (1)$$

donde:

- $\rho$  es la densidad del material [ $\text{kg}/\text{m}^3$ ],
- $c_p$  es el calor específico a presión constante [ $\text{J}/(\text{kg} \cdot \text{K})$ ],
- $k$  es la conductividad térmica del medio [ $\text{W}/(\text{m} \cdot \text{K})$ ],
- $Q$  representa una fuente o dissipación interna de calor [ $\text{W}/\text{m}^3$ ],
- $u$  es la temperatura en función del espacio y del tiempo [K].

Esta ecuación modela la difusión de calor considerando tanto la capacidad calorífica del material como su capacidad de conducción térmica. La presencia del término fuente  $Q$  permite representar elementos activos como resistencias calefactoras.

En el contexto de este trabajo, uno de los objetivos es modelar la distribución de temperatura dentro de un rack. El mismo está compuesto por distintos materiales —como aluminio, polietileno de alta densidad y aislantes térmicos— lo que implica que los coeficientes  $\rho$ ,  $c_p$  y  $k$  varían espacialmente. Además, el rack contiene resistencias internas que actúan como fuentes de calor controladas.

El sistema completo consta de tres de ellos ubicados dentro de un robot que no está térmicamente aislado. Por este motivo, cada rack no solo presenta una distribución de temperatura interna compleja, sino que también se ve influenciado por la proximidad de los otros, así como por elementos del entorno como un ventilador y el ambiente general del laboratorio. Estas condiciones generan un acoplamiento térmico indirecto entre racks, especialmente relevante dado que su disposición no es simétrica dentro del robot, lo que produce una interacción térmica heterogénea entre ellos.

Desde el punto de vista del modelado, resulta fundamental especificar condiciones de borde apropiadas. Para este tipo de problema, se consideran principalmente las siguientes:

- **Condiciones de Dirichlet:** Se podrían imponer en zonas de contacto térmico fijo con el entorno, como superficies en contacto con disipadores o partes estructurales del robot que mantengan una temperatura constante conocida.
- **Condiciones de Neumann:** Son útiles para modelar flujos de calor impuestos, como el que proviene directamente de las resistencias internas. También pueden usarse para modelar aislamiento térmico local (flujo nulo).
- **Condiciones de Robin (mixtas):** Permiten modelar el intercambio de calor por convección con el ambiente, lo cual es apropiado para las superficies expuestas al aire del laboratorio o a la corriente del ventilador. Tienen la forma:

$$-k \frac{\partial u}{\partial n} = h(u - u_{\text{amb}}),$$

donde  $h$  es el coeficiente de convección y  $u_{\text{amb}}$  es la temperatura del aire circundante.

Estas condiciones de borde, junto con la variabilidad espacial de los parámetros materiales y las fuentes internas de calor, permiten capturar de manera realista la dinámica térmica dentro de cada rack y entre ellos. El modelo obtenido será la base para simulaciones numéricas y para el diseño de estrategias de control térmico mediante métodos como el aprendizaje por refuerzo.

## 2.2. Ecuaciones de Navier-Stokes

En el contexto de este trabajo, los racks que contienen la región de muestras no se encuentran térmicamente aislados. De hecho uno de ellos, está ubicado en las cercanías de un ventilador, cuya acción introduce un flujo de aire caliente que puede influir en la evolución térmica del sistema. Este fenómeno implica la existencia de convección forzada, que debe ser modelada adecuadamente para estimar el efecto del movimiento del fluido sobre la temperatura de las muestras.

Para este tipo de análisis, se utilizan las ecuaciones de Navier-Stokes, que describen el comportamiento dinámico de los fluidos. Estas ecuaciones surgen de la aplicación de los principios de conservación de masa y cantidad de movimiento. Para un fluido incompresible y newtoniano, se escriben como:

$$\rho \left( \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = -\nabla p + \mu \nabla^2 \mathbf{v} + \mathbf{f}, \quad (2)$$

$$\nabla \cdot \mathbf{v} = 0, \quad (3)$$

donde:

- $\mathbf{v}(\mathbf{x}, t)$  es el campo de velocidades del fluido,
- $p(\mathbf{x}, t)$  es el campo de presiones,
- $\rho$  es la densidad del fluido [ $\text{kg}/\text{m}^3$ ],
- $\mu$  es la viscosidad dinámica [ $\text{Pa} \cdot \text{s}$ ],
- $\mathbf{f}$  representa fuerzas externas como la gravedad o interacciones térmicas.

El término no lineal  $\mathbf{v} \cdot \nabla \mathbf{v}$  representa la auto-advección del fluido, y es responsable de buena parte de la complejidad del sistema. La ecuación (3) impone la conservación del volumen para fluidos incompresibles, como el aire a velocidades moderadas.

La incorporación del flujo generado por el ventilador se modela resolviendo el sistema de ecuaciones de Navier-Stokes acoplado con la ecuación de calor, lo que permite capturar el transporte de energía térmica por convección forzada. Este acoplamiento resulta especialmente relevante, ya que el rack no se encuentra térmicamente aislado: cualquier variación en la temperatura ambiente del laboratorio, como la proximidad a un aire acondicionado u otras corrientes de aire, puede influir directamente en la temperatura de las muestras.

El análisis del comportamiento térmico en presencia de flujo de aire es, por lo tanto, fundamental para comprender la distribución real de temperaturas en el sistema y diseñar estrategias de control térmico precisas y robustas.

No obstante, debido a la complejidad computacional que implica la simulación completa del sistema acoplado y el tiempo requerido para su resolución, este enfoque se presenta aquí como una posible extensión del modelo. Por lo tanto, no será considerado en las primeras simulaciones del trabajo.

### 2.3. Elementos finitos y formulación débil

El método de los elementos finitos (FEM, por sus siglas en inglés) es una técnica numérica utilizada para aproximar soluciones de ecuaciones diferenciales parciales, especialmente en dominios con geometrías complejas o condiciones de contorno variadas. Su principal ventaja es que permite descomponer un dominio físico arbitrario en una colección de subdominios más simples llamados *elementos* (por ejemplo, triángulos en 2D o tetraedros en 3D), sobre los cuales se aproximan las soluciones mediante funciones base definidas localmente.

En el contexto de la conducción de calor, el FEM permite simular cómo evoluciona la temperatura en el tiempo y el espacio dentro de un objeto sometido a fuentes térmicas y disipación. Para ello, se parte de la ecuación de calor en su forma fuerte:

$$\rho c_p \frac{\partial u}{\partial t} = \nabla \cdot (k \nabla u) + Q \quad \text{en } \Omega \times (0, T], \quad (4)$$

Para aplicar el FEM, se reformula la ecuación anterior en su versión *débil*, lo cual implica multiplicar por una función de prueba  $v$  y aplicar el teorema de la divergencia (o integración por partes), obteniendo:

$$\int_{\Omega} \rho c_p \frac{\partial u}{\partial t} v \, d\mathbf{x} = - \int_{\Omega} k \nabla u \cdot \nabla v \, d\mathbf{x} + \int_{\Omega} Q v \, d\mathbf{x} + \int_{\partial\Omega} k \nabla u \cdot \mathbf{n} v \, ds, \quad (5)$$

Esta expresión se interpreta como: buscar una función  $u(t) \in V$ , con  $V$  un espacio funcional adecuado (por ejemplo,  $H^1(\Omega)$ ), tal que la ecuación se cumpla para toda función de prueba  $v \in V$  y para cada instante de tiempo  $t \in (0, T]$ .

Luego, el dominio  $\Omega$  se discretiza mediante una malla de elementos finitos y las funciones  $u$  y  $v$  se aproximan como combinaciones lineales de funciones base locales. Esto da lugar a un sistema algebraico de la forma:

$$M \frac{d\mathbf{u}}{dt} + K\mathbf{u} = \mathbf{f} \quad (6)$$

donde:

- $M$  es la matriz de masa (asociada al término temporal),
- $K$  es la matriz de rigidez (asociada al flujo térmico),
- $\mathbf{f}$  representa el vector de fuentes,
- $\mathbf{u}$  es el vector de temperaturas nodales.

Este sistema se resuelve mediante técnicas de integración temporal como Euler implícito, Crank-Nicolson o Runge-Kutta, dependiendo del compromiso entre precisión y estabilidad numérica.

El uso de FEM resulta especialmente útil en problemas como el presente, donde se involucran materiales con distintas propiedades térmicas, geometrías no triviales y fuentes de calor distribuidas, permitiendo modelar con alta fidelidad la evolución térmica del sistema.

## 2.4. Aprendizaje por refuerzos

El concepto de aprendizaje automático está definido de manera amplia a fin de incluir cualquier programa que mejore su performance en una tarea dada por medio de la experiencia. En su libro clásico, Tom Mitchell define que decimos que una computadora aprende de una experiencia  $E$ , respecto a un conjunto de tareas  $T$  y una medida de performance  $P$ , si esta medida de la performance  $P$  aumenta para las tareas  $T$  con la experiencia  $E$  [3]. Podríamos decir que el aprendizaje automático es un conjunto de técnicas que consisten en obtener funciones o asignaciones a partir de un conjunto de datos. Una clasificación frecuente dentro del aprendizaje automático consiste en dividir las técnicas en tres categorías principales: aprendizaje supervisado, aprendizaje no supervisado, y aprendizaje por refuerzos.

El aprendizaje supervisado está genéricamente enfocado en problemas de clasificación y regresión. Esto es, dado un conjunto de datos de entrada  $X$  y sus correspondientes salidas  $Y$ , queremos encontrar una función  $f : X \rightarrow Y$ . Si  $Y$  es discreto y finito decimos que tenemos un problema de clasificación, mientras que en caso contrario decimos que tenemos un problema de regresión. La medida de performance utilizada para estos problemas suele ser una medida del acuerdo entre las salidas  $y' = f(x)$ ,  $x \in X$  proporcionadas por la función encontrada  $f$  y las salidas buscadas  $y \in Y$  correspondientes a ese valor  $x \in X$ .

El aprendizaje no supervisado aborda problemas de agrupamiento (clustering) y estratificación. Esto es, contamos con un conjunto de datos  $X$  y queremos encontrar subconjuntos (i.e., una partición) conveniente. Una posible medida de performance para este caso premiaría la similitud de datos en el mismo subconjunto a la vez que desalentaría la similitud entre datos en subconjuntos distintos.

El aprendizaje por refuerzos (RL por sus siglas en inglés), por su parte, permite abordar *problemas de decisión*, es decir, elegir una acción dado un estado del sistema. En este caso, lo que buscamos aprender es una política  $\pi$  definida del espacio de estados observables del sistema en un conjunto de acciones  $\pi : S \rightarrow A$ . Podemos pensar que esta política refleja la toma de decisiones (o acciones) de un agente que interactúa con su entorno. Iterativamente, nuestro agente realiza una observación del estado de su entorno - que puede ser un reflejo parcial o infiel del estado real, selecciona una acción, la ejecuta, y recibe una retroalimentación del entorno como resultado de su acción que llamamos *recompensa* (ver Fig. 3). Esta recompensa puede ser negativa, penalizando la acción del agente, o positiva, recompensándola. En aprendizaje por refuerzo, la medida de performance (la función a extremar, en este caso maximizar) es la recompensa total recibida por el agente.

Vale la pena destacar que en aprendizaje por refuerzos no contamos en general con un conjunto de datos  $X$  que actúen como ejemplos o experiencia. En estos casos, es el agente quien debe generarlos interactuando con el entorno.

Inicialmente uno podría definir la política  $\pi(s)$  como una función que devuelve

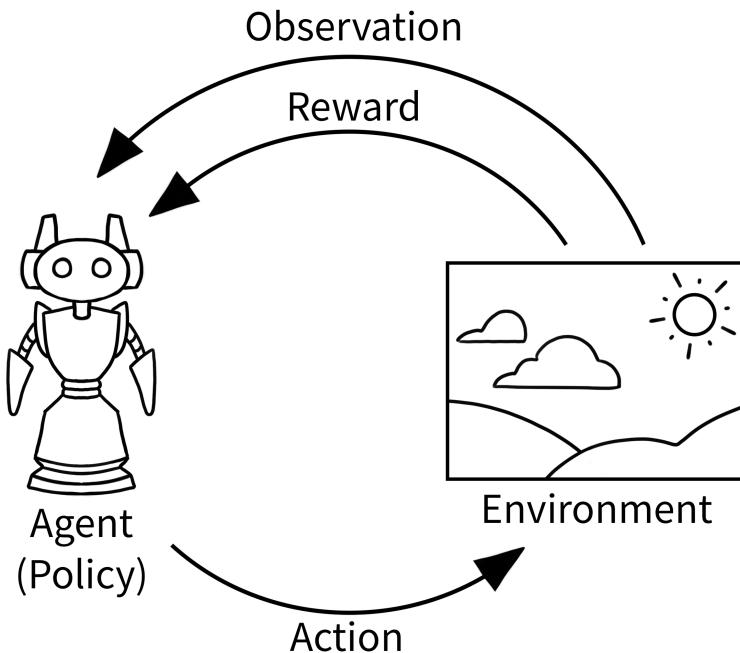


Figura 3: *Ciclo de interacción agente-entorno.* Tomado de la documentación de Gymnasium.

la acción  $a$  a realizar dado un estado  $s$ . Sin embargo, en general resulta conveniente definir la política como una distribución de probabilidad  $\pi(a|s)$  sobre las acciones posibles  $a$ , condicionada al estado del sistema  $s$ . De este modo, el agente buscará aprender la política  $\pi$  que maximizará su recompensa. Esto es, lo que obtendremos del entrenamiento es precisamente esta política.

En casos donde el estado actual del sistema es suficiente para decidir la acción a tomar a continuación (o equivalentemente, el estado actual resume el historial de estados), podemos plantear matemáticamente el problema como un proceso de decisión de Markov o MDP por sus siglas en inglés (Markov-Decision Process). Este consta de:

- un conjunto de estados del sistema  $S$ ;
- un conjunto de acciones disponibles para el agente  $A$ ;
- una función de recompensa  $R(s, a)$ , asociada a la acción que se acaba de ejecutar y al estado del sistema en ese momento;
- una distribución de probabilidad de transición  $P(s_{t+1} = s'|s, a)$  del estado actual, dada la acción que se acaba de tomar, al nuevo estado;
- y posiblemente un factor de descuento  $\gamma$  usado para que la recompensa acumulada no diverja cuando los episodios son muy largos.

Una cantidad que resulta útil definir es la *función de Estado-Valor*  $V^\pi(s)$ . Esta se define como el valor esperado de la recompensa acumulada en lo que queda del episodio, dado que me encuentro en el estado  $s$ . Su utilizad radica en ser una medida

de lo que vale para el agente llegar a dicho estado  $s$ . De hecho, un conjunto de técnicas muy exitosas para resolver problemas de aprendizaje por refuerzo son los llamados algoritmos actor-crítico. Dichos algoritmos combinan las ventajas de la generación de políticas (actor) con las de estimación del valor (crítico). Entrenan una función  $\pi(s, a)$  para seleccionar acciones, y un crítico  $\hat{V}(s)$  que estima el valor de los estados.

El algoritmo usado en nuestra propuesta para resolver el problema de aprendizaje por refuerzos es Proximal Policy Optimization (PPO) [5], el algoritmo estrella de gradiente de política. Combina el enfoque de los algoritmos actor-crítico con un mecanismo para limitar cuánto pueden variar las nuevas políticas. A su vez, este permite trabajar tanto con espacios de acciones continuas como discretas. En pocas palabras, se trata de un algoritmo de aprendizaje profundo (Deep RL) que utiliza redes neuronales para aprender tanto la política  $\pi(s, a)$  como el crítico (función estado-valor)  $V(s)$  a partir del estado del sistema. La desventaja más destacable es, quizás, que al ser un algoritmo *on-policy* no es posible usar un buffer de experiencias (memoria), lo cual incrementa el coste de entrenamiento.

### 3. Metodología

En esta sección se detalla la metodología adoptada por el equipo para abordar la problemática planteada. Nuestra propuesta se divide en dos enfoques complementarios:

- **Un enfoque cualitativo**, centrado en el análisis y rediseño estructural del sistema, en particular del rack que alberga la región de interés térmico, proponiendo mejoras en los materiales y disposición de los elementos que lo componen.
- **Un enfoque cuantitativo**, motivado por el margen de error térmico buscado por la empresa, en el que desarrollamos una estrategia de control basada en simulaciones físicas del entorno acopladas con algoritmos de aprendizaje por refuerzo. Esta aproximación busca optimizar el comportamiento del sistema ante diferentes escenarios térmicos.

Ambos enfoques fueron concebidos de manera complementaria: mientras el primero apunta a mejoras estructurales que reduzcan las perturbaciones externas y faciliten un entorno más estable, el segundo busca diseñar un sistema inteligente capaz de adaptarse y compensar dinámicamente dichas perturbaciones, logrando un control más robusto y eficiente de la temperatura en la zona de muestras.

#### 3.1. Enfoque cualitativo

La empresa nos proporcionó los planos detallados de cada rack (ver imagen 4 o en el repositorio), incluyendo la descripción precisa de los materiales que componen cada sección. Además, compartieron un mapa de calor correspondiente al rack B (ver imagen 5, asociado a la figura 2(a)), obtenido experimentalmente mediante el desplazamiento de un sensor a lo largo de los tubos, a la altura aproximada de la región de muestras  $\Gamma$ .

Este mapa permite observar variaciones significativas de temperatura dentro del rack, lo cual evidencia la complejidad del entorno térmico. Dada la ubicación del rack B dentro del sistema, es razonable inferir que su distribución térmica está influenciada por el rack A, ubicado en su lateral izquierdo, y el rack C, posicionado en su parte superior. A su vez, el sistema no se encuentra aislado, por lo que el rack también experimenta transferencia de calor por convección desde el ambiente exterior, como puede observarse por las diferencias térmicas en las zonas periféricas del dominio.

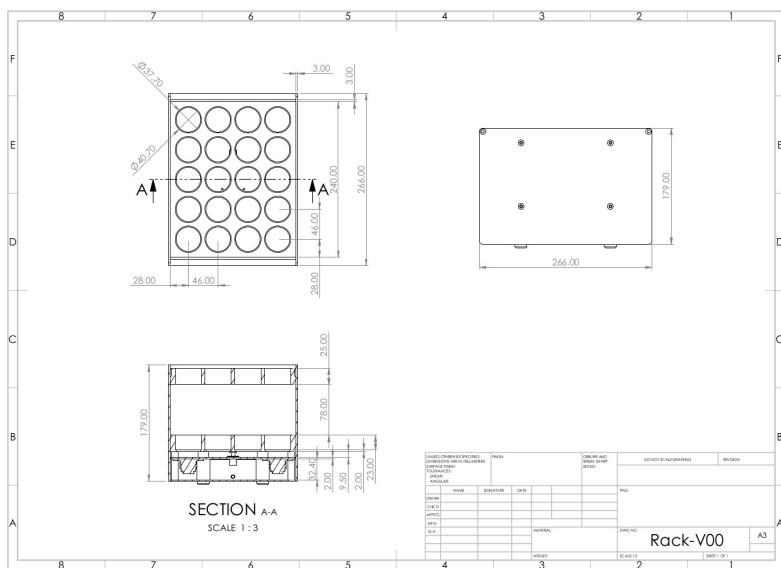


Figura 4: Plano rack.

Frente al escenario planteado, lo primero que consideramos como posible solución al problema fue introducir modificaciones en el diseño del sistema. En primer lugar, sugerimos evitar que el ventilador del resonador dirija directamente el flujo de aire caliente hacia los racks, ya que esto altera significativamente la distribución térmica en su interior.

En segundo lugar, proponemos una revisión en la selección de materiales. Actualmente, las paredes externas del rack están construidas en aluminio, seguidas de una delgada capa de aislante térmico, y en el interior, las estructuras están hechas de polietileno de alta densidad (HDPE), ver imagen 6. Esta configuración presenta un diseño térmicamente contraproducente: el aluminio, siendo un excelente conductor térmico, favorece la transmisión del calor desde el exterior, mientras que el HDPE, un mal conductor, dificulta la homogeneización de la temperatura interna.

Por tal motivo, recomendamos invertir esta lógica constructiva. Sugerimos utilizar materiales con alta conductividad térmica en el interior, como el aluminio o aleaciones livianas conductoras, para lograr una distribución de temperatura más homogénea en la zona de muestras, permitiendo un mejor control térmico. Por otro lado, en el exterior del rack, proponemos reemplazar el aluminio por materiales ligeros pero con baja conductividad térmica, como paneles sándwich con núcleos de espuma rígida de poliuretano (PUR), aerogel encapsulado o estructuras compuestas con revestimientos aislantes, que limiten las pérdidas o ganancias de calor del

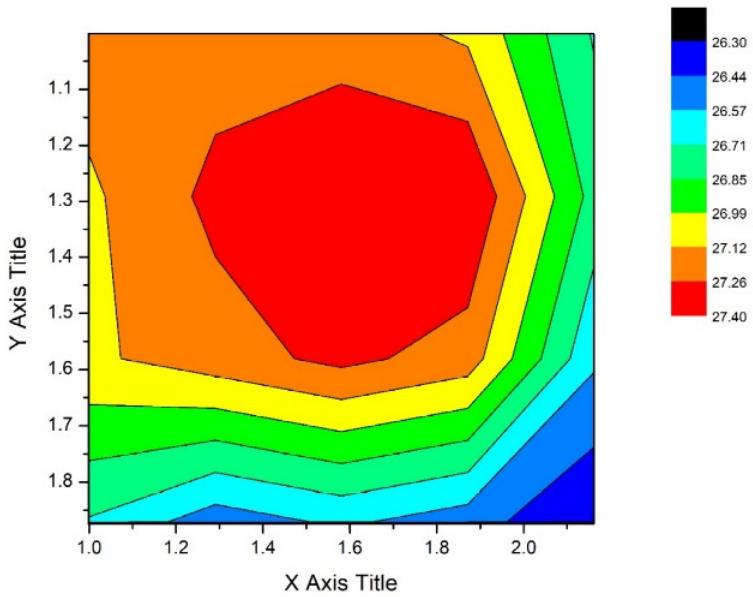


Figura 5: Distribución de temperatura tubo por tubo del rack B, ver 2(a).

entorno.

Este rediseño no sólo reduciría la influencia de condiciones externas —como el flujo de aire del ventilador o cambios en la temperatura ambiente del laboratorio—, sino que además permitiría que las fuentes de calor internas actúen de manera más eficiente y controlada, facilitando así una distribución térmica interna más estable y homogénea.

Finalmente, realizamos una recomendación respecto al diseño de las fuentes de calor. Actualmente, el sistema cuenta con una capa de aluminio colocada sobre las resistencias, cuyo propósito es distribuir el calor y lograr una temperatura más uniforme en la zona de muestras. Si bien esta estrategia ayuda parcialmente, sugerimos avanzar hacia un diseño de fuente de calor que, desde su geometría, promueva una distribución térmica más homogénea.

Una alternativa viable sería emplear una placa calefactora con resistencias dispuestas en un patrón optimizado, como un trazado en zigzag, meandro o espiral, de modo que el calor se entregue de forma más equitativa a lo largo de toda la región. Este tipo de configuración permite reducir los gradientes térmicos locales y mejorar la eficiencia del control.

Además, la elección del material sobre el que se coloca la resistencia y su contacto térmico con la región de interés también deben ser cuidadosamente diseñados para evitar puntos calientes o zonas frías. Esta propuesta, combinada con la selección adecuada de materiales conductores en el interior y aislantes en el exterior, contribuirá a mantener la temperatura de las muestras estable, uniforme y menos influenciada por factores externos.

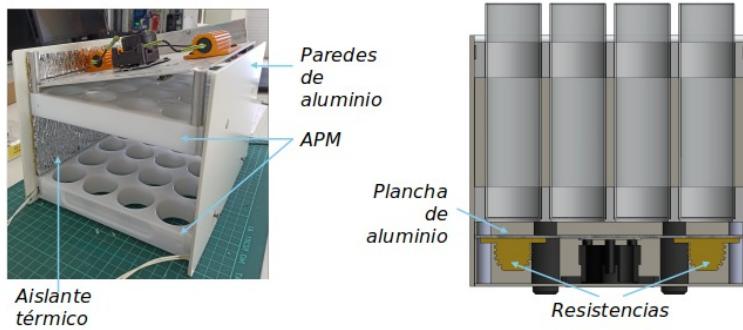


Figura 6: Distribución de materiales en el rack.

### 3.2. Enfoque cuantitativo

Dentro de este enfoque, orientado a implementar una solución basada en aprendizaje por refuerzo, nuestro objetivo fue desarrollar una simulación lo más realista posible de la evolución térmica dentro del robot. La idea es que el algoritmo aprenda a partir de escenarios representativos del entorno real, de modo que la política óptima aprendida, denotada por  $\pi$ , sea robusta y eficaz frente a las distintas condiciones operativas del sistema.

Un aspecto fundamental en este proceso es contar con simulaciones que incluyan distintas condiciones iniciales, lo que permite que el agente explore una amplia variedad de “trayectorias” posibles. Esto favorece la obtención de una política robusta, capaz de identificar acciones adecuadas incluso ante perturbaciones o configuraciones no previstas durante el entrenamiento.

Sin embargo, el desarrollo de una simulación con estas características presenta numerosos desafíos. El primero de ellos es el modelado detallado del sistema. Aunque la región de interés  $\Gamma$ , en la cual buscamos controlar la temperatura, se encuentra dentro de los racks, hemos evidenciado que el comportamiento térmico del robot completo influye significativamente en dicha región. Factores como la geometría del sistema, la interacción térmica entre los distintos racks, la presencia de fuentes de calor externas (como resistencias o el ventilador) y la exposición al ambiente deben ser tenidos en cuenta para lograr una representación fiel del fenómeno.

Para llevar a cabo una simulación por el método de elementos finitos, es necesario contar con un dominio modelado mediante mallas tipo TET (tetraédricas) o HEX (hexaédricas), en las cuales se puedan identificar claramente las distintas superficies para aplicar condiciones de borde diferenciadas y diferenciar los volúmenes que representen los distintos materiales. En nuestro caso, no contábamos con un modelo completo del robot, sino únicamente con el de un rack aislado proporcionado por la empresa. Pero no teníamos control directo sobre su geometría ni sobre cómo había sido discretizado, al no haber sido generada por nosotros, resultaba difícil adaptarla al flujo de trabajo de nuestro código FEM.

El segundo gran desafío fue el modelado físico de la termodinámica del sistema. La simulación completa requiere la resolución acoplada de la ecuación de calor, para describir la evolución térmica de los racks bajo diferentes condiciones iniciales, y de las ecuaciones de Navier-Stokes, para capturar la influencia del flujo convectivo inducido por el ventilador y otras corrientes de aire sobre la distribución de

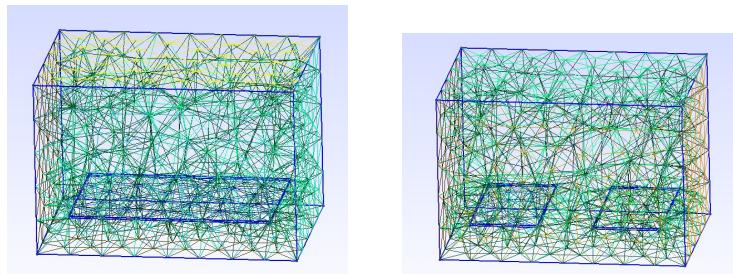
temperaturas.

Las complejidades mencionadas, tanto físicas como computacionales, representan una barrera significativa para alcanzar una simulación realista y eficiente en este enfoque cuantitativo de la solución. Dado que el proyecto se desarrolla en el marco de la VI TAMI, un taller con una duración de una semana, hemos optado por simplificar el modelado termodinámico del sistema. En paralelo, se decidió abordar el problema de control mediante técnicas de aprendizaje por refuerzos, tratando ambos aspectos de forma desacoplada. Esta estrategia permite avanzar en el desarrollo y validación del enfoque de control, mientras se pospone la incorporación de una simulación física más compleja para etapas posteriores del proyecto.

### 3.2.1. Simulación de la evolución de la temperatura

Para abordar la simulación termodinámica del sistema, se adoptó una estrategia incremental: comenzar con una versión altamente simplificada del problema e ir incorporando complejidades de manera progresiva. En esta primera etapa, se decidió simular un único rack de forma aislada, con una geometría simplificada y asumiendo que estaba compuesto por un único material homogéneo. Estas decisiones se fundamentan en las limitaciones de tiempo de una semana.

Para la construcción del dominio geométrico, se utilizaron dos grillados tridimensionales generados con la herramienta *Gmsh* [2], correspondiente a un volumen rectangular con las dimensiones aproximadas del rack. La diferencia entre ambos modelos radica en la configuración de las fuentes de calor: uno con una única placa emisora, simulando la propuesta de resistencia en zigzag mencionada en la sección anterior, y otro con dos placas emisoras separadas, emulando la disposición actual del sistema, ver imagen 7. Las grillas se encuentran disponibles en el repositorio.



Un rack con una fuente de calor.

Un rack con dos fuentes de calor.

Figura 7: Grillas

La resolución numérica se llevó a cabo en el lenguaje *Julia*, haciendo uso de la biblioteca *Gridap* [7], que permite implementar métodos de elementos finitos para ecuaciones en derivadas parciales. Para ello, se formuló la ecuación de calor en su forma débil, como se muestra en la ecuación 5, lo que da lugar a un problema diferencial-algebraico de primer orden en el tiempo, ecuación 6, resoluble mediante técnicas estándar de integración temporal.

Un aspecto central de esta simulación es la especificación de las condiciones de borde e iniciales, ya que definen tanto el estado térmico inicial del sistema como las

restricciones físicas que lo gobiernan. En esta instancia, se asumió que el contorno externo del rack está en contacto térmico con el ambiente, imponiéndose condiciones de Dirichlet a temperatura constante correspondiente a la del laboratorio. Las resistencias, por su parte, se modelaron como fuentes de flujo de calor constante mediante condiciones de Neumann, con una potencia fija que no depende del tiempo. Esta elección se ve respaldada por la libertad otorgada por la empresa para experimentar con diferentes configuraciones de potencia y distribución de las fuentes térmicas.

Cabe destacar que el código desarrollado permite modificar de forma sencilla tanto las condiciones de borde e iniciales como la geometría del dominio. Esto facilita su extensión a escenarios más complejos, como la incorporación del robot completo o el conjunto de racks, así como la modificación de la forma y distribución de las fuentes de calor, lo cual será clave en etapas futuras del proyecto.

### 3.2.2. Control dinámico mediante RL

En primer lugar, tenemos que plantear el problema de RL. Para esto, debemos definir los elementos del MDP. El primero es el conjunto de estados del sistema. En principio, este es un espacio continuo (conjunto denso) dado por la función de temperatura en cada punto  $T(\vec{x})$ ,  $\vec{x} \in \Omega \subset \mathbb{R}^3$ , donde  $\Omega$  es el dominio espacial de la simulación física. Por supuesto, en la implementación computacional utilizaremos una grilla que discretice el espacio continuo, pero los valores de temperatura posibles en cada punto seguirán siendo valores de punto flotante (i.e., números reales). Con esto, un estado  $s$  del sistema estará dado por un conjunto de valores de la temperatura para cada valor del dominio  $\{T(\vec{x}) | \vec{x} \in \Omega\}$ . Por supuesto, el conjunto de estados del sistema  $S$  será el conjunto de todos los estados  $s$  posibles.

Consideramos que el dispositivo tiene  $n$  actuadores independientes que controlan resistencias de distintas formas. Además, consideramos todas las resistencias conectadas al mismo actuador como una única resistencia, posiblemente no conexa (y usaremos los términos intercambiablemente). A su vez, consideramos que el dispositivo cuenta con  $m$  sensores de temperatura que reportan un promedio local en el entorno en el que están ubicados. De este modo, si al tiempo  $t$  la resistencia  $j \in \{1, \dots, n\}$  recibe la indicación de fijar la corriente al valor  $I^*$ , i.e.,  $I_j(t) = I^*$ , el sistema recibirá el correspondiente calor de Joule  $P_j(t) = I_j(t) \times \Delta V = I_j^2(t) \times R_j$ , donde  $R_j$  es la resistencia del  $j$ -ésimo resistor. Las corrientes varían de un mínimo de 0 a un valor máximo de  $I_{Max}$ .

El agente a considerar dispondrá de un conjunto continuo de acciones

$$\tilde{A} = \{a | a \in \mathbb{R}_{[0, I_{Max}]}^n\}$$

del cual escoge una acción particular para cada tiempo  $t$ , a aplicar durante un tiempo  $\Delta t$ . La duración de dicho tiempo  $\Delta t$  depende de la simulación física del entorno y en esta sección lo consideraremos como dado. Por supuesto, usar valores más pequeños para  $\Delta t$  nos permite un control más granular, pero valores demasiado chicos pueden no tener sentido ya que la potencia inyectada no llegaría a cambiar la temperatura del sistema. A esto hay que sumarle las limitaciones técnicas de lograr un control con tanta resolución temporal, y las fluctuaciones en las mediciones de temperatura

durante intervalos tan chicos por parte de los sensores. De este modo,  $\Delta t$  debería elegirse como el valor más grande que nos permita mantener el nivel de control deseado.

Una formulación alternativa se obtiene discretizando el espacio de acciones (i.e., discretizando los valores de la corriente). Si la resolución de la corriente es  $\Delta I$ , tendremos un máximo valor discreto  $N_{Max} = I_{Max}/\Delta I$  para la corriente a aplicar, por lo que el nuevo espacio de acciones  $A$  será

$$A = \{a | a \in \mathbb{Z}_{[0, N_{Max}]}^n\}$$

En ambas formulaciones, el resultado de una acción será aplicar una corriente dada en las resistencias de cada actuador. Estos valores de la corriente se usan luego como parámetros para la simulación física que nos da la evolución temporal en un  $\Delta t$ . Por simplicidad, en esta parte incluimos las corrientes en la simulación como fuentes, pero en una simulación más realista como la descripta en la sección anterior, debería usarse para fijar las condiciones de contorno de von Neumann. En cualquier caso, es importante notar que las acciones no cambian de manera directa (o discontinua) los estados, sino que los modifican a través del término de fuente de la ecuación de calor, o del flujo impuesto en las superficies con condición de contorno de von Neumann.

Con el arreglo actual, donde las resistencias son de 220 ohm y están conectadas en serie, al alimentarlas con 220V, la potencia máxima sería de 110W (55W en cada resistencia). El control de potencia se hace alimentandolas con un PWM (modulación por ancho de pulso) que va de 0 a 255, siendo 0 una potencia de 0w y 255 una potencia de 110W , lo que nos da una resolución de 0.43W. Esto implica que en este arreglo  $P_{Max} = I_{Max}^2 R = 110W$  y  $N_{Max} = 255$ . Dado el funcionamiento actual del robot, optaremos por esta formulación discretizada del espacio de acciones.

Lo siguiente a definir en nuestro problema es la función recompensa a usar. Sea  $\Gamma \subset \Omega$  la sección del espacio físico que comprende las muestras, la recompensa que daremos al agente por su acción a un tiempo dado será el opuesto de la suma de los desvíos de temperatura en cada punto, respecto a la temperatura objetivo, i.e.,

$$\tilde{R}(s_{t-1}, a_{t-1}) = R(s_t) = \int_{\Gamma} [T(\vec{x}, t) - T_{target}]^2 d^3x$$

donde  $T_{target}$  es la temperatura a la que deseamos mantener las muestras. Recorremos que el efecto de una acción sobre un estado se da a través de las condiciones de contorno de von Neumann para la ecuación de calor (o en este caso simplificado, a través del término de fuente de la ecuación).

Respecto a la distribución de probabilidad de transición, usaremos la evolución determinística de elementos finitos como regla de transición de un tiempo al siguiente. Cabe destacar que para obtener una política robusta ante cambios imprevistos del entorno, debería exponerse al sistema a condiciones con cierta aleatoriedad. Esto podría ser ruido, pero lo recomendable sería incluir también cambios macroscópicos como la entrada de un frente frío de aire que cambie bruscamente las temperaturas por convección.

Habiendo planteado el problema, la metodología de este enfoque consistirá en resolver dicho problema mediante PPO. Como nuestro objetivo es ilustrar el funcionamiento de una solución de este tipo, corremos un modelo simplificado de acciones discretas binarias. Usamos la implementación de PPO de Stable-Baselines3, y usamos un modelo simplificado del entorno que consiste en resolver la ecuación de calor con 3 fuentes gaussianas y 3 sensores desplazados diagonalmente respecto a las fuentes. Estas simplificaciones son para evitar el costo computacional de la simulación completa, pero la extensión es directa. El código puede verse en el repositorio.

## 4. Resultados

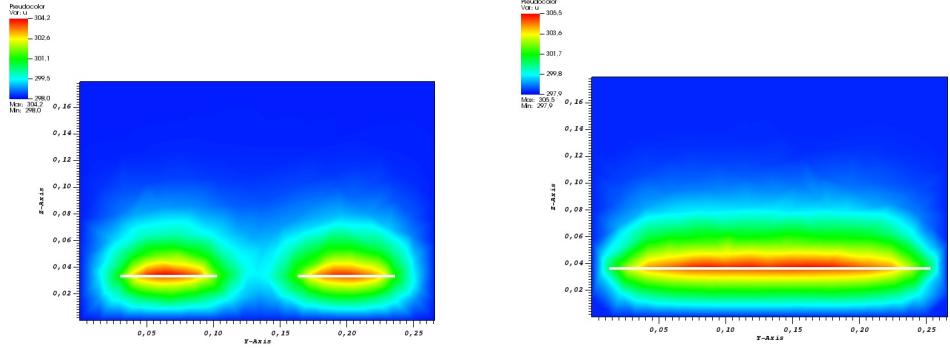
El abordaje inicial debería constar en una optimización del diseño del equipo teniendo en cuenta la eficiencia térmica del mismo. Las recomendaciones generales más importantes son utilizar materiales aislantes en el exterior y materiales conductores en el interior. Esto hará que la temperatura se homogeneize rápidamente en el interior, a la vez que reducimos la pérdida de calor en las superficies exteriores. En la misma línea, es importante evitar las fuentes de asimetrías siempre que sea posible. Un ejemplo de esto es la presencia de un frente de aire caliente dado por el ventilador que apunta al rack C, pero también podría serlo un posible frente frío dado por la desafortunada ubicación de un aire acondicionado en dirección al equipo.

Respecto a la optimización estática, el código desarrollado para la simulación térmica de un rack ofrece resultados consistentes con lo esperado. En la imagen 8 se muestra un corte bidimensional representativo de la distribución de temperatura para un instante de tiempo dado, evidenciando la evolución térmica dentro del dominio simulado, dadas las condiciones iniciales y de borde mencionadas en la sección anterior.

Una de las principales fortalezas del código es su flexibilidad: permite modificar con facilidad la disposición y geometría de las fuentes de calor, el diseño de las grillas, así como las condiciones de borde e iniciales. Este control total sobre el dominio, los parámetros físicos y la salida de la simulación es fundamental para su integración con el enfoque de aprendizaje por refuerzos. En este marco, el algoritmo debe ser capaz de observar el estado del sistema y predecir su evolución térmica a partir de diferentes acciones, con el fin de guiarlo hacia el estado deseado. Más allá de esto, el simple hecho de tener una simulación térmica permite una evaluación cuantitativa de del impacto de cambios tanto en los materiales como en la disposición espacial de los componentes del equipo, permitiendo guiar sucesivas mejoras en su diseño.

Si bien herramientas como COMSOL [4] permiten realizar simulaciones físicas de manera robusta, en nuestro caso es imprescindible tener control absoluto sobre el modelado. Esto incluye la definición precisa de las superficies donde se imponen las condiciones de contorno, la segmentación de volúmenes para asignar propiedades térmicas diferenciadas, la posibilidad de proponer y evaluar modificaciones de diseño, y la manipulación directa de la salida del campo de temperatura. Todo esto resulta esencial para acoplar la simulación física con el algoritmo de control basado en aprendizaje automático.

Respecto al enfoque de control dinámico, el agente entrenado con PPO aprende a hacer oscilar las señales para intentar mantener la temperatura lo más homogénea



Rack con una fuente de calor.

Rack con dos fuentes de calor.

Figura 8: Distribuciones de temperatura en un corte dos dimensional del rack. Ver GIF de evolución en el repositorio.

possible dentro de la región de interés (ver Fig. 9 y GIF en el repositorio). Esta estrategia es esperable, ya que es una forma muy intuitiva de disminuir la diferencia de temperatura entre la fuente y el borde del dominio. La ventaja respecto a un protocolo determinista de corriente oscilante es que esta policy sería en principio capaz de reaccionar a un cambio en las temperaturas debido a causas externas (en lugar de sus propias acciones), como reflejado por las mediciones de los sensores. Repitiendo este procedimiento con el entorno físico más realista podríamos obtener una política robusta que controle los desvíos de temperatura en tiempo real. Una opción más simple que RL podría ser fijar una señal de corriente oscilante cuya amplitud se hace crecer o decrecer en función de las lecturas de los sensores. Sin embargo, si se requiriera un control muy fino del equipo, RL presenta un enfoque más escalable. Esto es así ya que los agentes que controlan los actuadores de los distintos racks pueden ser el mismo o ser agentes cooperantes. Ambas opciones permiten a su vez compartir información sobre el estado total del sistema. Dado que la temperatura en un rack no es independiente de la temperatura en los demás, esta información compartida permite estimar con mayor fidelidad el estado completo del sistema, y por lo tanto mejorar la toma de decisiones. No sólo esto, sino que si la influencia de la temperatura de un rack en otro fuera lo suficientemente grande, uno podría tener en cuenta el efecto de prender la resistencia de un rack sobre los demás a la hora de decidir el amperaje a aplicar. Esta estrategia de RL multi-agente ha sido ya evaluada en la literatura, notablemente en el caso de la cimatisación de edificios, donde se estima que gran parte del consumo eléctrico se debe a la calefacción, ventilación y refrigeración (HVAC - Heating, Ventilation and Air Conditioning) de los ambientes. Por ejemplo, Liang Yu et al. [8] reportan que entre el 40 % y el 50 % del consumo eléctrico en edificios comerciales se debe a HVAC. Motivados por esto, plantean el problema como un juego de Markov y proponen un algoritmo de control de HVAC basado en Deep RL con múltiples agente, que logra efectividad, robustez y escalabilidad en entornos simulados en base a datos de edificios reales. Por supuesto, dada la complejidad de este enfoque hay una amplia gama de desafíos a enfrentar para lograr un uso exitoso y seguro de los mismos. Como se señala en la revisión sistemática realizada por Khalil Al Sayed et al. [1], el mayor desafío

consiste en proveer entornos con la diversidad suficiente para reflejar condiciones reales durante el entrenamiento, y el coste computacional que esto conlleva. Sin embargo, los autores identifican el meta-RL como una posible manera de superar estas limitaciones y sugieren que trabajos futuros en el área deberían investigar en esta dirección. Todo esto nos vuelve optimistas respecto a un enfoque basado en RL para este tipo de problemas, si los beneficios ameritan la inversión.

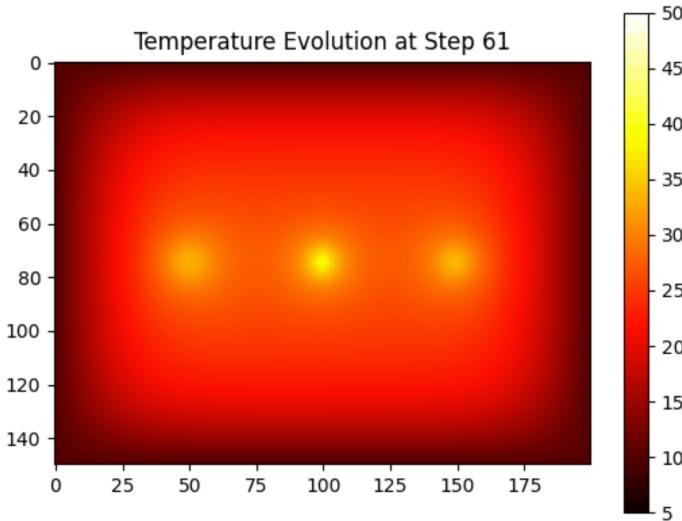


Figura 9: *Snapshot de la simulación de la interacción del agente con el entorno.* Vea el GIF en el repositorio.

## 5. Conclusión y Trabajo Futuro

A partir del registro de temperaturas provisto por la empresa en uno de los racks (imagen 5), se observa que las temperaturas oscilan en torno a los 27°C deseados, pero con un margen de error de aproximadamente  $\pm 1^\circ\text{C}$ . Considerando esto, creemos que los cambios en los materiales y su distribución, tal como se propusieron en este trabajo, podrían contribuir significativamente a alcanzar una distribución térmica más homogénea en la zona de muestras con la precisión deseada.

Por otro lado, hemos demostrado que el enfoque cuantitativo propuesto —basado en simulaciones térmicas y aprendizaje por refuerzos— es factible. Sin embargo, debido a las limitaciones de tiempo y recursos en el marco del VI TAMI, la implementación realizada fue altamente simplificada y pretende ser simplemente una guía a la hora de construir la solución real. Aún resta analizar con mayor profundidad el costo computacional y la complejidad del sistema completo para evaluar su viabilidad a gran escala.

Como línea de trabajo futuro, proponemos a la empresa implementar las modificaciones de diseño aquí sugeridas y evaluar experimentalmente su impacto en la distribución térmica dentro de los racks. En caso de que dichas modificaciones no logren el control deseado de la temperatura en la región de interés, se sugiere

avanzar en una modelación física más detallada del sistema, incorporando todas las complejidades mencionadas a lo largo del informe. Disponer de simulaciones realistas del entorno físico permitirá explorar cambios de diseño evaluando el impacto cuantitativo en las temperaturas. Si esto no fuese suficiente o si se deseara un control robusto y en tiempo real del sistema, el disponer de una simulación realista del sistema físico permitiría entrenar un algoritmo de aprendizaje por refuerzos capaz de identificar estrategias de control dinámico que garanticen una temperatura estable y homogénea en la zona de muestras. Las estrategias para llevar esto a cabo son variadas y diversas, y la selección de la misma debería realizarse incorporando los aprendizajes de áreas más maduras del control de temperatura como la climatización de edificios, que presenta un problema análogo en otra escala.

## Referencias

- [1] Khalil Al Sayed, Abhinandana Boodi, Roozbeh Sadeghian Broujeny, and Karim Beddiar. Reinforcement learning for hvac control in intelligent buildings: A technical and conceptual review. *Journal of Building Engineering*, 95:110085, 2024.
- [2] Christophe Geuzaine and Jean-François Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79:1309 – 1331, 09 2009.
- [3] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Science/Engineering/Math, 1997.
- [4] COMSOL Multiphysics. Multiphysics software, 2025.
- [5] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms, 2017.
- [6] Spinlock S.R.L. Spinlock. magnetic resonance solutions, 2025.
- [7] Francesc Verdugo and Santiago Badia. The software design of gridap: A finite element package based on the julia jit compiler. *Computer Physics Communications*, 276:108341, 2022.
- [8] Liang Yu, Yi Sun, Zhanbo Xu, Chao Shen, Dong Yue, Tao Jiang, and Xiaohong Guan. Multi-agent deep reinforcement learning for hvac control in commercial buildings. *IEEE Transactions on Smart Grid*, PP:1–1, 07 2020.