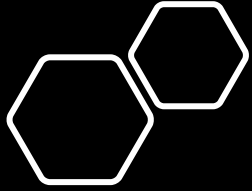


# Presentación final trabajo de título.

Diseño de sistemas de control de climatización en edificios basados en aprendizaje reforzado.

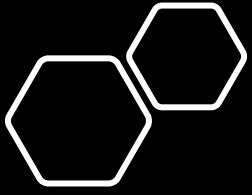


- Profesor del curso:  
Francisco Rivera
- Profesor Guía:  
Diego Muñoz C.
- Integrante comisión:  
Óscar Cartagena V.
- Alumno:  
Ricardo López D.



# Agenda

- Objetivos.
- Problema a resolver
- Aprendizaje reforzado.
- Configuración de los algoritmos.
- Estudios por simulación.
- Trabajo futuro.



# Objetivos

- **Objetivo general:**
  - Controlar mediante aprendizaje reforzado un sistema de climatización de un edificio, mejorando la eficiencia energética sin perjudicar a las personas en su interior.
- **Objetivos específicos**
  - Resolver distintas variantes del problema mediante aprendizaje reforzado con algoritmo tipo model free.
  - Analizar el rendimiento de diversos algoritmos de aprendizaje reforzado.
  - Utilizar Matlab para el diseño del sistema de control generando un principio de toolbox.

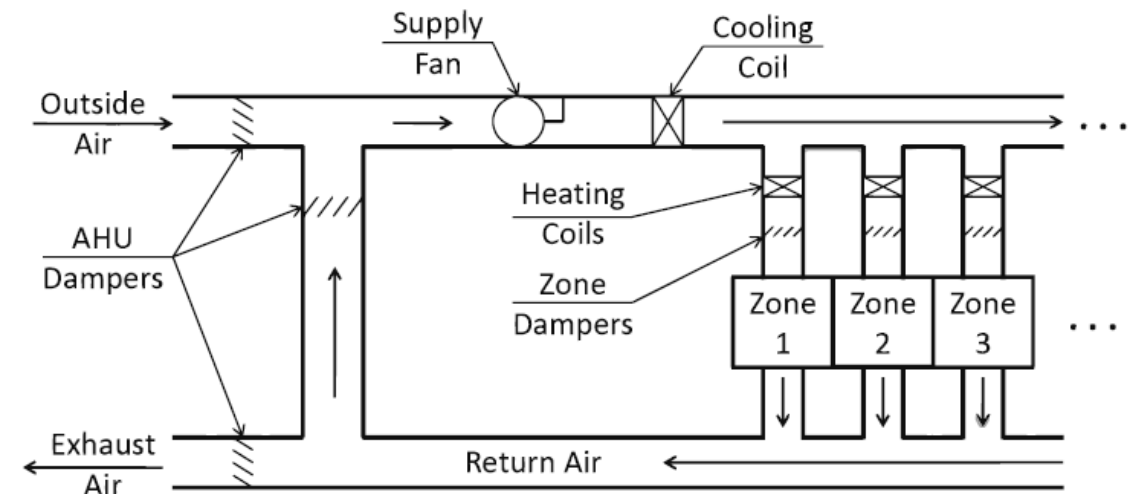
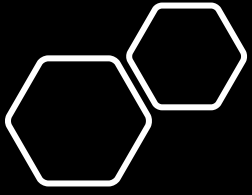


Fig.1: Diagrama HVAC[1]



# Problema a resolver

- El sistema a controlar consta de 4 variables que permiten el control de las restricciones impuestas.
- Rangos restricciones:
  - Temperatura: 22 a 25 [°C].
  - Concentración CO<sub>2</sub>: menor a 1200 [ppm].
  - Humedad relativa: 35 a 65 [%].

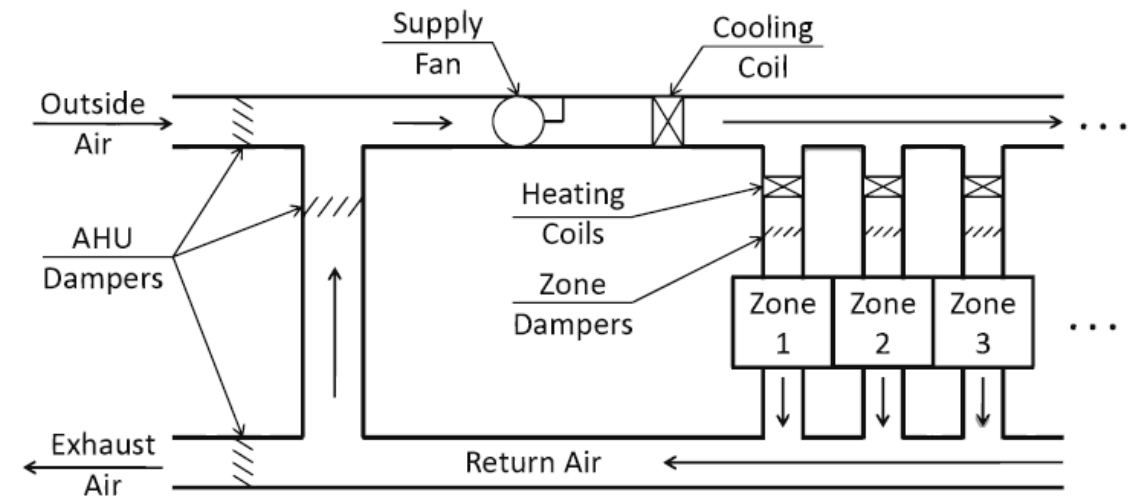
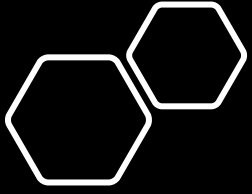


Fig.1: Diagrama HVAC[1]



# Problema a resolver

- Dinámica de temperatura y CO2:

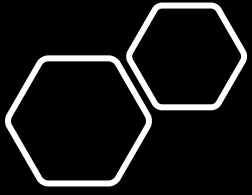
$$C_1^j \dot{T}_1^j = c_p \dot{m}_s^j (T_s^j - T_1^j) + \frac{T_2^j - T_1^j}{R^j} + \frac{T_a - T_1^j}{R_a^j} + P_d^j + \sum_{i \in N^j} \frac{T_1^i - T_1^j}{R_{ij}}$$

$$C_2^j T_2^j = \frac{T_1^j - T_2^j}{R_{12}^j}$$

$$T_s^j = \delta \frac{\sum_{i \in R} \dot{m}_s^i T_1^i}{\sum_{i \in R} \dot{m}_s^i} + (1 - \delta) T_a - \Delta T_c + \Delta T_h^j$$

$$\rho V^j \dot{C}_{CO_2}^j = \dot{m}_s^j (C_{CO_2,s} - C_{CO_2}^j) + g_{CO_2} N_{people}$$

$$C_{CO_2,s} = \delta \frac{\sum_{i \in R} \dot{m}_s^i C_{CO_2}^i}{\sum_{i \in R} \dot{m}_s^i} + (1 - \delta) C_{CO_2,a}$$



# Problema a resolver

- Dinámica de Humedad:

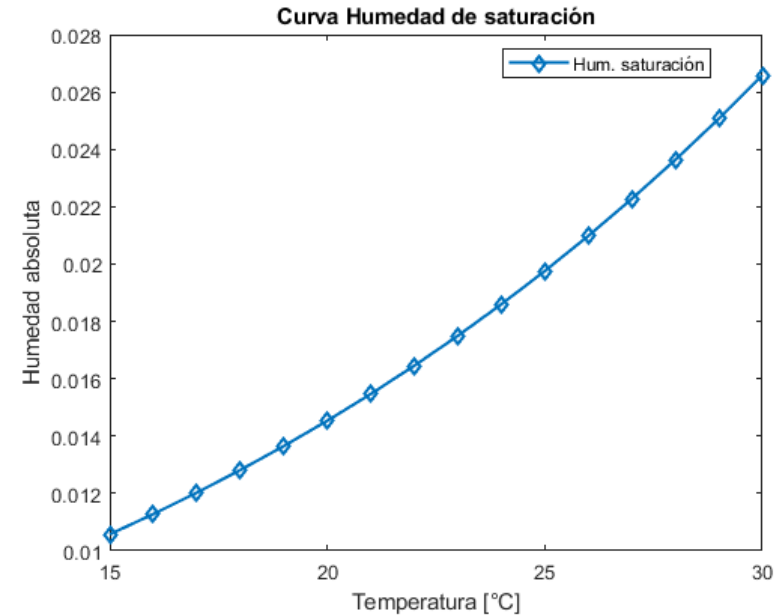
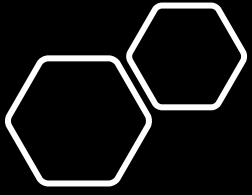


Fig.2: Humedad de saturación[2]

$$\rho V^j \dot{H}^j = \dot{m}_s^j (H_s - H^j) + g_H N_{people}$$

$$H_m = \delta \frac{\sum_{i \in R} \dot{m}_s^j H^i}{\sum_{i \in R} \dot{m}_s^j} + (1 - \delta) H_a$$

$$H_s = \min(H_m, H_{sat, T_s})$$



# Aprendizaje reforzado

- Área de ML y DataM.
- Hace un agente inteligente que toma acciones para maximizar una cierta recompensa promedio.
- Aprendizaje en base a experiencias.
- Se utilizan los algoritmos DDPG, TD3 y PPO.

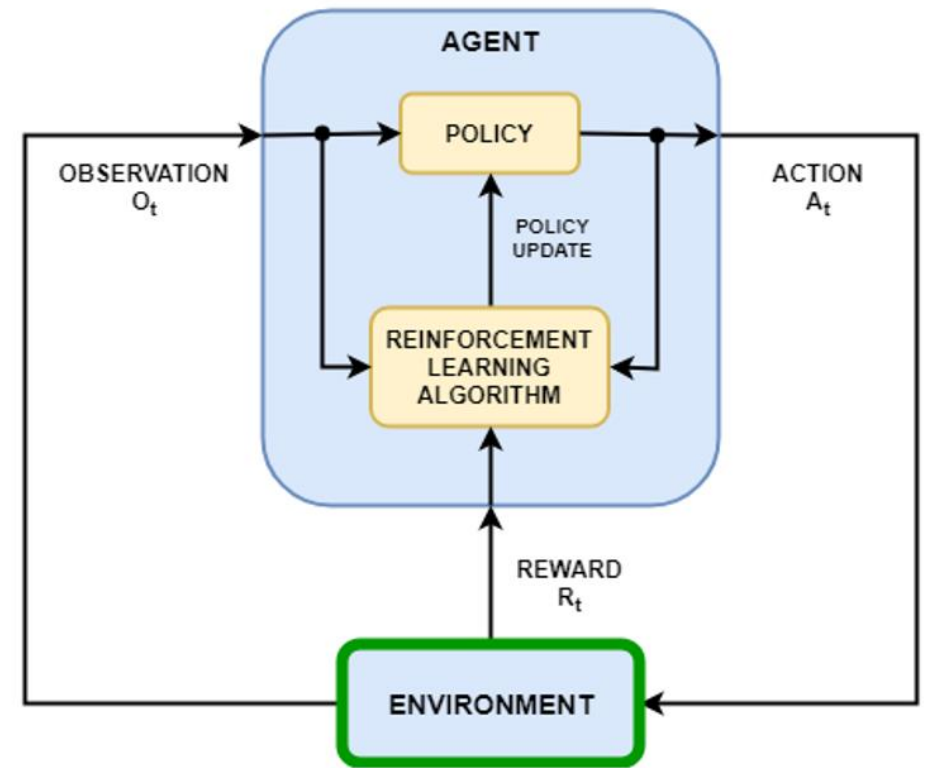
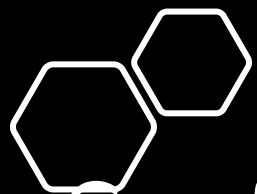
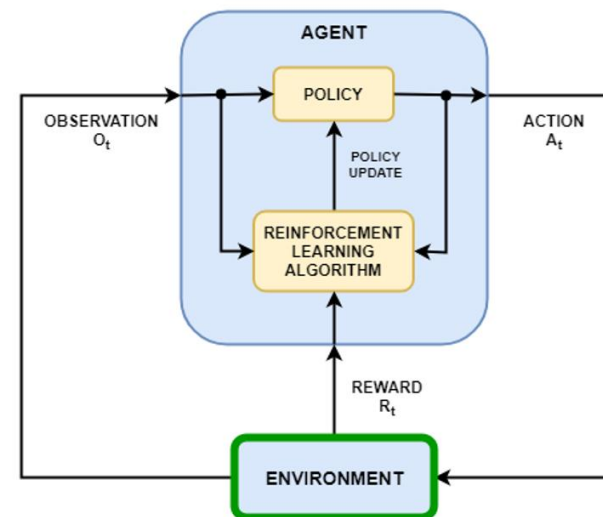


Fig.3: Diagrama RL-Matlab[2]



# Configuración general de los algoritmos

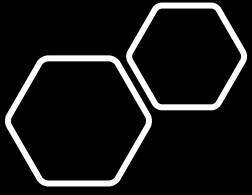
- Tiempos de simulación de 6 y 48 horas.
- Tiempo de muestreo de 10 min.
- Redes neuronales acorde a la magnitud del problema.
- Observaciones son los estados y perturbaciones medidas.
- Función de recompensa dada por:



$$R_t = E \cdot \beta + \begin{cases} 0, & \text{if Cumple las restricciones.} \\ -\Delta t, & \text{if No cumple las restricciones} \end{cases}$$

$$E = \frac{c_p}{\eta_c} \cdot m_{s,k} \Delta T_k + \frac{c_p}{\eta_h} \cdot m_{s,k} \Delta T_h + \kappa_f \left( \sum_i m_{s,k}^i \right)^2$$





# Estudios por simulación

- Se abordan distintos casos de simulación.
- Se escala en cantidad de variables de control, de 1 a 4.
- Se escala en cantidad de restricciones, partiendo por temperatura.

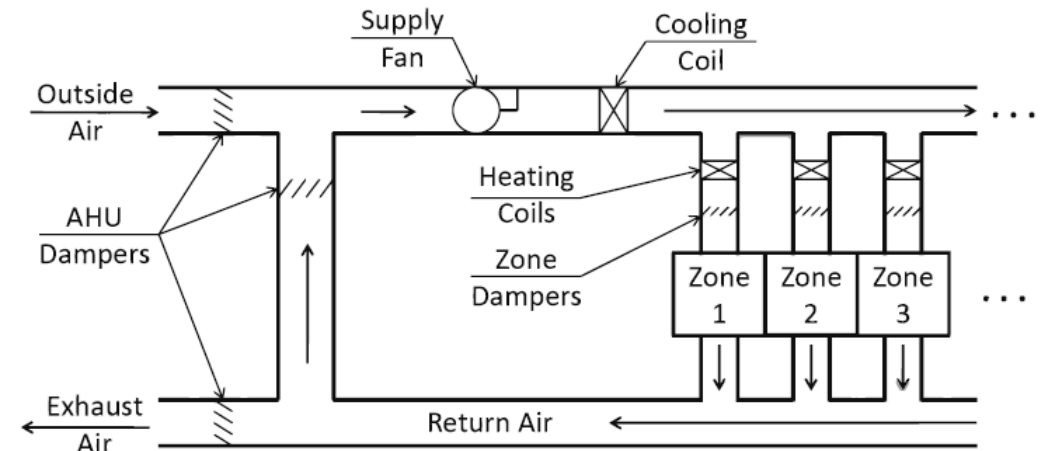
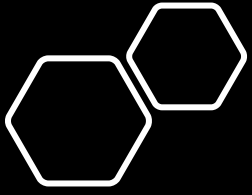


Fig.1: Diagrama HVAC[1]



# Estudios por simulación

- Se aborda el problema en el caso más simple posible.
- Se utiliza 1 zona, controlando solo temperatura.
- Se consideran acciones control solo de enfriamiento.
- Se tiene 1 variable como acción de control, y 3 acciones constantes.

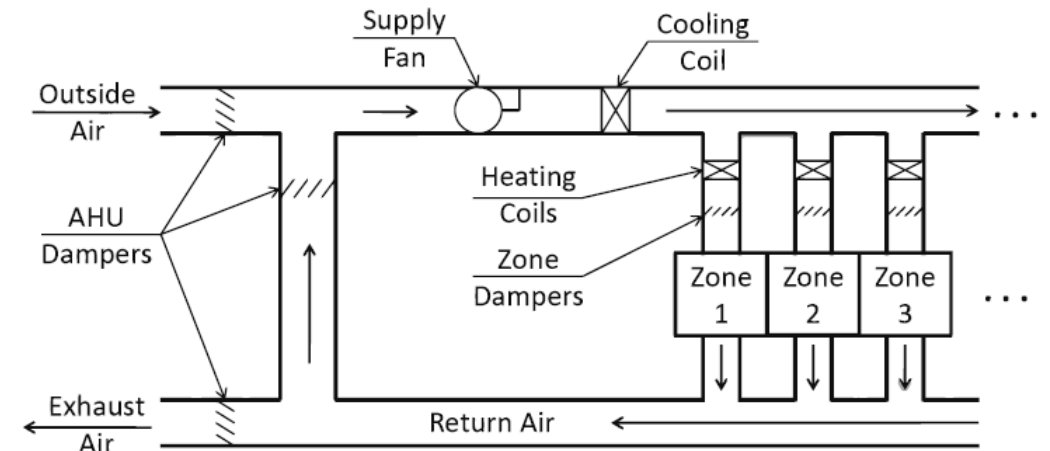
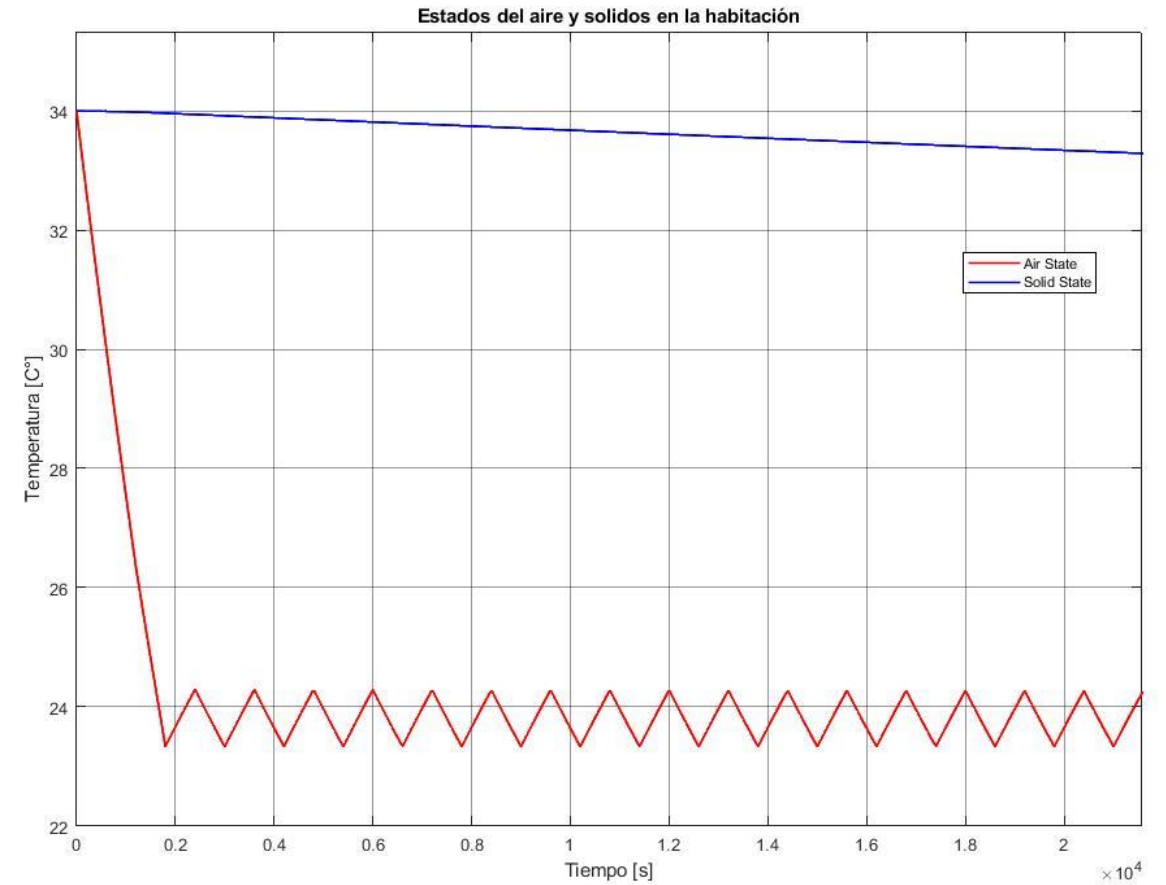
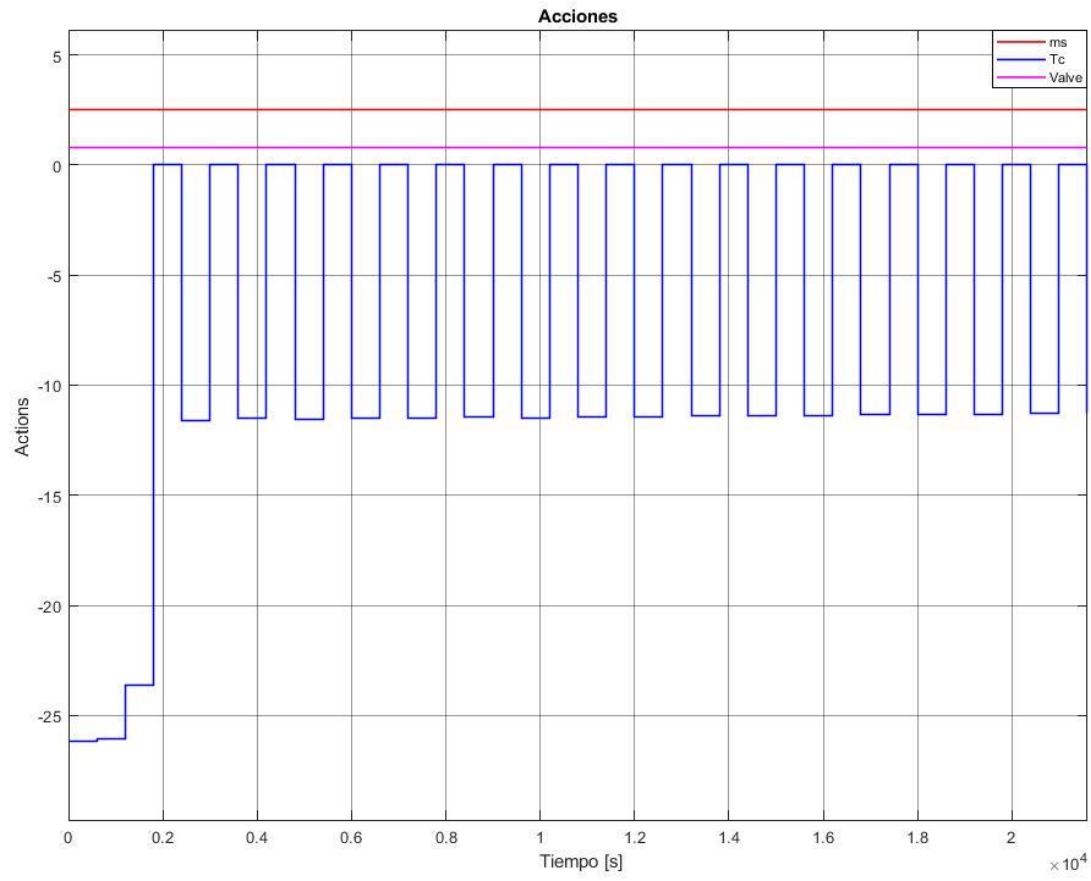


Fig.1: Diagrama HVAC[1]

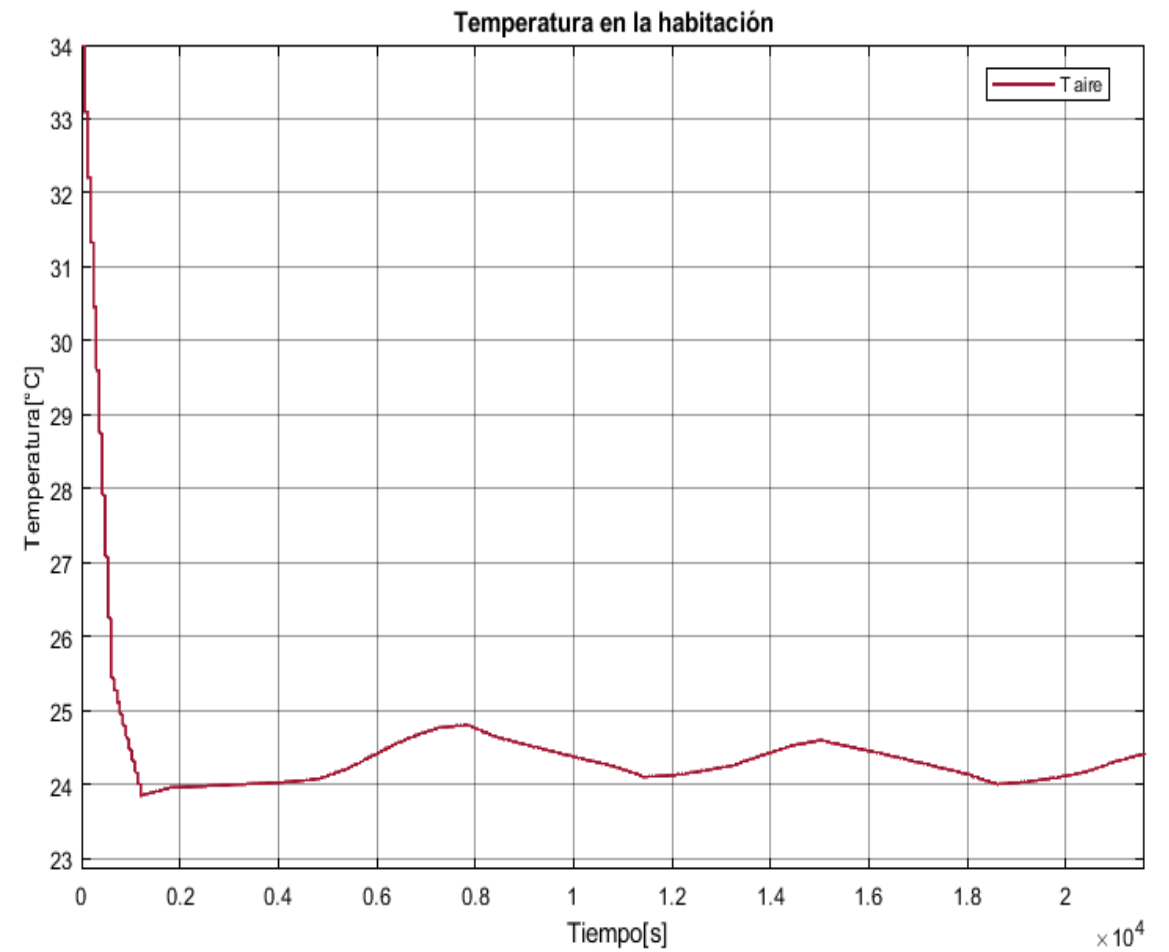
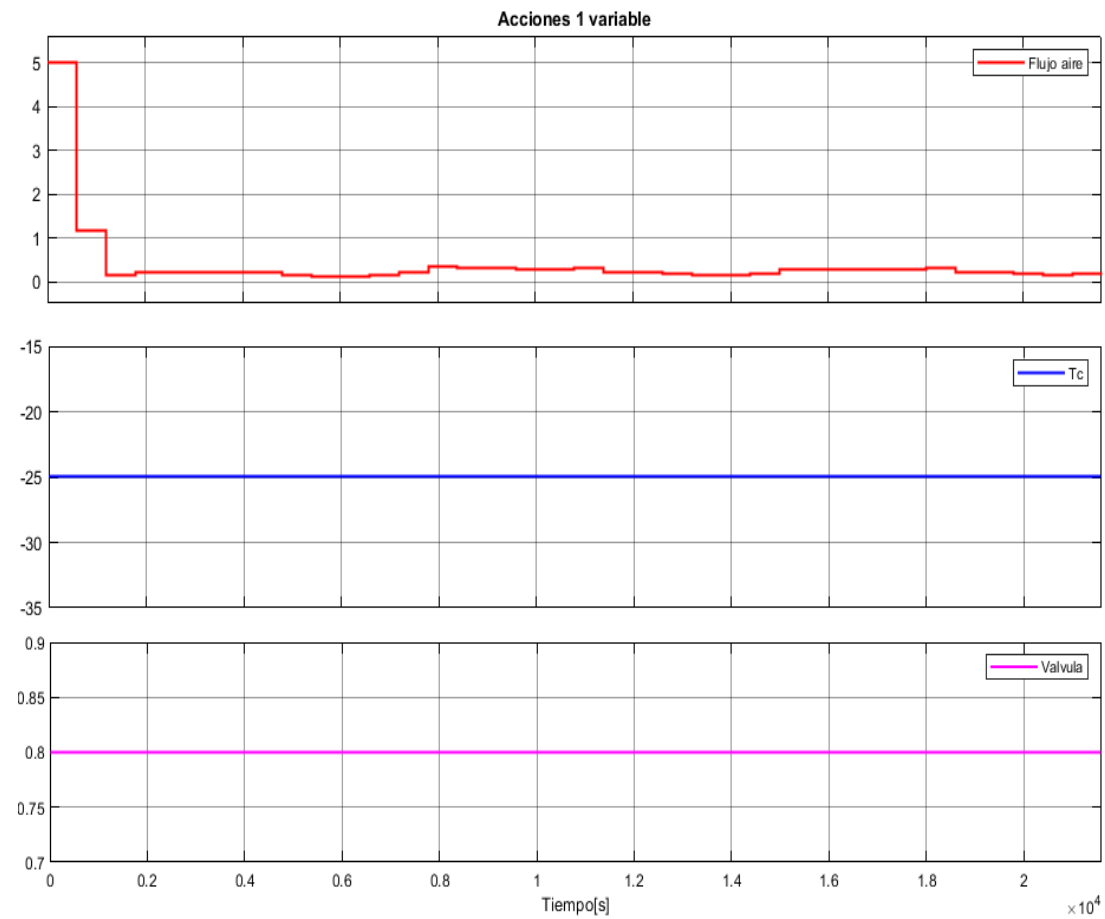
# Resultados

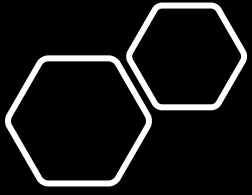


# Resultados



# Resultados





# 2 variables de control

- Se utiliza 1 zona, controlando solo temperatura.
- Se consideran acciones control solo de enfriamiento.
- Se tienen 2 variables como acciones de control, y 1 acción constante.
- Se prueba el sistema con consumo energético.

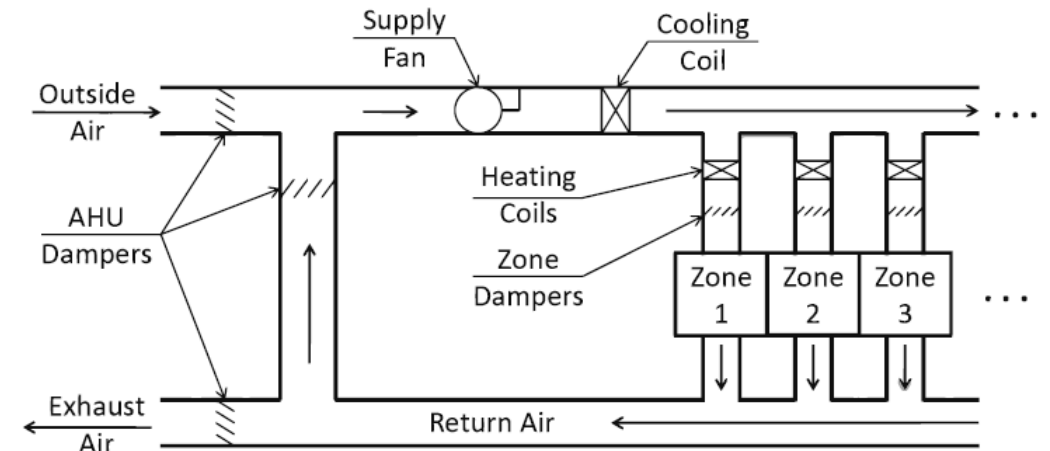
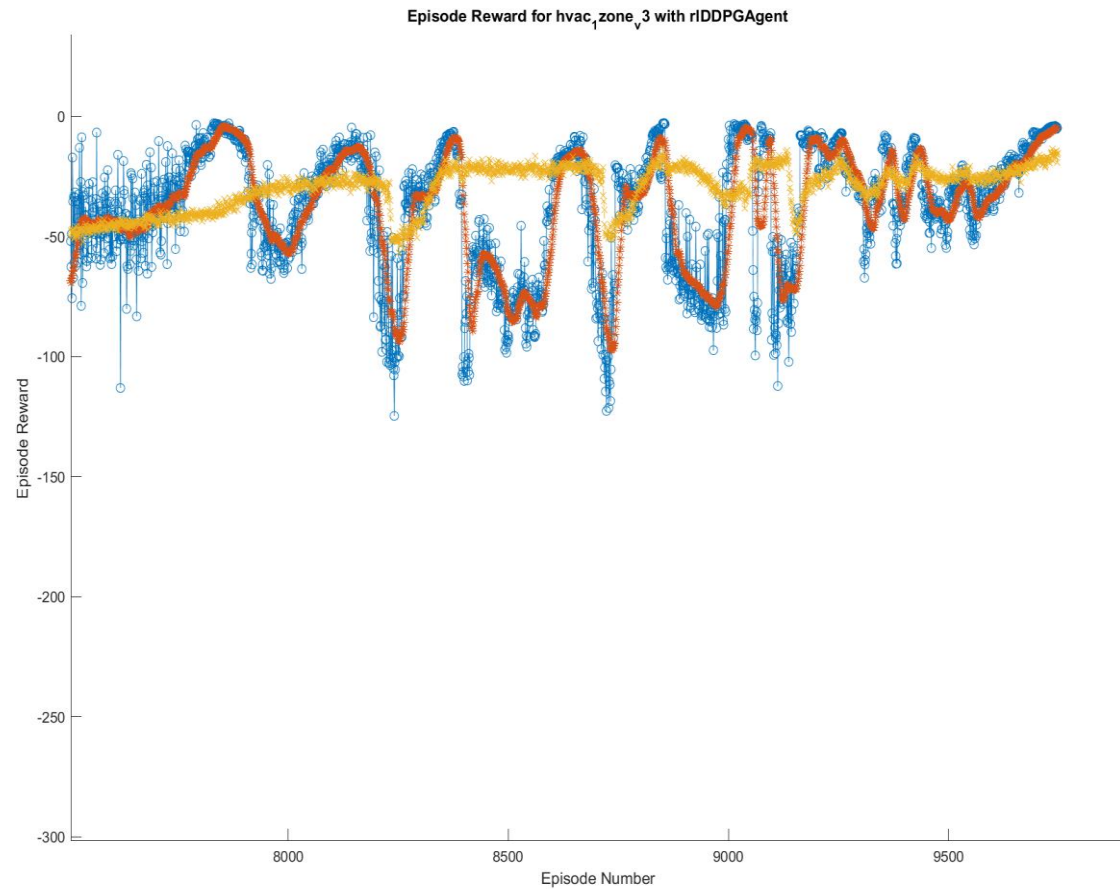
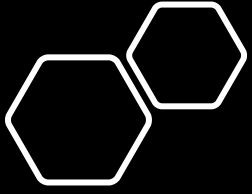


Fig.1: Diagrama HVAC[1]

## 2 variables de control



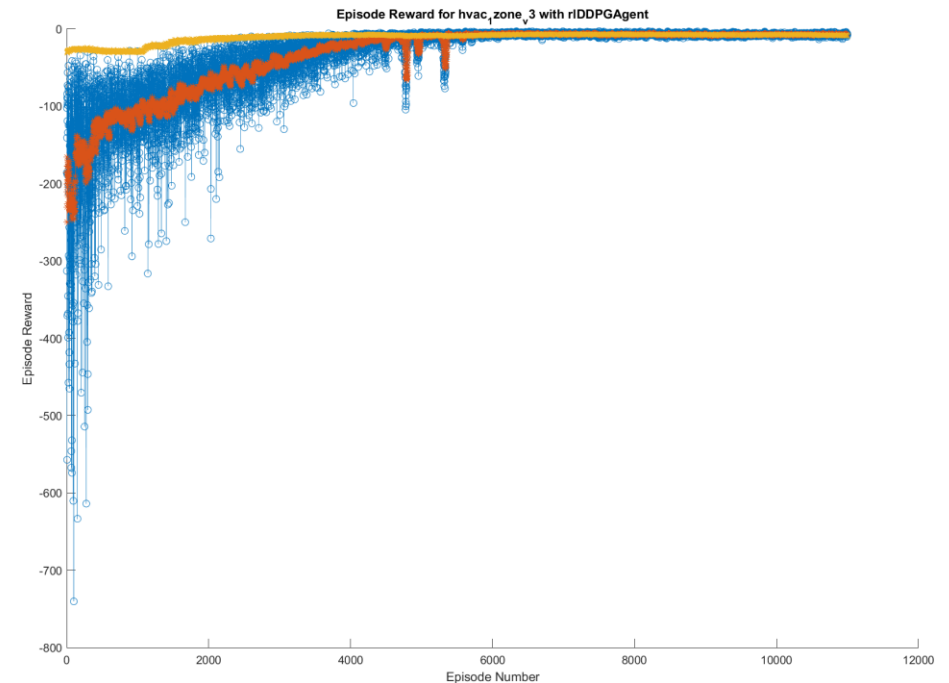


# Cambios en el algoritmo

- Se actualiza la función de recompensa
- Se mejora la red neuronal:
  - Función de activación sigmoid de salida.
  - Escalamiento de las variables.

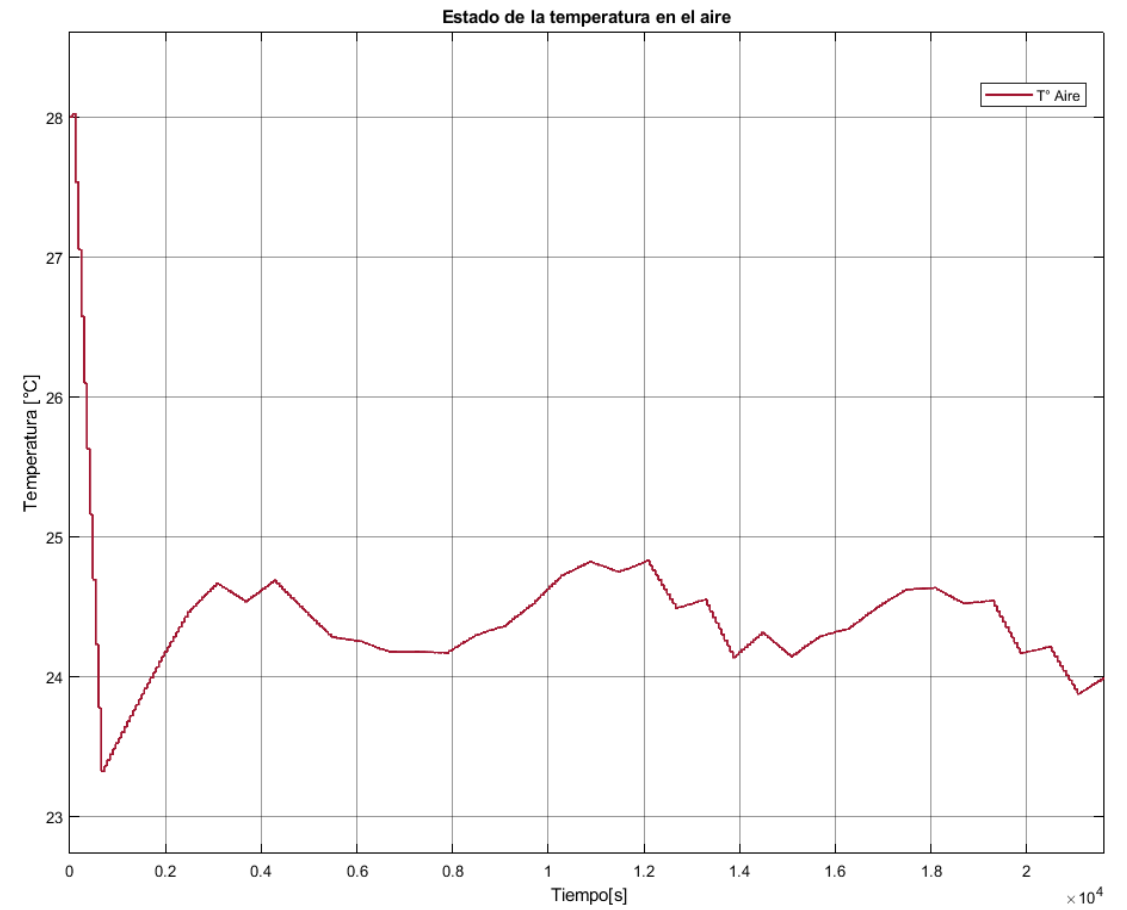
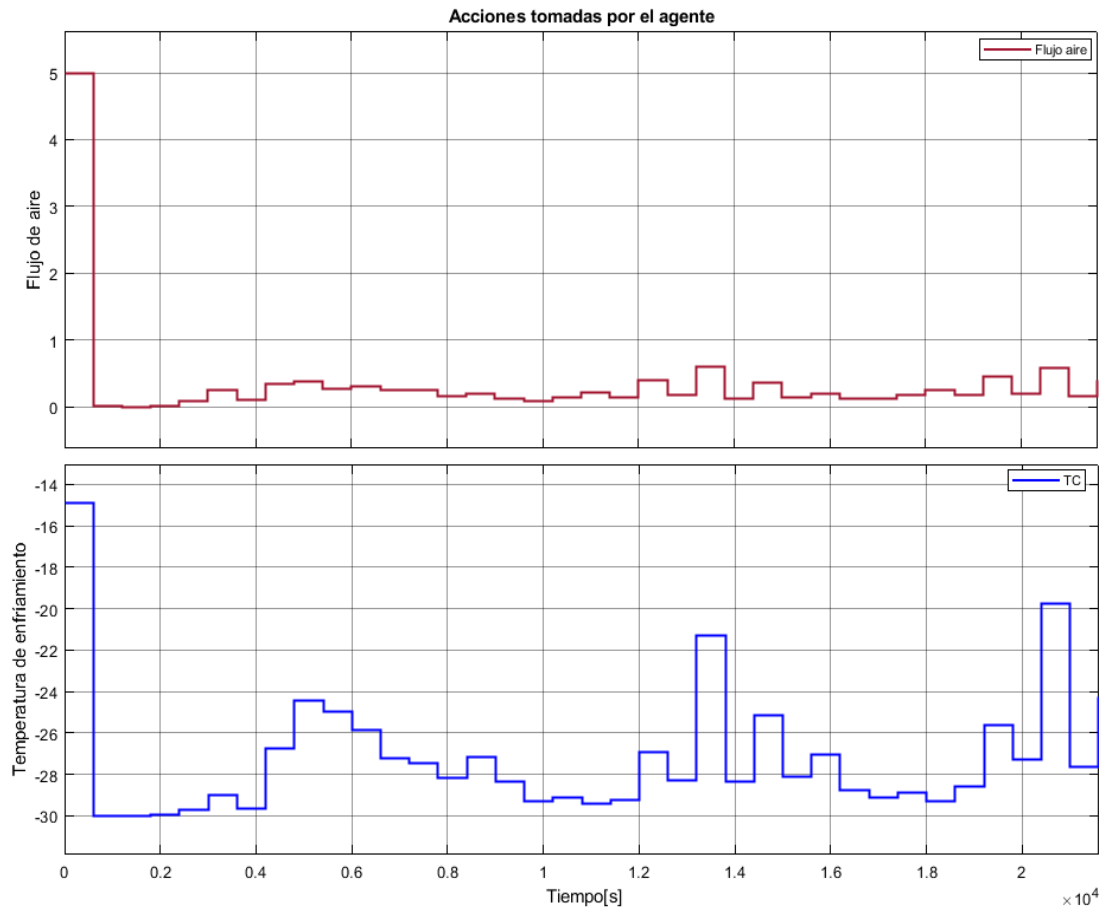
$$w_t = \min(|t - t_{\min}|, |t - t_{\max}|)$$

$$r(t, \alpha) = -0.1 \cdot E + \gamma \cdot (w_t + w_t^2)$$

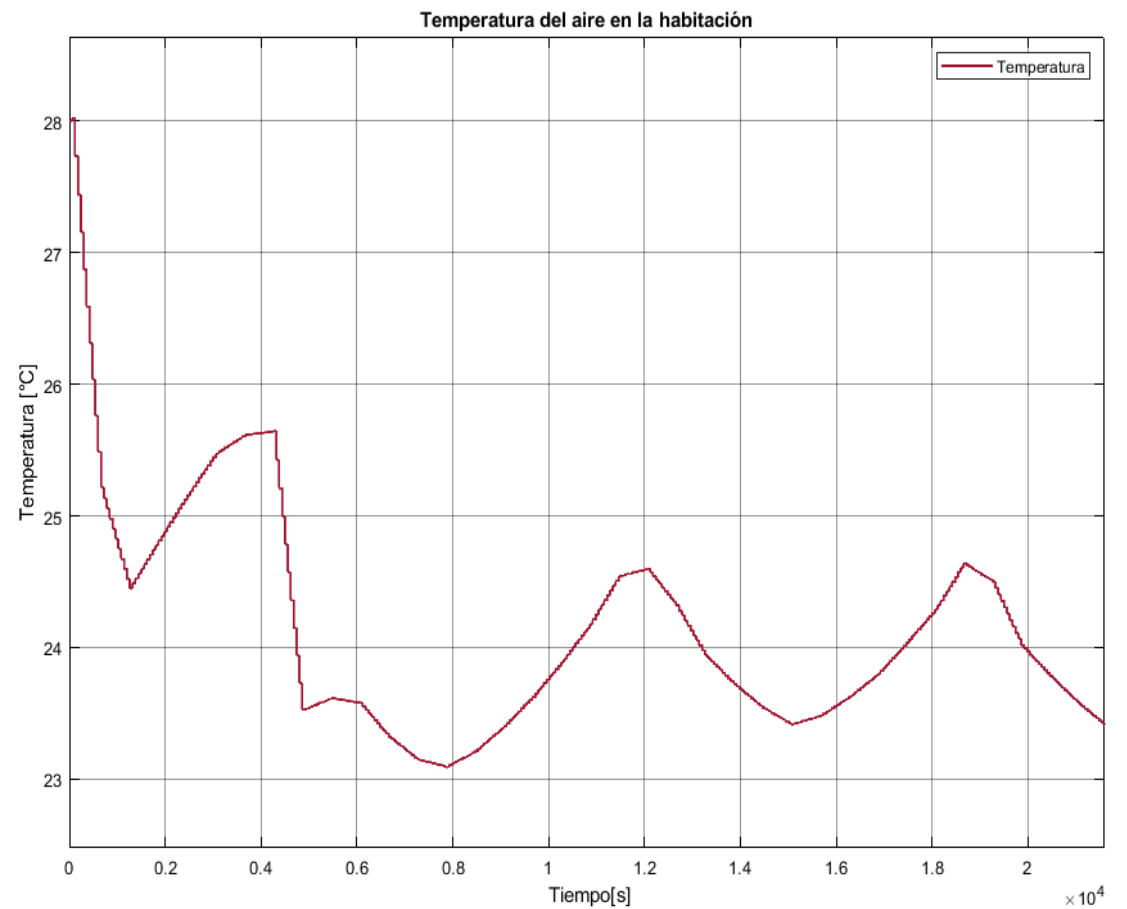
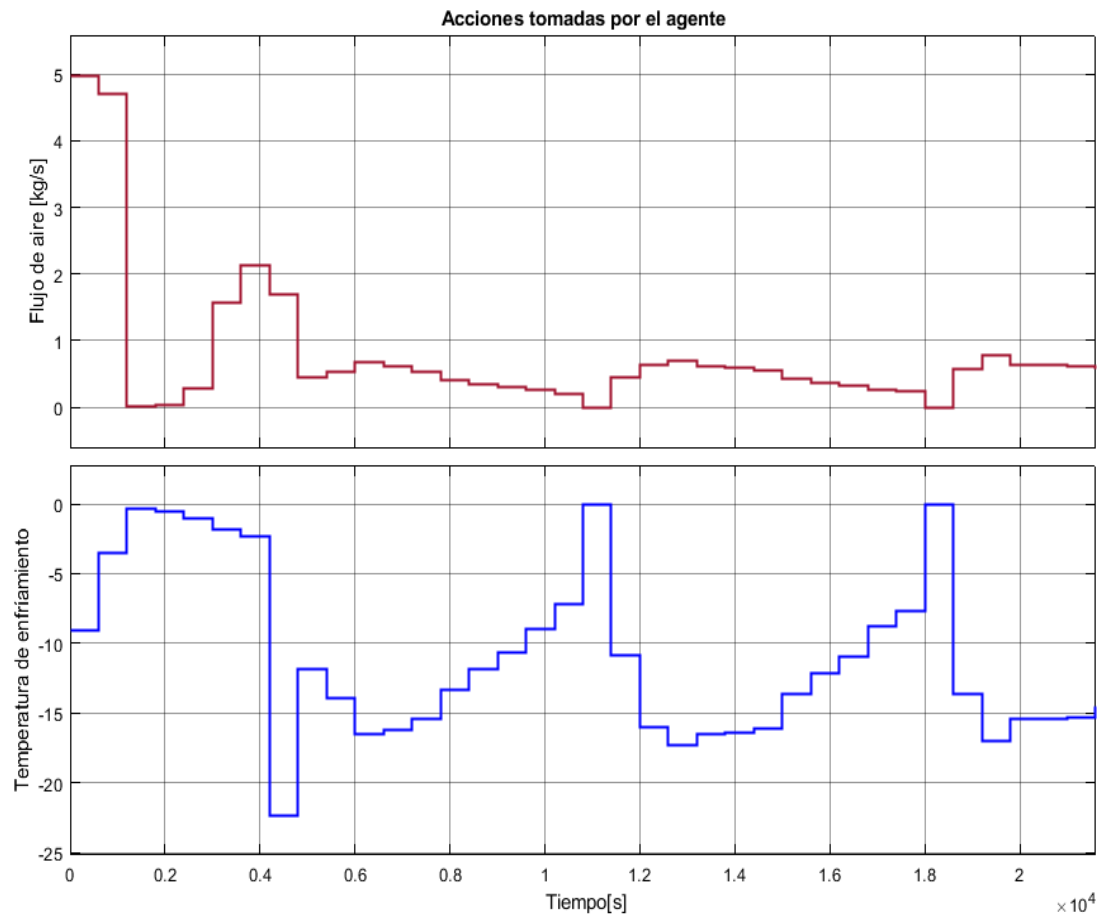


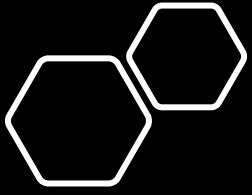


## 2 variables de control incluyendo energía



## 2 variables de control incluyendo energía – TD3





# 4 variables de control y CO2

- Se utiliza 1 zona, controlando solo temperatura.
- Se consideran las 4 acciones control.
- Se prueba el sistema con consumo energético.
- Se utilizan perturbaciones no constantes.
- Simulación de 2 días.

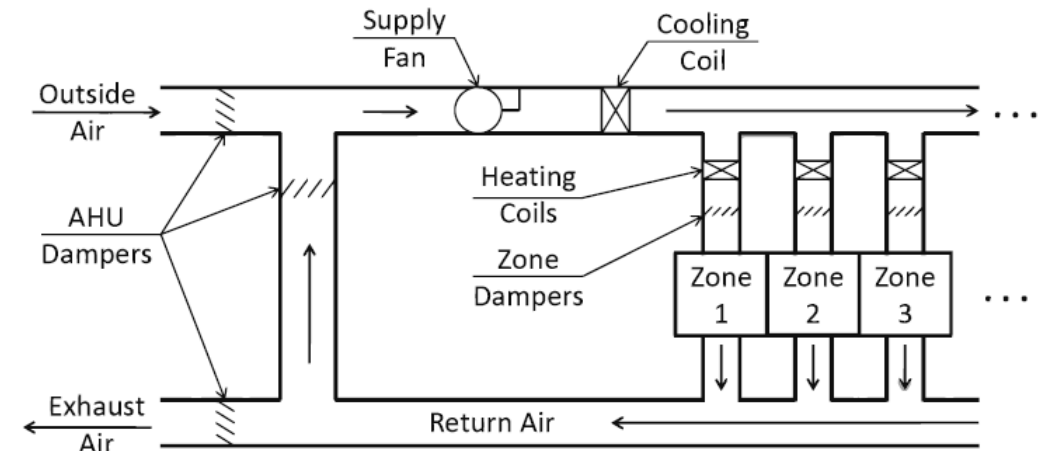
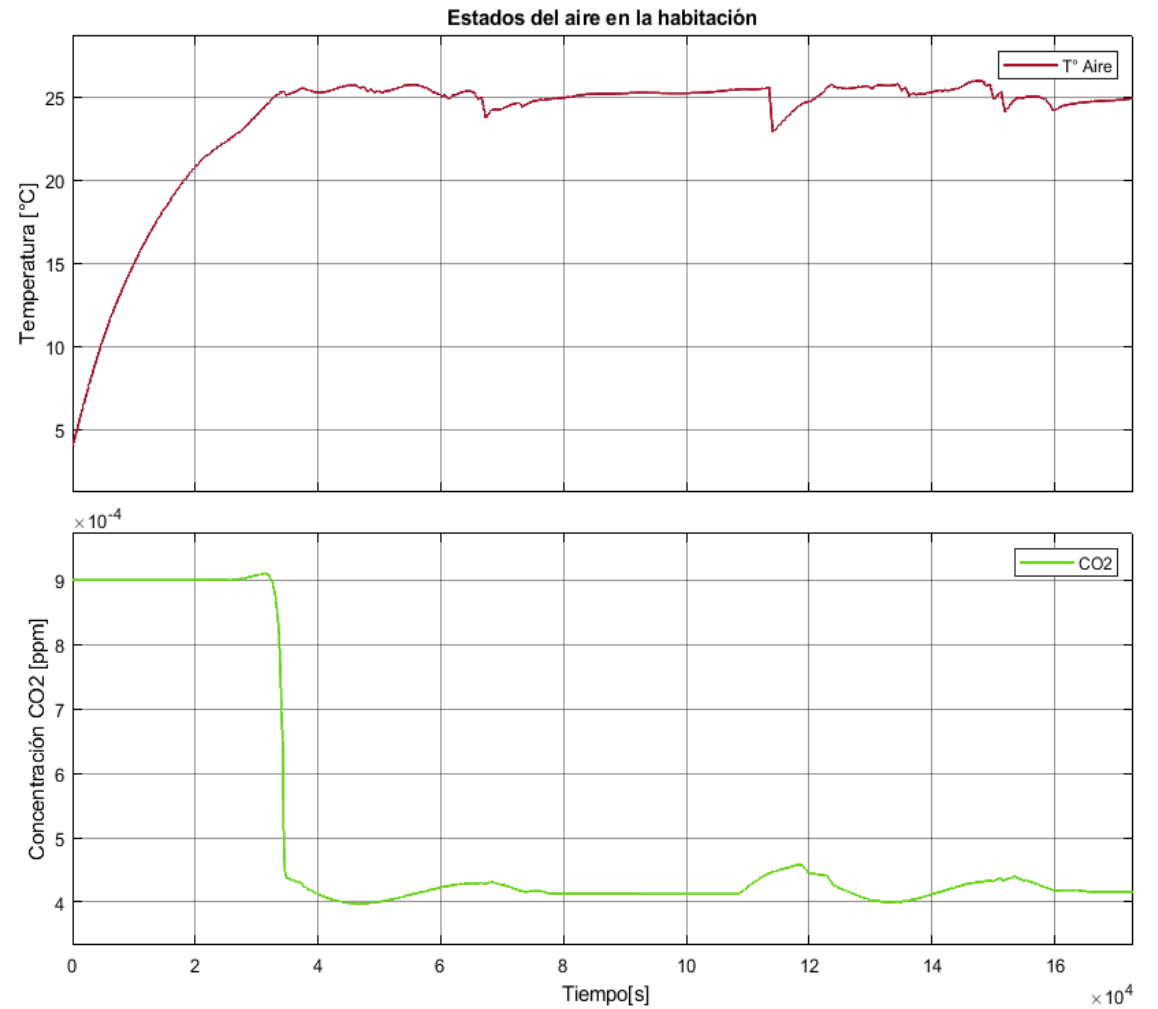
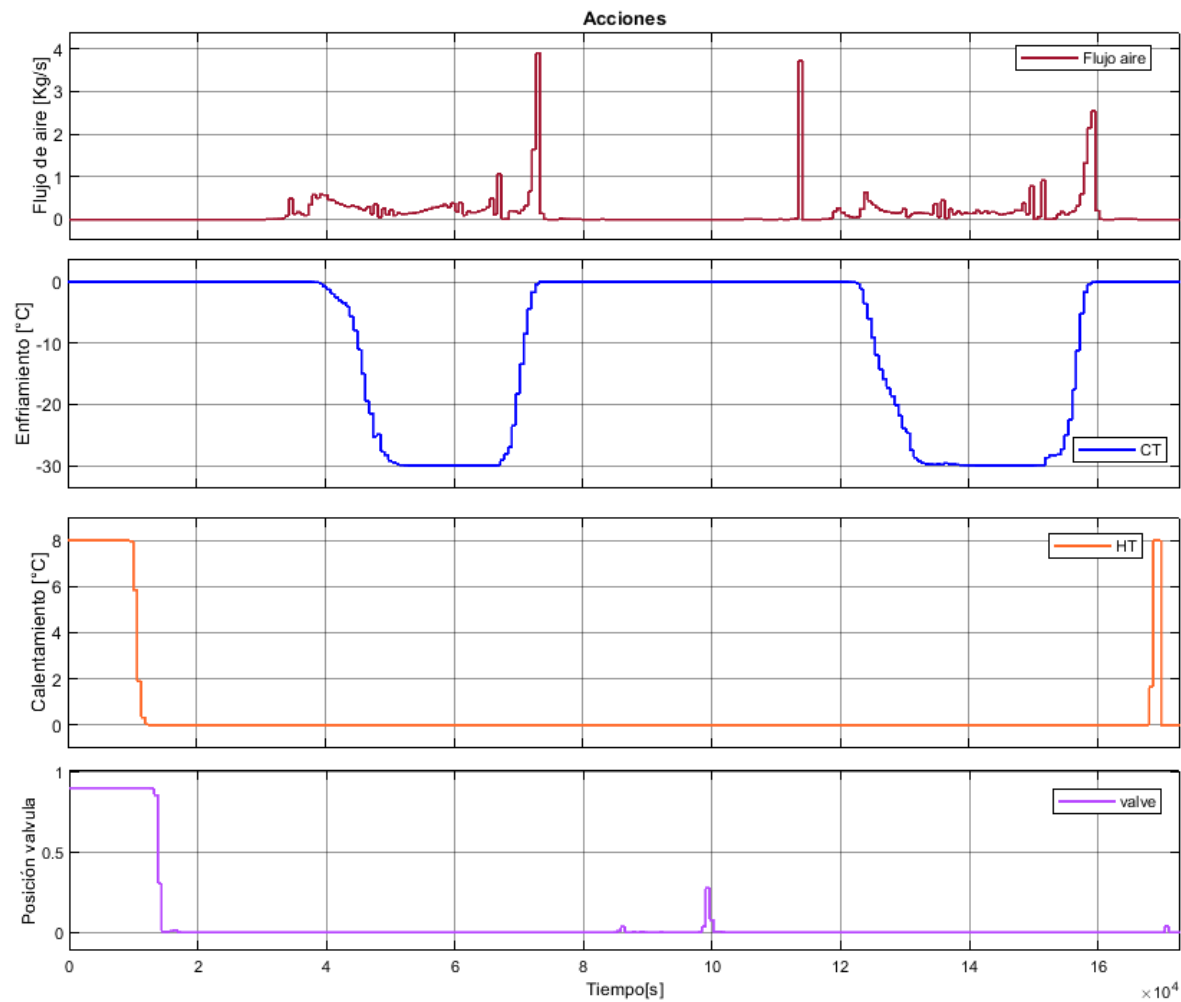
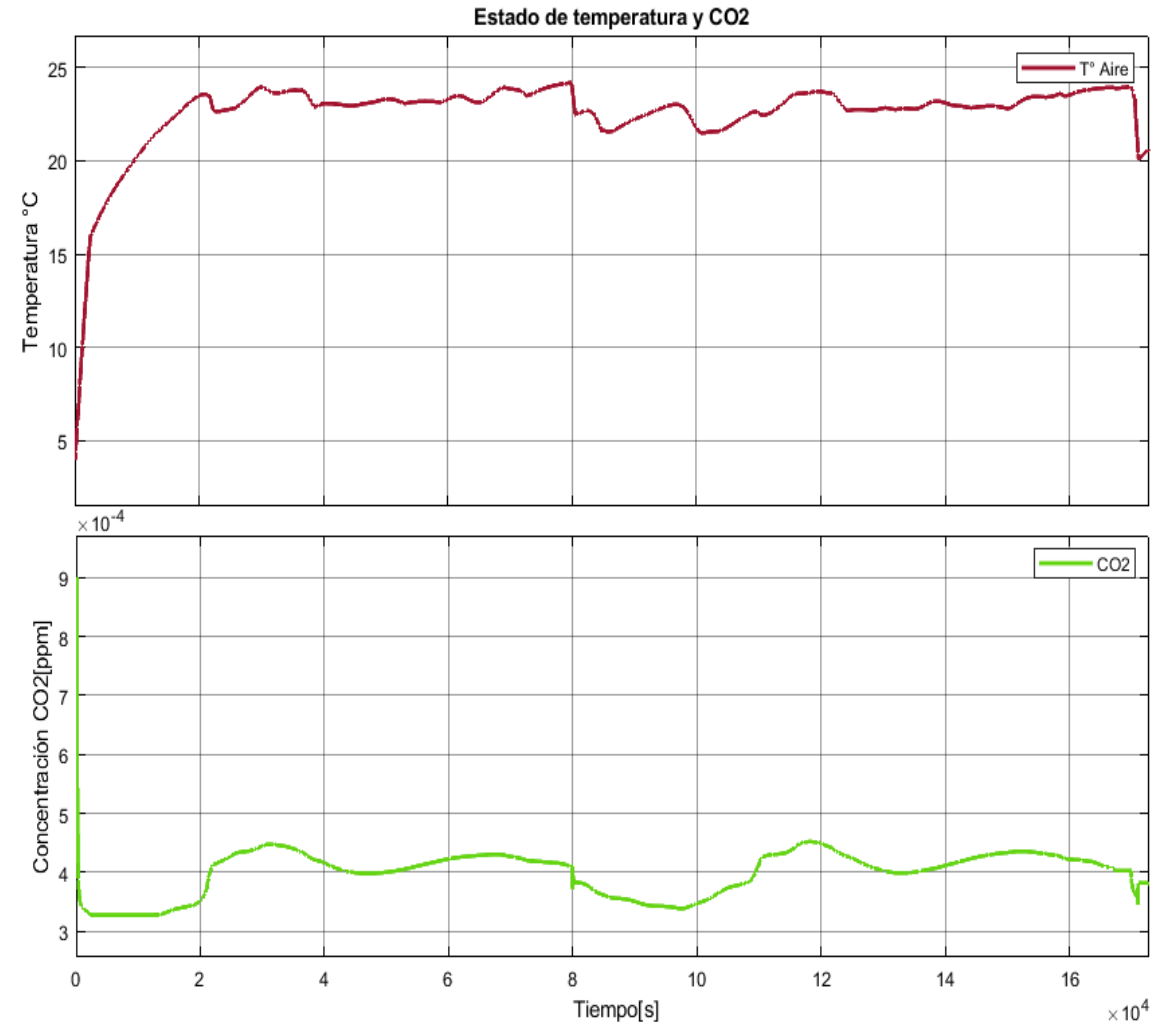
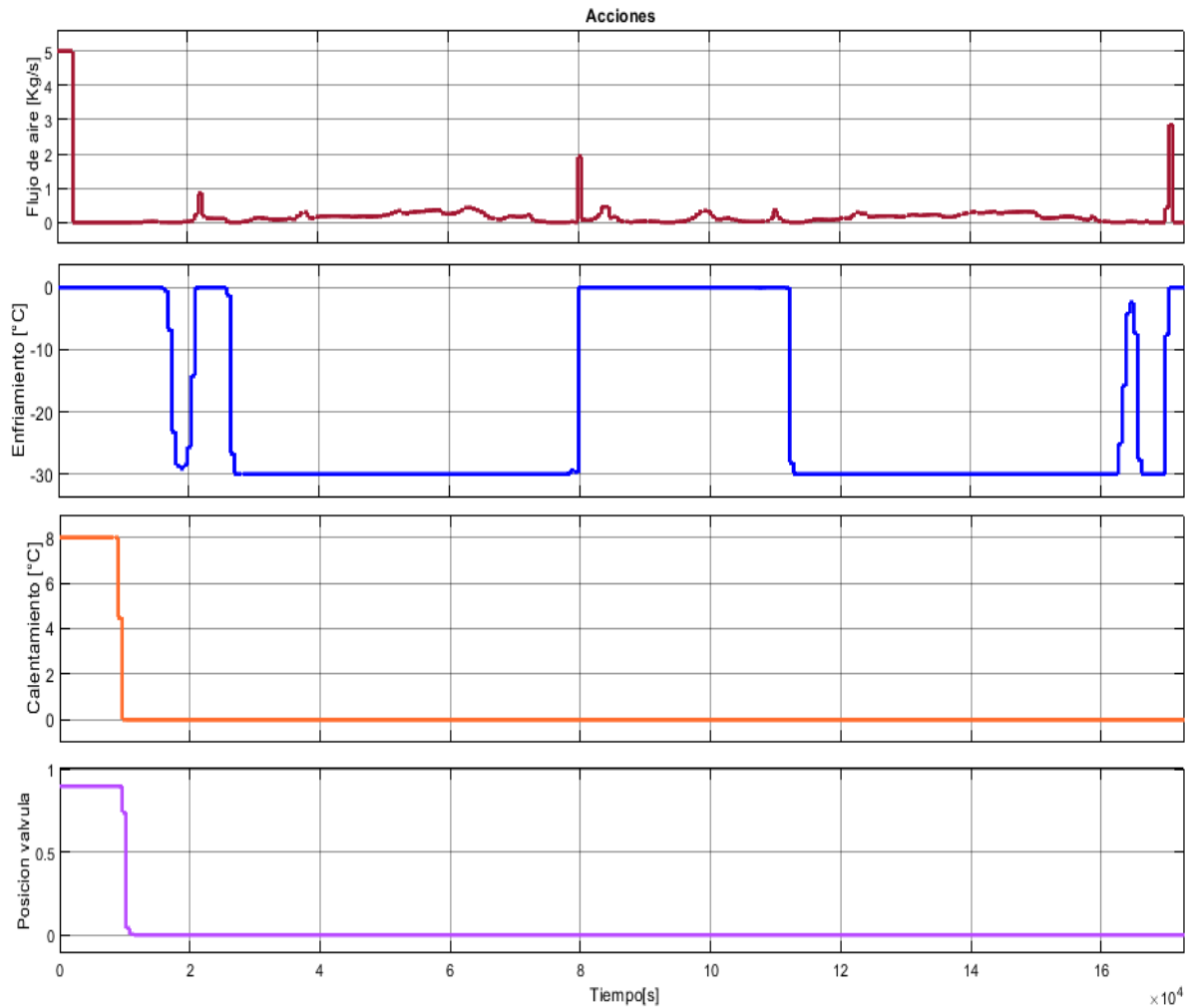


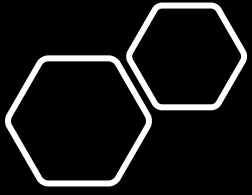
Fig.1: Diagrama HVAC[1]

# 4 variables y CO2



# 4 variables y CO2 – TD3





# 4 variables de control, CO2 y Humedad

- Se utiliza 1 zona, controlando solo temperatura.
- Se consideran las 4 acciones control.
- Se prueba el sistema con consumo energético.
- Se utilizan perturbaciones no constantes.
- Simulación de 2 días.
- Se controlan las 3 restricciones.

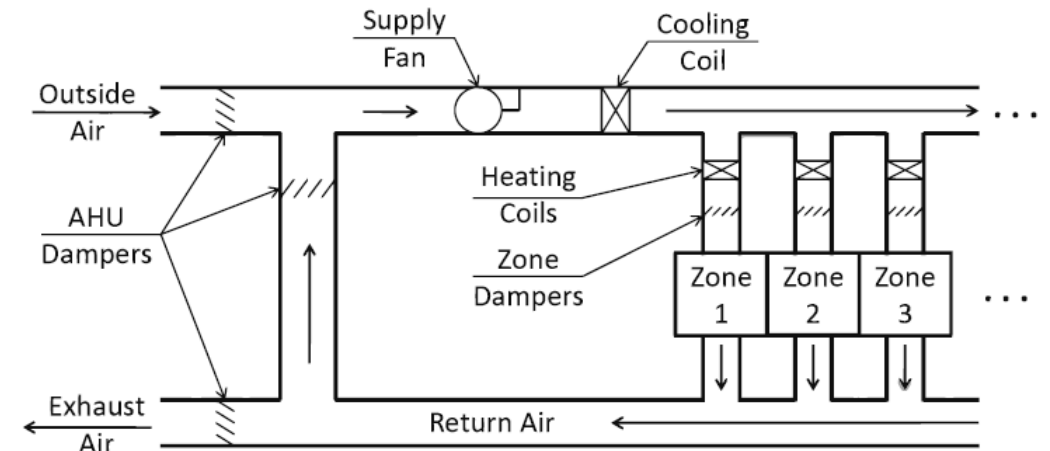
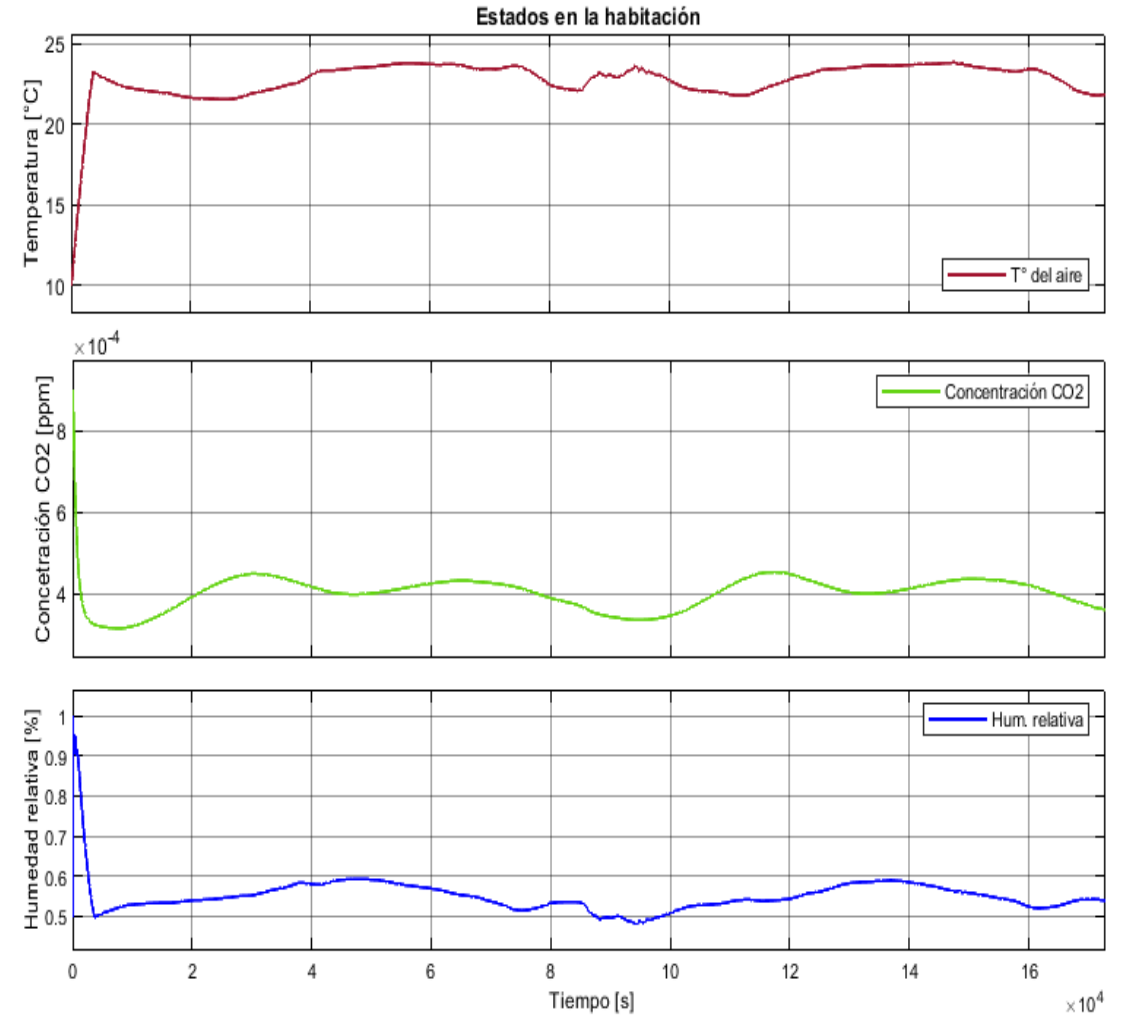
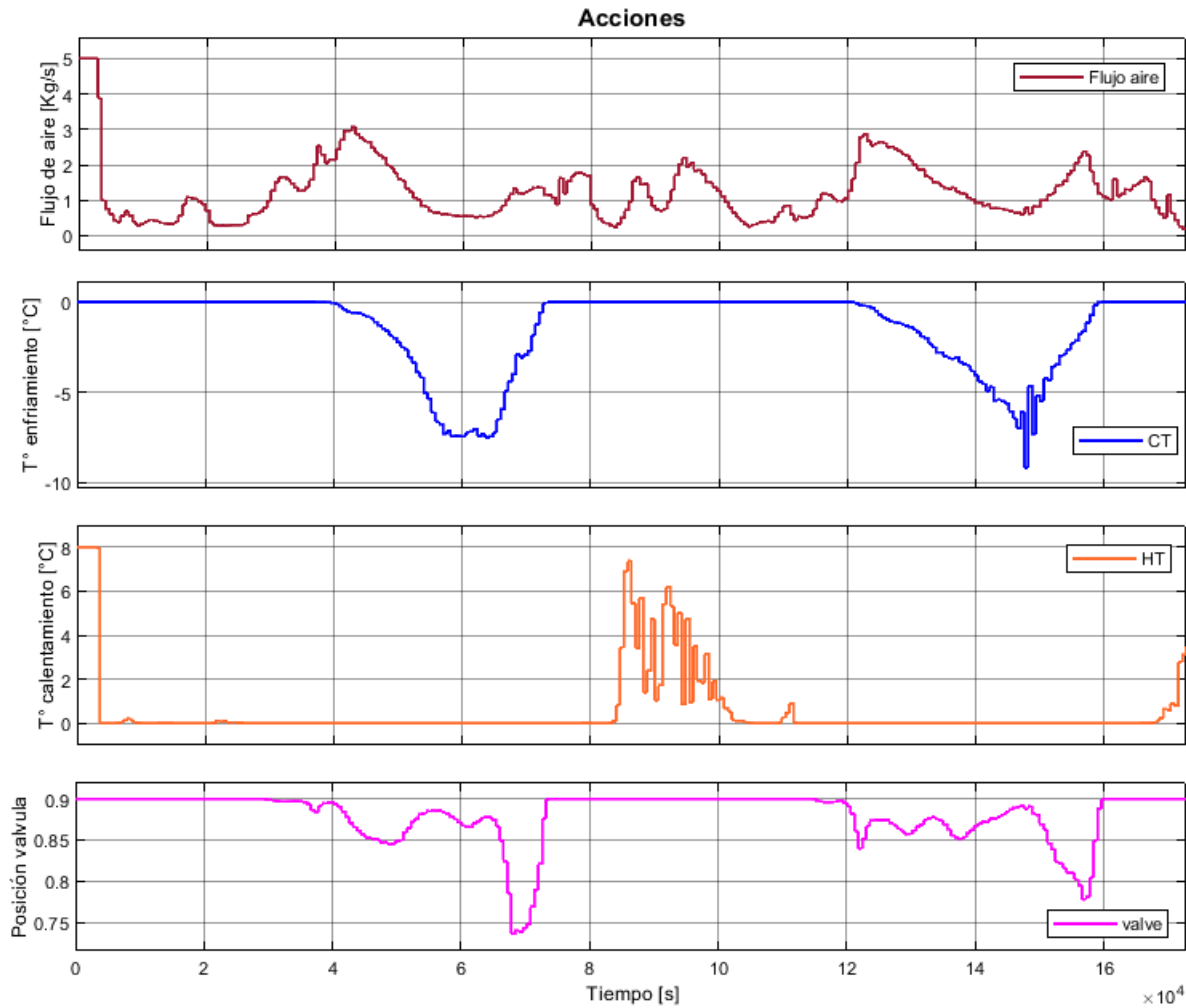
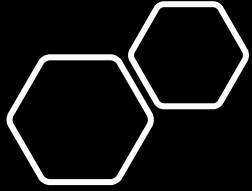


Fig.1: Diagrama HVAC[1]

# 4 variables, CO2 y Humedad – TD3





# Trabajo futuro

- Implementar un tercer algoritmo.
- Realizar comparaciones con otros métodos como MPC.
- Terminar de documentar el código.
- Obtener indicadores numéricos para comparar los distintos métodos.



# Presentación final trabajo de título.

Diseño de sistemas de control de climatización en edificios basados en aprendizaje reforzado.

- Profesor Co-Guía:  
Francisco Rivera
- Profesor Guía:  
Diego Muñoz C.
- Alumno:  
Ricardo López D.