

KEA experiment flow: observation-action sequence

episode_duration=7 | sampling_time=1 | T₀=60.0° | heating=+1/step | cooling=-1/step

■ heating ON

■ heating OFF

— temperature

-- T bounds

method calls

env.reset()

env.step()

env.step()

env.step()

env.step()

env.step()

env.step()

env.step()

observations

obs[0]

obs[1]

obs[2]

obs[3]

obs[4]

obs[5]

obs[6]

obs[7]
(terminal)

scenario data

scn[0]
€0.12

scn[1]
€0.15

scn[2]
€0.18

scn[3]
€0.22

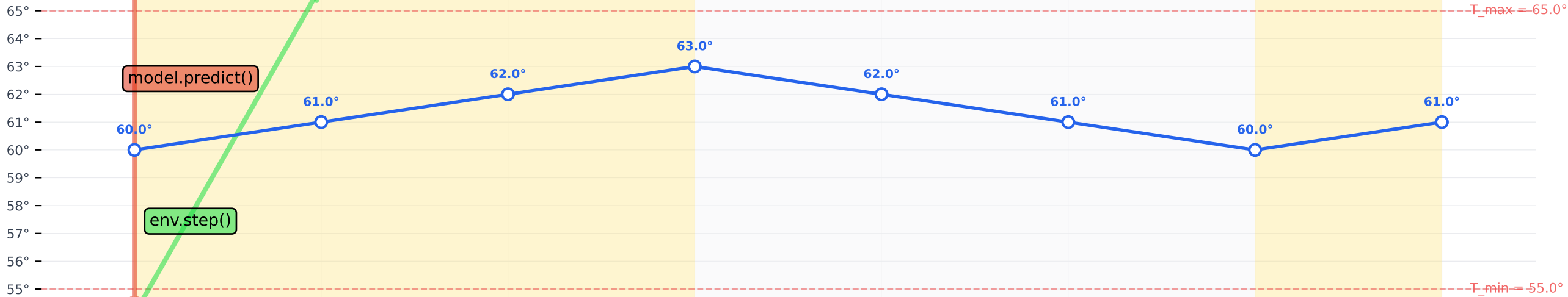
scn[4]
€0.25

scn[5]
€0.20

scn[6]
€0.14

scn[7]
€0.10

Tank temperature (°C)



model.predict()

env.step()

actions

act[0]=ON

act[1]=ON

act[2]=ON

act[3]=OFF

act[4]=OFF

act[5]=OFF

act[6]=ON

t=0

t=1

t=2

t=3

t=4

t=5

t=6

t=7

Timestep

observations: 8 entries (0..7) | actions: 7 entries (0..6) | rewards: 7 entries (0..6) | scenario: 8 entries (0..7)