Q-values estimation using Dyna agent after 2000 episodes