

Q-learning, $N = 101$, $m = 7$, $s = 5$

