Reinforcement learning scheme with proposed model M5

