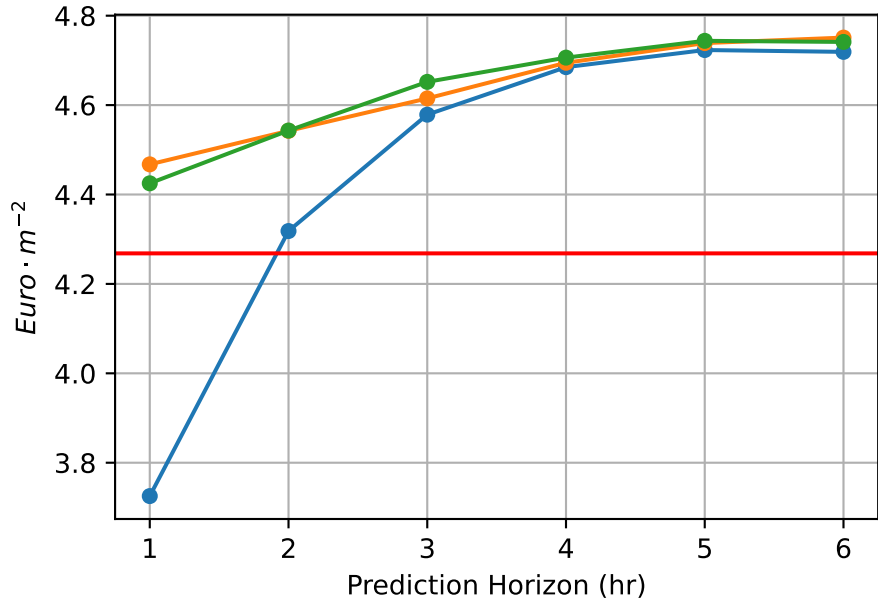
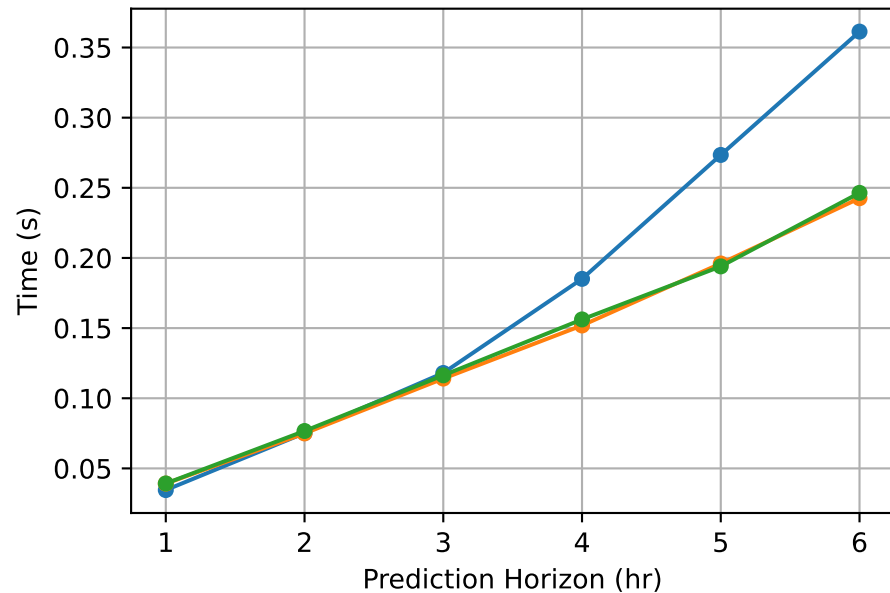


Final Cumulative Reward



Computational Time of Control Input



—●— Nominal MPC —●— RL-MPC 3 with $(\tilde{x}_{k|k}, \tilde{u}_{k|k})$ —●— RL-MPC 3 with $(\hat{x}_{k|k}, \hat{u}_{k|k})$ — RL Agent