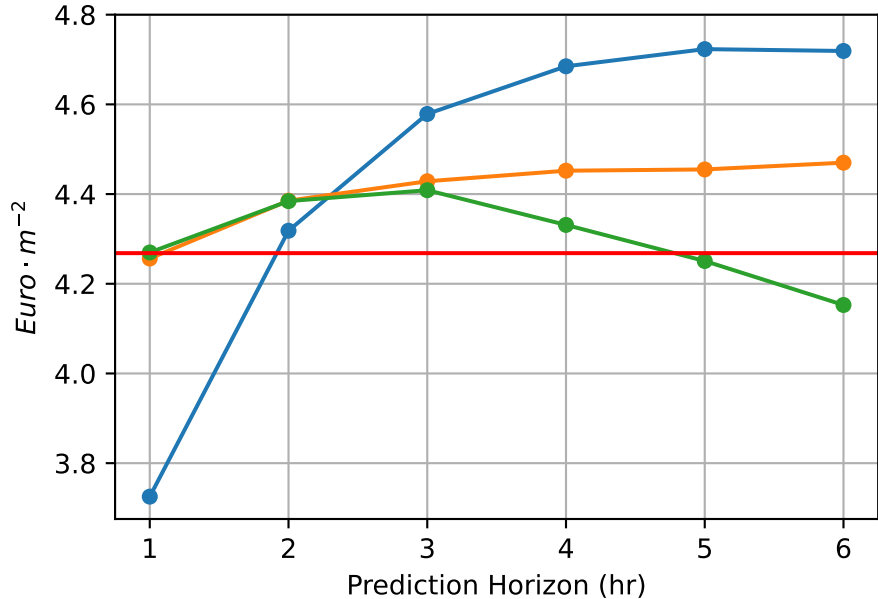
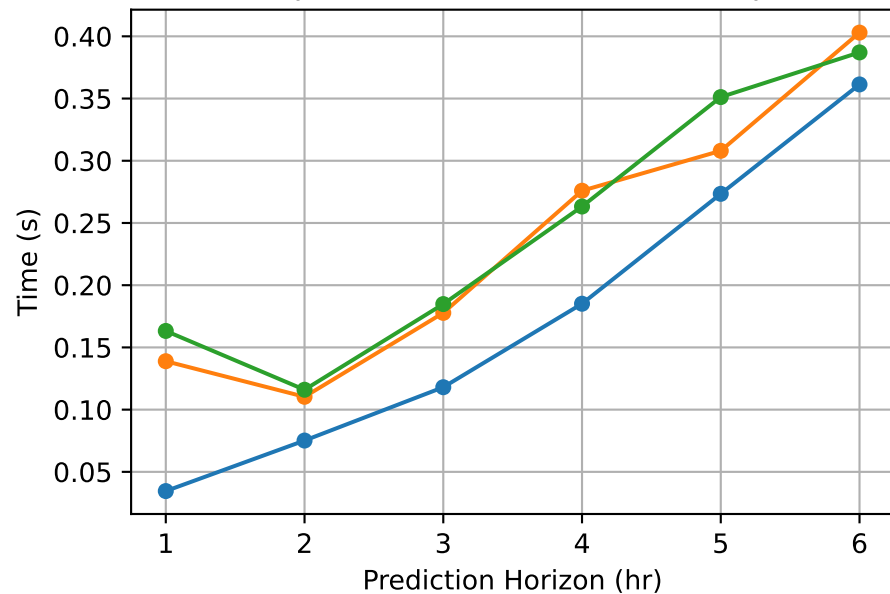


Final Cumulative Reward



Computational Time of Control Input



—●— Nominal MPC    —●— RL-MPC 2 with  $(\tilde{x}_{k|k}, \tilde{u}_{k|k})$     —●— RL-MPC 2 with  $(\hat{x}_{k|k}, \hat{u}_{k|k})$     — RL Agent