Journal of Zhejiang University-SCIENCE A

Time control entry guidance method for hypersonic glide vehicles based on deep reinforcement learning

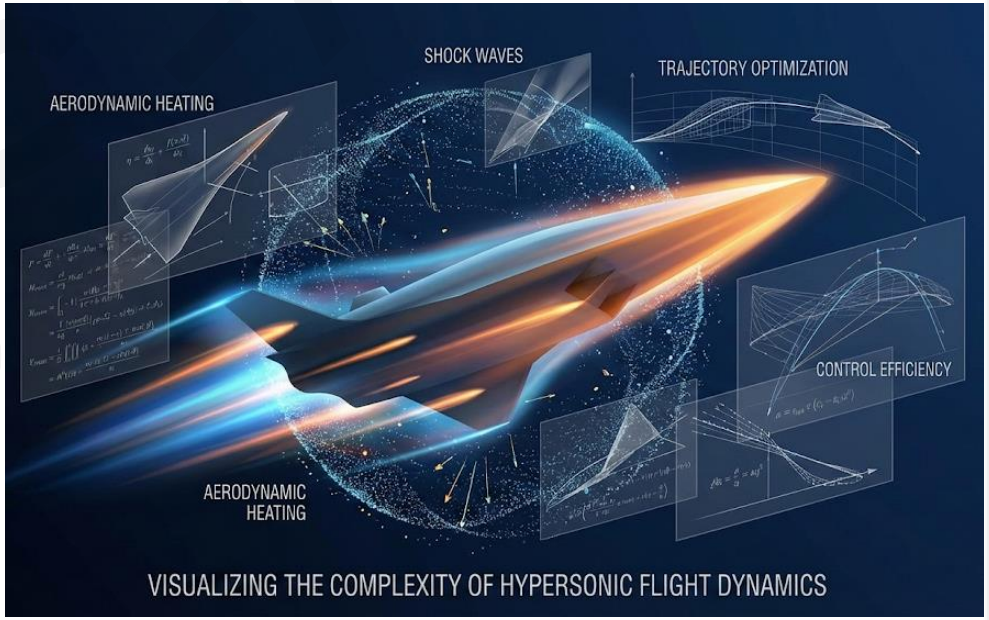
Key words:

Hypersonic glide vehicles; Entry guidance; Reinforcement learning; Time coordination; Deep neural network

Cite this as: Zhenyu LIU, Gang LEI, Yong XIAN, Leliang REN, Shaopeng LI, Daqiao ZHANG, 2026. Time control entry guidance method for hypersonic glide vehicles based on deep reinforcement learning. *Journal of Zhejiang University-SCIENCE A*, 27(4):365-383. https://doi.org/10.1631/jzus.A2500144

Research Background & Objectives

- Challenges: Strong nonlinearity, and strict terminal constraints make HGV entry guidance difficult.
- Traditional Limitations: Numerical methods lack real-time performance; Analytical methods struggle with precision.
- DRL Advantage: Offline training enables online real-time decision making with strong robustness.
- Objective: Achieve precise time-control entry guidance (TCEG) for multi-HGV coordination.



Complexity in Hypersonic Flight Dynamics

Methodology Framework

1. DRL Guidance Agent

- PPO Algorithm
- Learn reference profile tracking command

2. Remaining Flight Time Prediction

- Analytical Profile Estimation
- DNN Residual Correction

3. Online Correction

- Dynamic heading error threshold
- Real-time trajectory adjustment

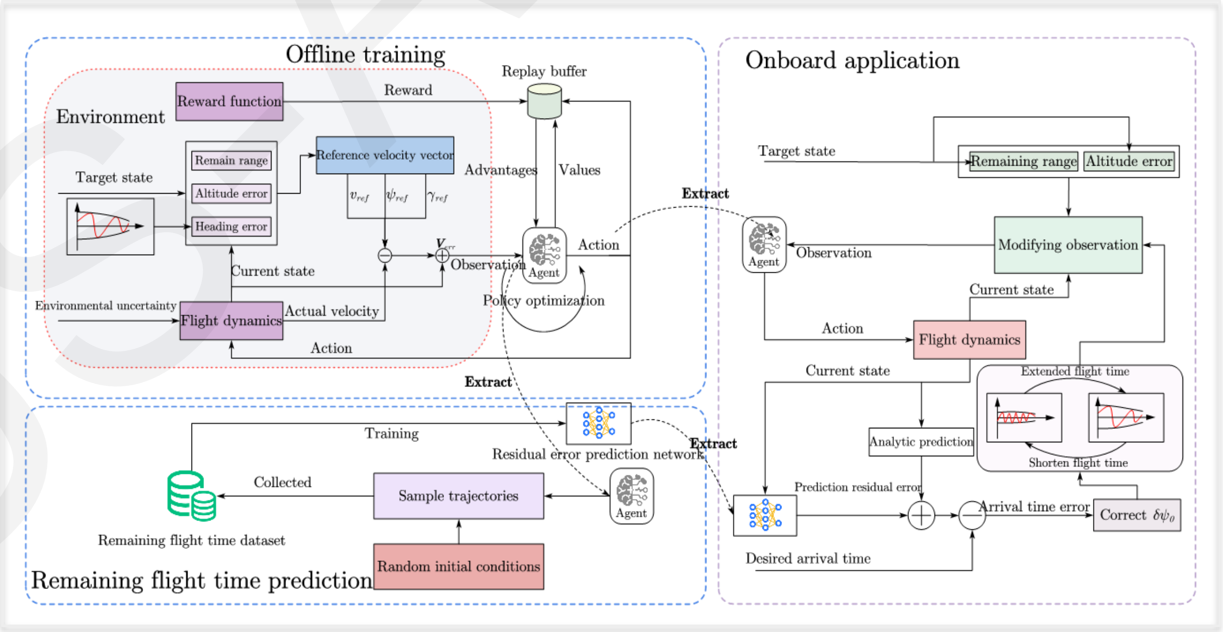


Figure 1: Framework of the time control entry guidance method

Key Technologies

Reference Profile Design

Designed longitudinal and lateral profiles serve as the baseline for training, ensuring physical feasibility and simplifying the learning task.

PPO + GRU Networks

PPO ensures stable updates. GRU handle POMDP by retaining history of states (velocity, range, etc.).

Hybrid Time Prediction

Combines **Analytical Formula** for rough estimation with a **Deep Neural Network (DNN)** to predict residual errors, balancing speed and accuracy.

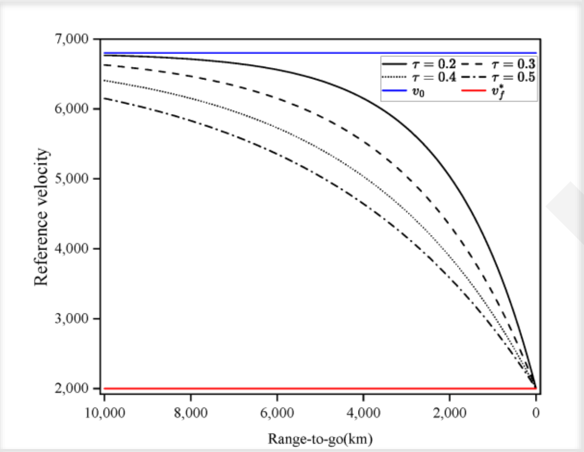


Fig. 2: Reference velocity profile with varying τ

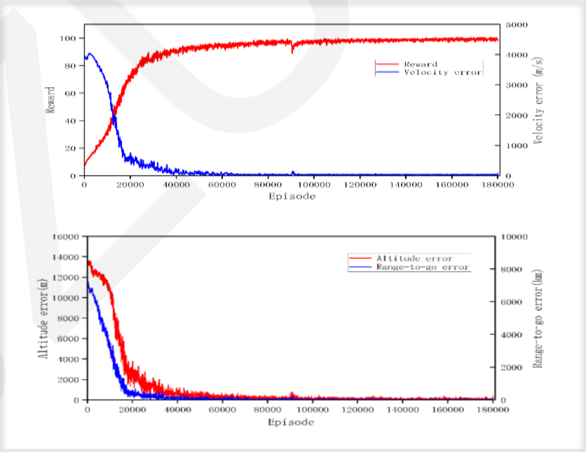


Fig. S3: Reward and terminal error during training process

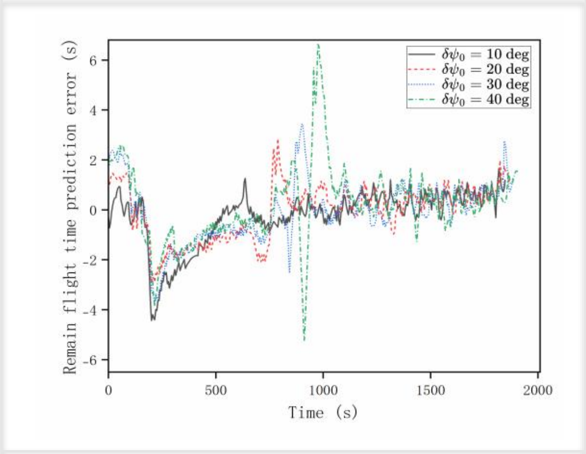


Fig. 7: Estimation error for the remaining flight time

Main Results & Performance

Terminal Time Error

< 1 s

Single HGV achieves high precision terminal time control

Command computation time

55 ms

Meet the real-time requirements of engineering

Termina State Error

- Range error <50m
- Altitude ±100m.
- Velocity ±50m/s

Verified robustness under perturbation conditions

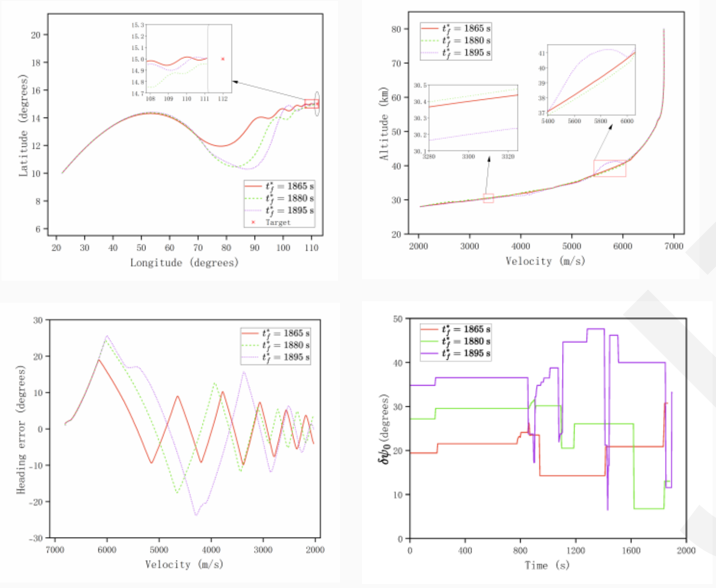


Fig. 8: Simulation results of single HGV with varying terminal time

Table 4: Statistical results of the command computation time

Methods	Mean	Minimum	Maximum
Proposed (s)	0.055	0.019	0.478
Analytical based (s)	0.003	0.001	0.021
Numerical based (s)	0.917	0.073	5.050

Table 5: Comparison of the terminal states

Methods	Proposed	Analytical based	Numerical based
Δs_f(m)	-4.895	-33.404	-14.038
Δh_f (m)	30.281	2744.676	-1412.726
Δv_f(m/s)	32.748	573.004	118.382
$\Delta \psi_f$ (deg)	3.603	6.956	5.747
Δt_f	-0.58	-64.58	-7.54

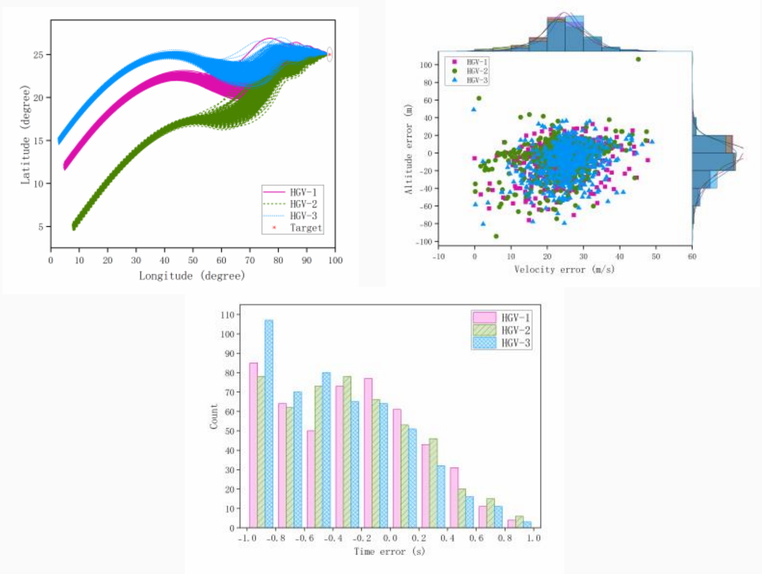


Fig. 16: Monte Carlo simulation results

Conclusion & Contributions

– Real-time Capability

Proposed DRL method eliminates computationally intensive online trajectory generation.

– Hybrid Remain Flight Time Prediction

Combines analytical formulas with DNNs to achieve both high speed and high accuracy.

– Modular Integration

Effectively bridges learning-based control with mission-level guidance objectives.