Journal of Zhejiang University-SCIENCE A 2026 27(4):365-383
www.jzus.zju.edu.cn; www.springer.com/journal/11582
E-mail: jzus_a@zju.edu.cn

Research Article
https://doi.org/10.1631/jzus.A2500144

Check for updates

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

Zhenyu LIU, Gang LEI, Yong XIAN, Leliang REN, Shaopeng LI✉, Daqiao ZHANG

Department of Missile Engineering, Rocket Force University of Engineering, Xi'an 710025, China

Abstract: To meet the requirement of simultaneous arrival for multiple hypersonic glide vehicles (HGVs), we propose a time control entry guidance (TCEG) method leveraging deep reinforcement learning. First, the entry guidance problem is solved with a reinforcement learning framework based on a designed reference flight profile. By appropriately designing the observation space and training environment, the well-trained agent demonstrates robust guidance performance under varying widths of the heading error corridor. Then, a novel method for predicting the remaining flight time is established, which consists of two main components. The first component estimates the remaining flight time using an analytical formula, while the second component employs a deep neural network (DNN) to predict the residual error between the estimated and the true value. Subsequently, based on the predicted terminal time error, the threshold of the heading error and the observation vector are corrected in real time, thereby guiding the agent to dynamically adjust its output actions. This enables precise control of the terminal time. Since the generation of guidance commands only requires forward computations by the neural network, the proposed method exhibits excellent real-time performance. Finally, the effectiveness and robustness of the method are demonstrated through numerical simulations in various scenarios.

Key words: Hypersonic glide vehicles (HGVs); Entry guidance; Reinforcement learning; Time coordination; Deep neural network (DNN)

1 Introduction

Hypersonic glide vehicles (HGVs) have been a recent focus in the aerospace field due to their long-range flight capabilities and high maneuverability (Guo et al., 2020). Current entry guidance methods cannot meet the increasingly diverse requirements of reentry missions, which can include simultaneous arrival or terminal time constraints. Consequently, time control entry guidance (TCEG) methods with high adaptability have attracted increasing research attention.

The primary objective of an entry guidance system is to guide the HGV from the entry interface (EI) to the required terminal area energy management (TAEM) interface under various constraints (Lu, 2014; Hu et al., 2022). Current entry guidance methods can be categorized into two main types: reference trajectory tracking guidance (Lu, 1997; Li et al., 2018) and predictor–corrector guidance (Xue and Lu, 2010; Liang et al., 2017; Zeng et al., 2018). The reference trajectory tracking guidance method consists of two parts: trajectory planning and tracking. This method has the advantage of low real-time computational load; however, due to its reliance on a predetermined reference trajectory, it has limitations in dealing with initial state deviations and environmental disturbances (Harpold and Gavert, 1983; Brunner and Lu, 2012). The predictor–corrector guidance method does not require a predetermined reference trajectory; thus, it is highly robust (Lu, 2014). However, this method involves repeated trajectory predictions and iterative calculations, which can hinder real-time performance (Xue and Lu, 2010).

Currently, there are numerous studies on time control guidance for tactical missiles (Kim et al., 2020; Hu et al., 2021) or unmanned aerial vehicles (UAVs) (Li et al., 2023; Suresh et al., 2023; Wang et al., 2024), but there are few investigations on TCEG methods. Li et al. (2019) proposed a TCEG method based on predictor–corrector guidance. Initially, the terminal

✉ Shaopeng LI, lishaopeng.2021@tsinghua.org.cn

Received Apr. 27, 2025; Revision accepted Nov. 7, 2025;
Crosschecked Jan. 26, 2026; Online first Mar. 12, 2026

time is adjusted roughly using a lateral guidance approach, then longitudinal guidance is used to exert precise control of the terminal time. However, this method faces significant challenges in terms of real-time performance. Yu et al. (2019) derived analytical formulas to predict the range and terminal time while also accounting for the Earth's rotation. The guidance commands were iteratively solved in the longitudinal flight profile to satisfy the terminal constraints. Moreover, Guo et al. (2020) derived analytical prediction formulas for time-to-go and range-to-go. In particular, they relaxed the terminal velocity constraint to simultaneously satisfy other terminal constraints. Also, Liu et al. (2025) proposed a TCEG method for adjusting the terminal time in the longitudinal channel. First, the height–velocity (H–V) profile that satisfies the terminal time constraint is obtained through a numerical optimization algorithm. Later, the terminal range is predicted by numerically integrating the reduced-order equations of motion, and the range-to-go error is corrected by adjusting the heading error threshold. Wang HN et al. (2022) developed a time coordination method that is based on a range coordination strategy. Similarly, it first plans a H–V profile in the longitudinal guidance that satisfies the terminal time constraint; then, the lateral guidance plans the flight trajectory based on Bézier curves to eliminate range errors.

Based on the analysis above, currently, TCEG methods can be achieved in two different ways. The first is a lateral guidance-based approach (Li et al., 2019; Yu et al., 2019), which achieves terminal time control by adjusting the length of the HGV's flight trajectory through lateral guidance, while longitudinal guidance is used to correct terminal range errors. The second is the longitudinal guidance-based approach (Guo et al., 2020; Wang HN et al., 2022; Liu et al., 2025), which involves designing a reference flight profile that satisfies the terminal time constraints in the longitudinal plane; in contrast, lateral guidance primarily aims to correct range-to-go errors. However, both methods require repeated trajectory prediction and correction, facing significant challenges in terms of real-time performance. Although using analytical formulas in the prediction process can significantly reduce the computational burden, the derivation of these formulas often relies on numerous assumptions and empirical parameters. This introduces different challenges, such as accurately estimating prediction errors, thereby necessitating continuous correction commands throughout the entire entry flight.

The application of deep reinforcement learning (DRL) to flight guidance and control (Cheng et al., 2021; Ren et al., 2023, 2025) has attracted substantial attention in recent years due to its potential for multi-objective optimization and sequential decision-making on high-dimensional data. For example, Wang NY et al. (2022) established an impact time control guidance law for tactical missiles and determined the guidance gain using the deep deterministic policy gradient (DDPG) algorithm. Lee et al. (2023) established a guidance law with a field of view (FOV) constraint by learning bias guidance commands in multiple learning stages. Furthermore, Qiu et al. (2024) established a recorded recurrent twin delayed deep deterministic (RRTD3) policy gradient algorithm to solve the guidance problem of intercepting endoatmospheric maneuvering missiles. Yang HW et al. (2024) presented a robust DRL-based low-thrust guidance method for approaching asteroids under process uncertainties.

Nevertheless, only a few studies have used reinforcement learning to address the entry guidance problem. Gaudet et al. (2022) used DRL to optimize an adaptive entry guidance system that is suitable for HGVs; however, terminal time constraints were not considered. Gao et al. (2024) established an integrated entry guidance law for HGVs with no-fly zone constraints based on DRL; nevertheless, the training environment only considered scenes with a specific no-fly zone. Bao et al. (2025) developed a hierarchical guidance framework for hypersonic morphing vehicles under multiple no-fly zone constraints, in which DDPG was used to optimize the vehicle's morphing control strategy. Unfortunately, this method can only be applied to scenarios with fixed no-fly zone distributions.

The foregoing analysis reveals a distinct gap in the current literature. On one hand, traditional TCEG methods can manage terminal time but often face issues with real-time computational burdens, limiting their practicality. On the other hand, DRL-based entry guidance methods have demonstrated remarkable real-time performance and robustness but have not yet adequately addressed the problem of terminal time control. Therefore, we propose a DRL-based TCEG method to solve the terminal time constraint for HGVs. The principal contributions of this study are summarized as follows.

(1) A real-time DRL-based entry guidance method is developed, which achieves high precision and robustness without the need for computationally intensive online trajectory predictions. Meanwhile, the agent is uniquely trained to adapt its policy to a variable heading error corridor, which serves as the direct interface for the time control commands.

(2) A hybrid time-to-go prediction method combining an analytical formula with a deep neural network (DNN)-based residual estimator is established. This approach offers higher accuracy than purely analytical models and far greater computational efficiency than numerical integration-based prediction, making it uniquely suitable for a real-time guidance loop.

(3) A novel TCEG method is proposed. By adaptively modulating the heading error corridor based on the feedback from the rapid time prediction module, it precisely controls the terminal time without altering the core DRL agent; thus, this method demonstrates an effective modular integration of learning-based control and mission-level objectives.

2 Entry problem formulation

2.1 Entry dynamics

Considering the scenario of a spherical and rotating Earth, the three-degree of freedom (DOF) dynamics equations of an HGV are described as follows (Shen and Lu, 2003):

$$\dot{r}=v\sin\gamma, \tag{1a}$$

$$\dot{\lambda}=\frac{v\cos\gamma\sin\psi}{r\cos\phi}, \tag{1b}$$

$$\dot{\phi}=\frac{v\cos\gamma\cos\psi}{r}, \tag{1c}$$

$$\dot{v}=-\frac{D}{m}-g\sin\gamma+C_v, \tag{1d}$$

$$\dot{\gamma}=\frac{1}{v}\left[\frac{L\cos\sigma}{m}+\left(\frac{v^2}{r}-g\right)\cos\gamma+C_\gamma\right], \tag{1e}$$

$$\dot{\psi}=\frac{1}{v}\left[\frac{L\sin\sigma}{m\cos\gamma}+C_\psi\right], \tag{1f}$$

$$\dot{\alpha}=\Delta\alpha, \tag{1g}$$

$$\dot{\sigma}=\Delta\sigma, \tag{1h}$$

where r is the distance between the HGV and the center of the Earth; v is the current velocity of the HGV; λ and ϕ are the current geocentric longitude and latitude of the HGV, respectively; γ and ψ are the current flight-path angle and heading angle of the HGV, respectively; α and σ are the attack angle and bank angle of the HGV, respectively; L and D are the values of lift and drag forces, respectively; m is the mass of the HGV; g is the gravitational acceleration; C_v, C_γ, and C_ψ are the components of the Earth's rotation, for which detailed expressions can be found in Liu et al. (2025). Details of the aerodynamic force calculation model, actuator delay, and uncertainty model are provided in Section S1 of the electronic supplementary materials (ESM).

2.2 Entry constraints

2.2.1 Path constraints

Path constraints must be considered to achieve safe entry flight trajectories. Typical path constraints can be formulated as follows:

$$\dot{Q}=k_Q\sqrt{\rho}\,v^{3.15}\leqslant\dot{Q}_{\max}, \tag{2a}$$

$$n=\frac{\sqrt{L^2+D^2}}{mg_0}\leqslant n_{\max}, \tag{2b}$$

$$q=\frac{1}{2}\rho v^2\leqslant q_{\max}, \tag{2c}$$

where $k_Q>0$ is a constant value associated with the vehicle; ρ is the atmospheric density; g_0 is the gravitational acceleration at sea level; $\dot{Q}$, n, and q are the heating rate, aerodynamic overload, and dynamic pressure, respectively; $\dot{Q}_{\max}$, $n_{\max}$, and $q_{\max}$ are the maximum allowed values of heating rate, aerodynamic overload, and dynamic pressure, respectively.

2.2.2 Control constraints

At the same time, the control variables must be restricted within reasonable limits, and the angular rates should also be considered, that is:

$$\alpha\in[\alpha_{\min},\,\alpha_{\max}], \tag{3a}$$

$$\sigma\in[\sigma_{\min},\,\sigma_{\max}], \tag{3b}$$

$$\dot{\alpha}\leqslant\dot{\alpha}_{\max}, \tag{3c}$$

$$\dot{\sigma}\leqslant\dot{\sigma}_{\max}, \tag{3d}$$

where $\alpha_{\min}$, $\alpha_{\max}$, $\sigma_{\min}$, and $\sigma_{\max}$ are the minimum and maximum permissible values of the attack angle and bank angle, respectively; $\dot{\alpha}_{\max}$ and $\dot{\sigma}_{\max}$ are the maximum limitations of change rates for the attack angle and bank angle, respectively.

2.2.3 Terminal constraints

Terminal constraints are imposed to ensure that the HGV successfully reaches the TAEM interface, and can be expressed as follows:

$$|t_f - t_f^*| \leqslant \delta t_f, \tag{4a}$$
$$|s_f - s_f^*| \leqslant \delta s_f, \tag{4b}$$
$$|h_f - h_f^*| \leqslant \delta h_f, \tag{4c}$$
$$|v_f - v_f^*| \leqslant \delta v_f, \tag{4d}$$
$$|\Delta\psi_f| = |\psi_f - \psi_{LOS}| \leqslant \delta\psi_f, \tag{4e}$$

where t_f is the terminal time; s_f, h_f, v_f, and ψ_f are the range-to-go, altitude, velocity, and heading angle at the terminal time t_f, respectively; t_f^* is the desired terminal time; s_f^*, h_f^*, and v_f^* are the desired range-to-go, altitude, and velocity at the desired terminal time t_f^*, respectively; $\Delta\psi_f$ is the heading error between the heading angle ψ_f and the line-of-sight (LOS) angle ψ_{LOS}. Specifically, ψ_{LOS} is the azimuth angle of the great circle connecting the current location and the target location. δt_f, δs_f, δh_f, δv_f, and $\delta\psi_f$ are the tolerance deviations of the terminal states.

3 Time control entry guidance algorithm

We develop a novel TCEG framework, as illustrated in Fig. 1, wherein longitudinal guidance satisfies terminal range constraints, and lateral guidance controls the flight time by adjusting the heading error corridor width. Additionally, reinforcement learning techniques are employed to train an agent to meet path constraints during the guidance, ensuring safe entry flight.

3.1 Reference flight profile design

Conventional longitudinal guidance converts path constraints into flight corridors, such as the drag–energy (D–E) and height–velocity (H–V) corridors, based on quasi-equilibrium glide assumptions. However, by neglecting Earth's rotation and employing an exponential atmospheric density model, there is a risk of violating these constraints due to environmental disturbances, even when the flight trajectory stays within the designated corridor. To address this issue, we design the reference profile without considering path constraints, instead focusing solely on terminal constraints. Subsequently, reinforcement learning techniques are used to train an agent to satisfy these path constraints while tracking the reference profile, which will be discussed in Section 3.2.

3.1.1 Longitudinal profile

First, to meet the terminal velocity constraint and the range constraints, the reference velocity v_{ref} profile is designed as follows (Gaudet et al., 2022):

$$v_{ref} = v_0 - (v_0 - v_f^*)\exp\left(-\frac{s}{\tau s_0}\right), \tag{5}$$

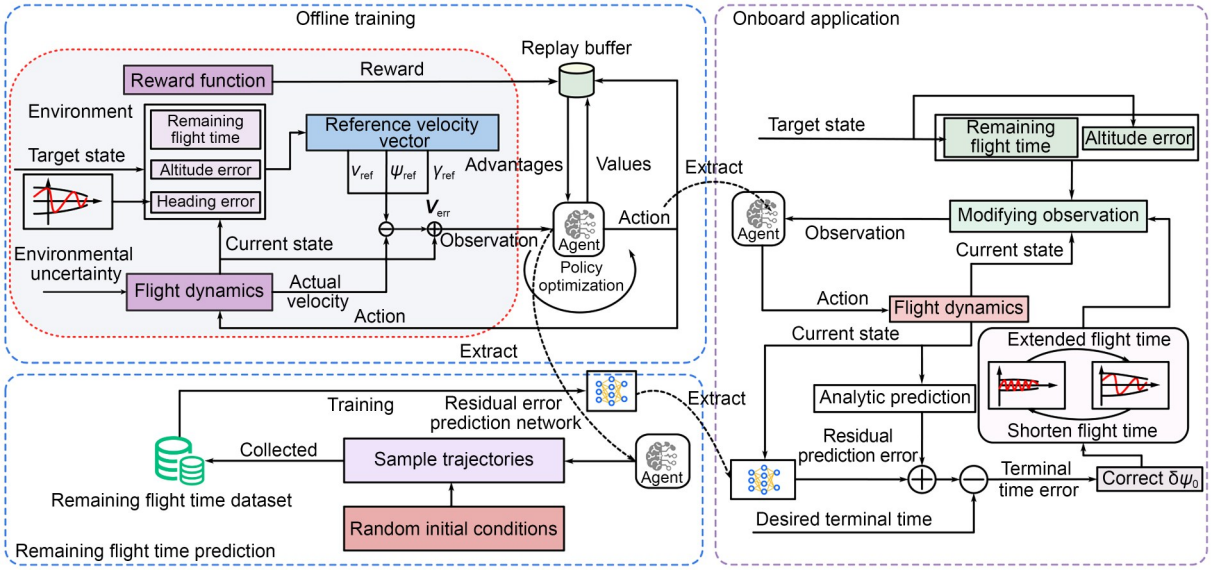


Fig. 1 Framework of the TCEG method. ψ_{ref} and γ_{ref} are the reference flight-path angle and lateral reference, respectively; V_{err} is the reference velocity tracking error vector; ψ_0 is the initial heading angle; $\delta\psi_0$ is the heading error corridor adjustment parameter

where v_0 is the initial velocity; s_0 is the initial range-to-go; τ is a constant value (which is set to 0.25 in this paper). The current range-to-go s can be calculated using the Haversine distance formula (Guo et al., 2020). It should be noted that different values of τ result in different reference profiles. As depicted in Fig. 2, although a larger value of τ leads to a smoother reference velocity profile, it also yields a larger error between the reference and initial velocity profiles. In other words, larger velocity error signals will be produced during the initial flight phase, which may affect the agent's ability to track the reference profile. After several simulation experiments, we found that setting τ=0.25 achieves a good balance between smoothness and tracking performance.

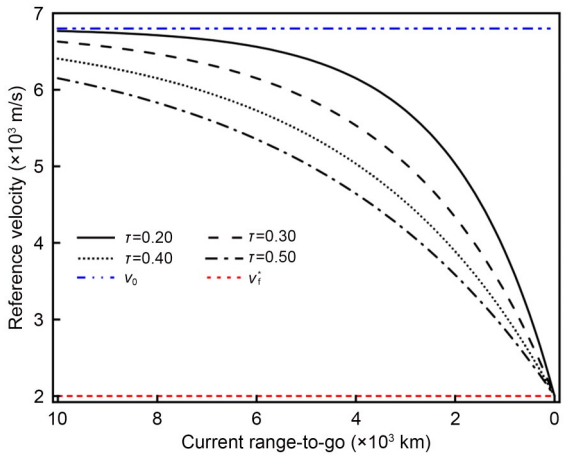


Fig. 2 Reference velocity versus range-to-go with varying τ values

Furthermore, to meet terminal altitude constraints and ensure the flight trajectory is as smooth as possible, the reference flight-path angle γ_{ref} profile is designed as follows:

$$\gamma_{\text{ref}}=\arcsin\left(\frac{h_{\text{f}}^{*}-h}{s_{\text{f}}}\right), \tag{6}$$

where $h=r-R_{\text{E}}$ is the current altitude of the HGV, and R_{E} is the radius of the Earth.

It is important to note that Eqs. (5) and (6) represent merely one solution among an infinite number of flight profiles satisfying the terminal constraints. In practice, any flight profile that meets the terminal conditions can serve as an alternative to Eqs. (5) and (6).

3.1.2 Lateral profile

The lateral reference profile must ensure that the HGV continuously heads towards the target and satisfies the terminal heading error constraints.

Obviously, the reference heading angle can be set directly as the line-of-sight angle. However, this may lead to frequent bank angle reversals, degrading the stability of the HGV. Therefore, based on traditional lateral guidance methods concerning heading error, the following lateral reference ψ_{ref} profile is established:

$$\psi_{\text{ref}}=\begin{cases}\psi, & \text{if } |\psi-\psi_{\text{LOS}}|\leqslant\delta\psi_{\text{thr}},\\ \psi_{\text{LOS}}, & \text{others},\end{cases} \tag{7}$$

$$\psi_{\text{LOS}}=\arctan\left(\frac{\sin(\lambda_{\text{f}}^{*}-\lambda)}{\cos\phi\tan\phi_{\text{f}}^{*}-\sin\phi\cos(\lambda_{\text{f}}^{*}-\lambda)}\right), \tag{8}$$

where λ_{f}^{*} and ϕ_{f}^{*} are the desired longitude and latitude at the desired terminal time t_{f}^{*}; $\delta\psi_{\text{thr}}$ is the heading error threshold, which is designed as a velocity-dependent function (Liu et al., 2025):

$$\delta\psi_{\text{thr}}(v)=\frac{\sqrt{v}-\sqrt{v_0}}{\sqrt{v_{\text{f}}^{*}}-\sqrt{v_0}}(\delta\psi_{\text{f}}-\delta\psi_0)+\delta\psi_0. \tag{9}$$

Eq. (9) indicates that a larger value of $\delta\psi_0$ corresponds to a wider heading error corridor. Consequently, this leads to a flight trajectory that is characterized by fewer bank angle reversals and longer flight times.

As shown in Eq. (7), $|\psi-\psi_{\text{ref}}|$ is equal to 0 when the value of $|\psi-\psi_{\text{LOS}}|$ is less than the heading error threshold $\delta\psi_{\text{thr}}$. This makes the HGV maintain the sign of the current bank angle, thereby preventing frequent reversals.

Remark 1 Based on the analysis above, it can be observed that focusing solely on terminal constraints significantly reduces the complexity needed to design the reference flight profile. In contrast, traditional design approaches inherently involve trimming and compressing the solution space of the entry guidance problem, making it challenging to ensure an optimal flight profile. The method proposed in this paper does not manipulate the solution space during the design phase, and thus theoretically enables the derivation of an optimal flight trajectory.

3.2 Tracking the reference profile via deep reinforcement learning

In this section, we employ reinforcement learning techniques for training agents to control the HGV, so as to track the designed reference flight profile while satisfying all path constraints.

3.2.1 Observation space and action space

To enable the agent to learn how to track a reference flight profile while simultaneously satisfying path constraints, the observation space must include the current state of the HGV as well as other constraint variables. Consequently, the observation vector $\boldsymbol{o}_t$ is defined as follows:

$$\boldsymbol{o}_t=[V_{err},\ s,\ v,\ \alpha,\ \sigma,\ h_{err},\ \Delta\psi_u,\ \Delta\psi_l,\ n_{ac}]^T, \quad (10)$$

where $h_{err}=h-h_f^*$ is the error between the current altitude and the desired altitude; $\Delta\psi_u$ is the difference between the heading angle and the upper boundary value of the heading angle; $\Delta\psi_l$ is the difference between the heading angle and the lower boundary value of the heading angle; n_{ac} is the actual aerodynamic overload. Detailed expressions for each variable are provided in Section S2 of the ESM.

By incorporating the actual aerodynamic overload into the observation vector, the agent is provided with essential inputs that aid in the identification of aerodynamic errors. This augmentation improves the adaptability of the agent to environmental disturbances. In addition, all elements of the observation vector are normalized to the range of $[-1, 1]$ in order to enhance the training stability and convergence speed of the reinforcement learning algorithm.

It is important to note that we assume the HGV is equipped with an onboard navigation system that provides precise real-time estimates of the flight state, including its position, velocity, and altitude. In this way, the observation vector can be calculated based on the current state of the HGV. While the navigation system uncertainties were not explicitly modeled, the robustness of the guidance policy to such errors is indirectly validated through extensive Monte Carlo (MC) simulations, as described in Section 4.

Next, based on the three-DOF dynamic equations of the HGV (as described in Eqs. (1a)–(1f)), the action space $\boldsymbol{a}_t$ is defined by the change rates of the attack angle and bank angle of the HGV, that is:

$$\boldsymbol{a}_t=[\dot{\alpha},\ \dot{\sigma}]^T. \quad (11)$$

Choosing the $\dot{\alpha}$ and $\dot{\sigma}$ values as actions was done because the inputs to vehicle control systems are typically command body rates (Chai et al., 2021). This ensures easy integration with current vehicle attitude control systems.

3.2.2 Reward function design

In reinforcement learning, the design of the reward function is essential for accelerating training convergence. To ensure that the HGV accurately tracks the reference flight profile and consistently meets the path constraints, the reward function r_t is formulated as follows:

$$r_t=r_{shaping}+r_{alt}+r_{ctrl}+r_{proc}+r_{rev}+r_{bonus}, \quad (12)$$

where $r_{shaping}$ is the shaping reward; r_{alt} is the altitude error reward signal; r_{ctrl} is the penalty for the control effort; r_{proc} is the path constraint penalty signal; r_{rev} is the penalty for bank angle reversal; r_{bonus} represents the terminal bonus reward.

3.2.3 Policy optimization

The proximal policy optimization (PPO) algorithm (Schulman et al., 2017) is among the most prominent methods in DRL, achieving outstanding performance on various tasks. Therefore, the PPO algorithm is employed to optimize the agent policy in this study.

Furthermore, given the effectiveness of recurrent neural networks (RNNs) in managing long-term dependencies and sequence processing (Chung et al., 2015), we integrate an RNN layer within both the policy and value networks. This approach aims to make the policy network more adaptable to various tasks by updating the RNN's hidden states (Gaudet et al., 2022). In this paper, the RNN layers are implemented using gated recurrent units (GRU) (Chung et al., 2015).

In order to motivate the agents to learn to actively satisfy the path constraints, the positive reward flow is ceased by terminating the current episode early in the training phase, when path constraints are violated. Through this design and penalty signals for path constraints, the agents are encouraged to actively satisfy the path constraints. This conclusion will be validated with numerical simulations in Section 4.

In particular, the value of $\delta\psi_0$ in Eq. (7) is randomly generated at the start of each episode during

the training phase. This ensures that the agent is well-trained and can adapt to the varying heading error corridors.

3.3 Terminal time control guidance method

3.3.1 Remaining flight time prediction

Taking the derivative of s with respect to time t yields:

$$\frac{\mathrm{d}s}{\mathrm{d}t}=-\frac{v\cos\gamma\cos\Delta\psi}{r}, \tag{13}$$

$$\mathrm{d}t=-\frac{r}{v\cos\gamma\cos\Delta\psi}\mathrm{d}s. \tag{14}$$

Integrating both sides, we obtain:

$$t_{\mathrm{go}}=t_{\mathrm{f}}-t=\int_{s_t^*}^{s}\frac{r}{v\cos\gamma\cos\Delta\psi}\mathrm{d}s, \tag{15}$$

where t_{go} is the remaining flight time. From the right-hand side of the above equation, one can see that r, v, and γ are longitudinal flight parameters, whereas $\Delta\psi$ is a lateral parameter.

Then, based on Eq. (1a), the differential of r with respect to s can be obtained as follows:

$$\frac{\mathrm{d}r}{\mathrm{d}s}=\frac{\mathrm{d}r}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s}=-v\sin\gamma\frac{r}{v\cos\gamma\cos\Delta\psi}=-\frac{r\tan\gamma}{\cos\Delta\psi}. \tag{16}$$

Assuming that the well-trained agent can perfectly track the reference flight profile, which was designed in Section 3.1.1, we substitute Eqs. (5) and (6) into Eqs. (15) and (16) and consider $\Delta\psi$ as a constant. Consequently, the estimated remaining flight time $\hat{t}_{\mathrm{go}}$ can be obtained through numerical integration of Eqs. (15) and (16).

Therefore, the remaining flight time t_{go} can be expressed in the following form:

$$t_{\mathrm{go}}=\hat{t}_{\mathrm{go}}+\delta t_{\mathrm{go}}, \tag{17}$$

where δt_{go} is the residual error between the estimated and actual remaining flight times.

From the preceding analysis, it is clear that the residual error δt_{go} mainly consists of two components: one is the tracking error between the actual flight trajectory and the reference flight profile; the other is the estimation error caused by assuming a constant heading error. Given the difficulty of obtaining an accurate value of the residual error δt_{go}, a DNN is introduced to predict this error.

We denote the forward computation of the residual error prediction network as a mapping function φ, where the estimated residual error $\widehat{\delta t}_{\mathrm{go}}$ can be expressed as:

$$\widehat{\delta t}_{\mathrm{go}}=\varphi(\boldsymbol{z}_t), \tag{18}$$

where $\varphi(\boldsymbol{z}_t)$ is the mapping function of $\boldsymbol{z}_t$; $\boldsymbol{z}_t$ is the input vector of the residual error prediction network, whose expression is as follows:

$$\boldsymbol{z}_t=[s,\ \dot{s},\ v,\ h_{\mathrm{err}},\ \gamma,\ \Delta\psi,\ \delta\psi_0]^{\mathrm{T}}. \tag{19}$$

Finally, the prediction of the remaining flight time $t_{\mathrm{go}}^{\mathrm{p}}$ is determined by:

$$t_{\mathrm{go}}^{\mathrm{p}}=\hat{t}_{\mathrm{go}}+\widehat{\delta t}_{\mathrm{go}}=\hat{t}_{\mathrm{go}}+\varphi(\boldsymbol{z}_t). \tag{20}$$

Notably, the training data for the mapping function φ is generated by sampling trajectories from the agent that was optimized according to the material in Section 3.2. The initial conditions and $\delta\psi_0$ were randomly generated within specified ranges at the beginning of each trajectory.

3.3.2 Online correction of the remaining flight time

As shown in Eq. (15), the remaining flight time t_{go} is influenced by both longitudinal and lateral guidance. This leads to two approaches for flight time control: one based on longitudinal guidance (Guo et al., 2020; Wang HN et al., 2022; Yang HD et al., 2024; Liu et al., 2025) and the other on lateral guidance (Li et al., 2019; Liang et al., 2023).

The proposed approach utilizes lateral guidance for flight time control for two principal reasons. First of all, time control within the longitudinal channel, as illustrated in Eq. (15), involves the coupled parameters r, γ, and v, which complicates selecting a single input parameter for precise regulation of the flight time. Second, Eq. (15) demonstrates that the remaining flight time t_{go} increases monotonically with respect to $|\Delta\psi|$, and thus precise flight time control can be achieved through adjustments to $\Delta\psi$.

Next, Eq. (20) can be reformulated as follows:

$$t_{\mathrm{go}}^{\mathrm{p}}=\hat{t}_{\mathrm{go}}+\varphi(\tilde{\boldsymbol{z}}_t,\ \delta\psi_0), \tag{21}$$

where $\varphi(\tilde{\boldsymbol{z}}_t,\ \delta\psi_0)$ is the mapping function of $\tilde{\boldsymbol{z}}_t$ and $\delta\psi_0$; $\tilde{\boldsymbol{z}}_t=[s,\ \dot{s},\ h_{\mathrm{err}},\ \gamma,\ \Delta\psi]^{\mathrm{T}}$ is the state vector that can be reduced from the current state of the HGV at time t.

Then the desired remaining flight time t_{go}^{d} can be calculated as follows:

$$t_{go}^{d} = t_{f}^{*} - t. \tag{22}$$

To minimize the discrepancy between the desired remaining flight time t_{go}^{d} and the predicted remaining flight time t_{go}^{p}, the Newton–Raphson method is employed to iteratively solve for $\delta\psi_0$, that is:

$$f(\delta\psi_0) = \frac{1}{2}(t_{go}^{d} - t_{go}^{p})^2, \tag{23a}$$

$$\delta\psi_0^{(k+1)} = \delta\psi_0^{(k)} - \lambda_k \frac{f(\delta\psi_0^{(k)})}{\frac{\partial f(\delta\psi_0^{(k)})}{\partial \delta\psi_0^{(k)}}}, \tag{23b}$$

where k is the iteration number; $\delta\psi_0^{(k)}$ and $\delta\psi_0^{(k+1)}$ are the values of $\delta\psi_0$ at the kth and (k+1)th iterations, respectively; $f(\delta\psi_0^{(k)})$ is the objective function value at the kth iteration; λ_k is the step-size parameter, and is set to $1/2^k$.

Upon updating $\delta\psi_0$ using Eqs. (23a) and (23b), $\delta\psi_{thr}$ is computed from Eq. (9). Next, $\Delta\psi_l$ and $\Delta\psi_u$ are correspondingly revised using Eq. (S6) of the ESM. This adjustment modifies the observation vector as defined in Eq. (10), thereby influencing the HGV's flight trajectory and enabling precise adjustment of the terminal time.

To reduce the computational cost of the online real-time guidance, we adopted an event-triggered mechanism M_{on} for correcting the remaining flight time. Specifically, the correction of the remaining flight time according to Eqs. (23a) and (23b) is only executed when the difference between the predicted remaining flight time and the desired remaining flight time exceeds a predefined threshold:

$$M_{on} = \begin{cases} 0, & |t_{go}^{d} - t_{go}^{p}| \leqslant \delta t_{thr}, \\ 1, & |t_{go}^{d} - t_{go}^{p}| > \delta t_{thr}, \end{cases} \tag{24}$$

where M_{on}=1 indicates that a correction of $\delta\psi_0$ is required in the current guidance cycle, and M_{on}=0 signifies that no modification to $\delta\psi_0$ is needed. δt_{thr}>0 is the terminal time error threshold.

To ensure the accuracy of terminal time control while minimizing the real-time computational load onboard, the terminal time error threshold δt_{thr} is designed as a linear function of velocity:

$$\delta t_{thr}(v) = \frac{v - v_f}{v_0 - v_f}(\delta t_0 - \delta t_f) + \delta t_f, \tag{25}$$

where δt_0 is a positive constant.

The rationale behind this linear design is to create a guidance strategy that intelligently allocates computational resources while ensuring high terminal precision. In the early and mid-phases of the entry flight, when the velocity is high, the terminal time error threshold δt_{thr} is larger. This allows for greater tolerance in the predicted terminal time error, thereby reducing the frequency of computationally intensive corrections. As the vehicle decelerates and approaches the terminal area, the threshold linearly decreases, forcing the system to make more frequent and detailed adjustments to achieve the required terminal accuracy.

3.3.3 Feasible terminal time range

For a specific initial condition and desired terminal states, there necessarily exist two flight trajectories corresponding to the minimum and maximum terminal times ($\tilde{t}_f^{min}$ and $\tilde{t}_f^{max}$), for which detailed expressions can be found in Section S2 of the ESM.

For time coordinated entry flight missions involving multiple HGVs, the feasible desired terminal time range for the ith HGV is $[\tilde{t}_{f,i}^{min}, \tilde{t}_{f,i}^{max}]$ (i=1, 2, ⋯, M, and M is the number of HGVs). Then, the coordinated time adjustment range can be obtained as follows:

$$t_f^* \in [T_{min}, T_{max}], \tag{26a}$$

$$T_{min} = \max\{\tilde{t}_{f,1}^{min}, \tilde{t}_{f,2}^{min}, \cdots, \tilde{t}_{f,M}^{min}\}, \tag{26b}$$

$$T_{max} = \min\{\tilde{t}_{f,1}^{max}, \tilde{t}_{f,2}^{max}, \cdots, \tilde{t}_{f,M}^{max}\}, \tag{26c}$$

where T_{min} and T_{max} are the minimum and maximum values of the desired terminal time for multi-HGV, respectively. Subsequently, each HGV can accomplish the coordinated entry flight mission by utilizing the time-control guidance algorithm described in Section 3.3.2.

The proposed guidance method involves the following steps:

Step 1: Given the initial conditions and target state, determine the coordinated flight time using Eqs. (26a)–(26c).

Step 2: Estimate the terminal time error and evaluate Eq. (24). If $M_{on}=1$, then execute Step 3; otherwise, execute Step 4.

Step 3: Calculate $\delta\psi_0$ iteratively using Eqs. (23a) and (23b), and update the heading error threshold as described in Eq. (9).

Step 4: Combine the current state of the HGV and Eq. (9), calculate the observation vector, and input it into the well-trained agent as described in Section 3.2 for forward computation of the guidance command.

Step 5: Check if the HGV has reached the terminal interface. If it has, the guidance algorithm ends; if not, go back to Step 2.

4 Experiments and results

In the simulations, the common aero vehicle-hypersonic (CAV-H) (Phillips, 2003) vehicle was adopted to show the effectiveness and robustness of our entry guidance method. The vehicle mass was 907.2 kg. The complete parameters of CAV-H are provided in Table 1.

Table 1 CAV-H parameters

Parameter	Value
Reference area, A_{ref} (m^2)	0.4837
Guidance cycle (s)	5.0
Attack angle, α (°)	[0, 25]
Bank angle, σ (°)	[−85, 85]
Change rate of attack angle, $\dot{\alpha}$ ((°)/s)	[−5, 5]
Change rate of bank angle, $\dot{\sigma}$ ((°)/s)	[−40, 40]
Maximum heating rate, $\dot{Q}_{max}$ (kW/m^2)	2500
Maximum dynamic pressure, q_{max} (kPa)	150
Maximum aerodynamic overload, n_{max} ($\times g_0$)	3.0
Aerodynamic lift coefficient deviation, ΔC_L (%)	[−5, 5]
Aerodynamic drag coefficient deviation, ΔC_D (%)	[−5, 5]
Atmospheric density deviation, $\Delta\rho$ (%)	[−5, 5]

The DRL tracking agent and DNN time-to-go prediction network employed in this paper were fully trained offline. Detailed training environment settings, network architectures, PPO hyperparameters, and training results of both components are provided in Section S3 of the ESM.

4.1 Results for reference profile tracking agent

To evaluate the adaptability of the well-trained agent under different $\delta\psi_0$ values, as well as to verify the rationality of the conclusions drawn in Section 3.3.1, we perform simulations, setting the following HGV conditions at the initial interface: the initial altitude h_0= 80 km, initial longitude λ_0=22°, initial latitude ϕ_0=10°, v_0=6800 m/s, initial flight-path angle γ_0=−0.5°, initial heading error $\Delta\psi_0$=1°, initial attack angle α_0=20°, and initial bank angle σ_0=0°. The desired terminal state is set to λ_f^*=112°, ϕ_f^*=15°, s_f^*=95 km, v_f^*=2000 m/s, h_f^*=28 km, and $\delta\psi_f$=5°. It should be noted that the initial and target positions were set outside the range that was adopted in the training. This was done to simultaneously test the generalization performance of the trained agent. The simulation results are shown in Figs. 3 and 4.

As can be observed in Figs. 3a and 3d, the smaller the value of $\delta\psi_0$, the narrower the corresponding heading error corridor, and the closer the flight trajectory is to the great-circle arc connecting the initial position and the target position. Conversely, the larger the value of $\delta\psi_0$, the greater the lateral maneuvering range of the corresponding flight trajectory. Moreover, as can be seen from Figs. 3b and 3c, the well-trained agent is capable of controlling the HGV to track the designed reference flight profile. Specifically, the reference velocity tracking error converges to within 40 m/s after approximately 1600 s of flight, while the reference flight-path angle tracking error remains within 0.3° after 500 s of flight.

Fig. 4 depicts the path constraint variation curves corresponding to different trajectories. It can be observed that the path constraints for all trajectories remain within the prescribed limits. This further confirms that the well-trained agent can satisfy the path constraints.

Furthermore, inspecting Table 2, the terminal state errors of all trajectories remain within the allowable range. Additionally, as the value of $\delta\psi_0$ increases, the flight times of the trajectories gradually increase. This validates the rationality of the conclusions proposed in Section 3.3.1. It is worth noting that these simulations used the remaining range as the termination condition.

In addition, to evaluate the robustness of the well-trained agent, we conducted 500 MC simulations under more stringent conditions to analyze the guidance performance of the well-trained agent. Specifically, the aerodynamic parameter error range was set to [−15%, 15%], and the atmospheric density error range was set to [−10%, 10%]. The well-trained agent is used to generate the guidance command in real time.

The MC simulation results are presented in Figs. 5 and 6. Fig. 5 shows the histogram of the terminal range-to-go errors for the trajectories, which are all less than 50 m. Meanwhile, Fig. 6 depicts the distribution

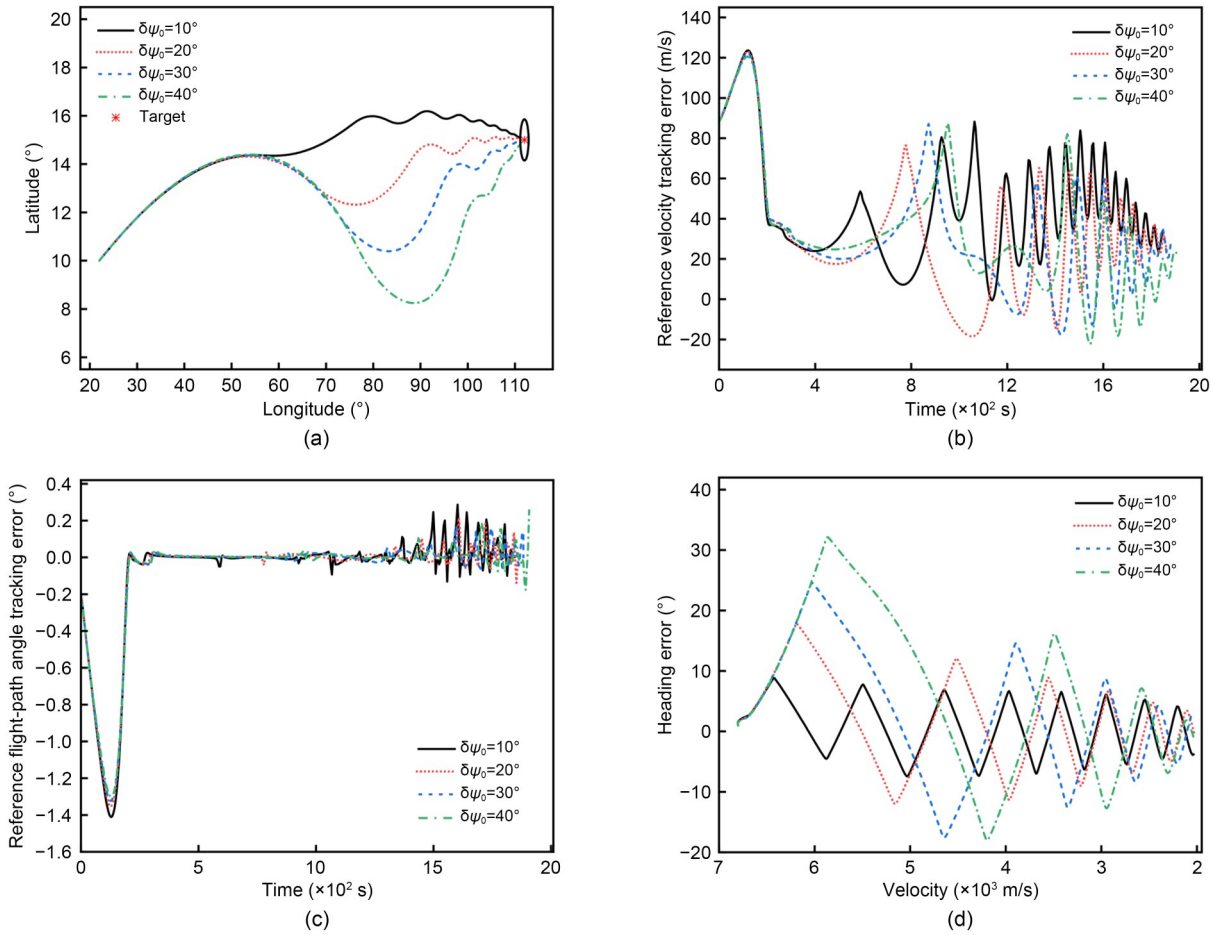


Fig. 3 Comparison of the simulation results across different $\delta\psi_0$ values: (a) ground tracking; (b) reference velocity tracking error; (c) reference flight-path angle tracking error; (d) heading error

of the terminal velocity and altitude errors for the trajectories. Notably, the terminal velocity errors of all trajectories are less than 50 m/s, while the terminal altitude errors lie within the interval [−90 m, 80 m]. These statistical results demonstrate that the proposed DRL-based guidance method achieves high guidance accuracy. Moreover, it is robust even in the presence of significant aerodynamic parameter uncertainties and environmental disturbances.

4.2 Results for the remaining flight time prediction

To evaluate the accuracy and adaptability of the remaining flight time prediction method and the residual prediction network, we utilized trajectory data generated under different values of $\delta\psi_0$ from Section 4.1 as test data. The test results are shown in Fig. 7. As one can observe, the estimation error for the remaining flight time remains within 2 s for most of the flight time. Moreover, the estimation error converges to within 1 s during the final 500 s of the flight. Furthermore, it is worth noting that when $\delta\psi_0$=40°, the relatively wide heading error corridor during the early phase of the flight can lead to larger heading errors, which in turn result in a higher estimation error. However, as the heading error corridor progressively narrows, the corresponding estimation error for the remaining flight time also gradually decreases.

4.3 Results for the time control entry guidance method

4.3.1 Single HGV time-control guidance results

The effectiveness of the proposed TCEG method and the generalization performance of the trained agent from Section 4.1 were validated through simulations under nominal conditions. The initial conditions were set as in Section 4.1. By setting $\delta\psi_0$ to a minimum of 5° and a maximum of 45°, the corresponding feasible

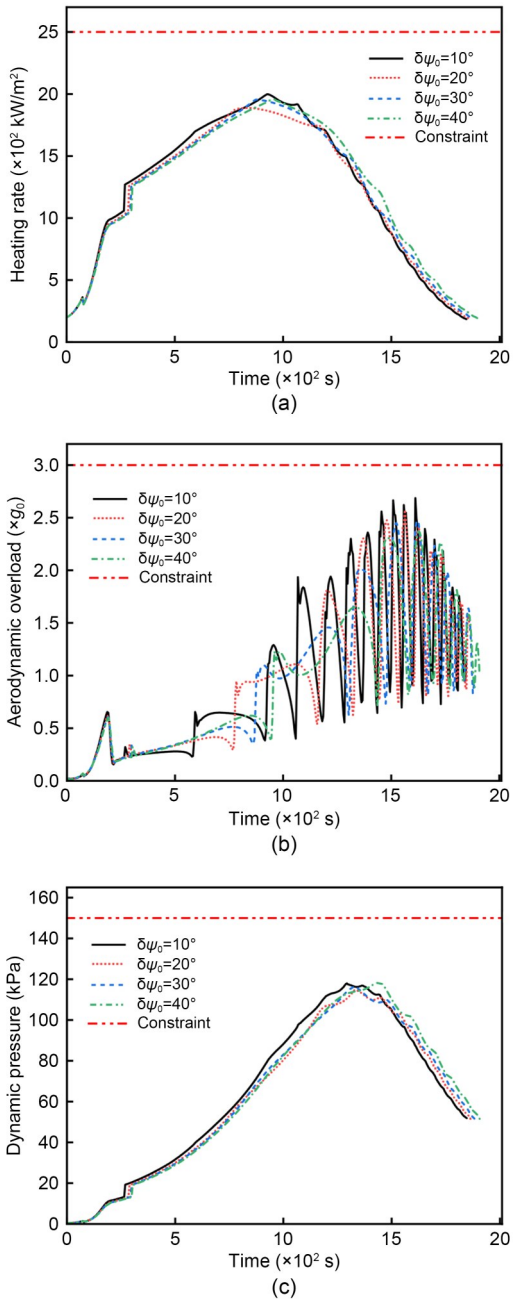


Fig. 4 Path constraint curves for different $\delta\psi_0$ values: (a) heating rate; (b) aerodynamic overload; (c) dynamic pressure

Table 2 Terminal state errors for different $\delta\psi_0$ values

$\delta\psi_0$ (°)	Δs_f (m)	Δh_f (m)	Δv_f (m/s)	$\Delta\psi_f$ (°)	Δt_f (s)
10	−13.624	−0.103	33.773	−3.792	1846.84
20	−11.694	−1.642	23.751	−0.572	1864.98
30	−2.902	5.612	29.245	2.956	1882.66
40	−26.735	6.927	24.123	−1.535	1907.96

Δs_f, Δh_f, Δv_f, $\Delta\psi_f$, and Δt_f are the range-to-go, altitude, velocity, heading, and terminal time errors, respectively

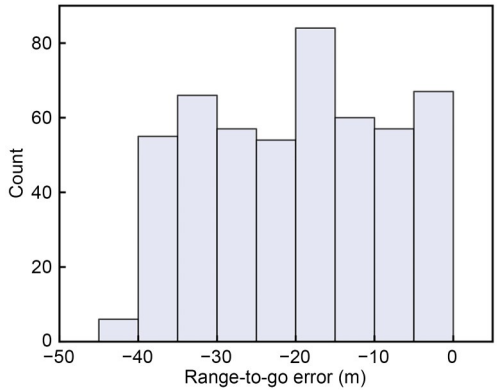


Fig. 5 Histogram of terminal range-to-go errors

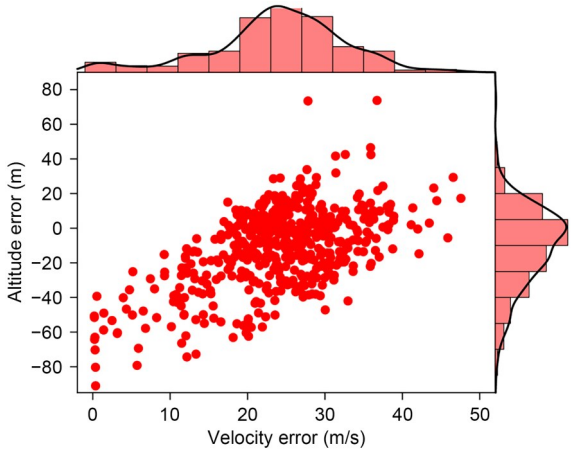


Fig. 6 Distribution of the terminal velocity and altitude errors

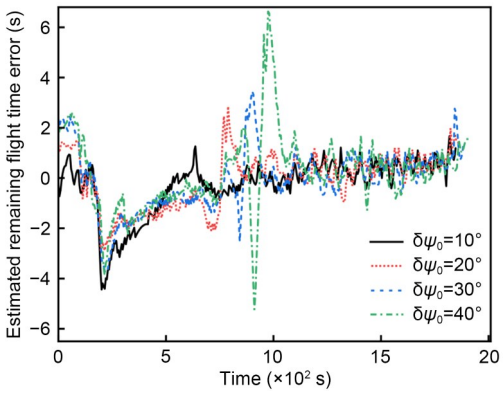


Fig. 7 Estimation error for the remaining flight time under different $\delta\psi_0$ values

range of terminal times is [1838 s, 1927 s]. As the safety factor k_t=0.25, the feasible range of terminal times is narrowed to [1860 s, 1904 s]. Figs. 8 and 9 present the simulation results under different terminal time constraints. Fig. 8a illustrates the ground tracking, from which it can be observed that the shorter the

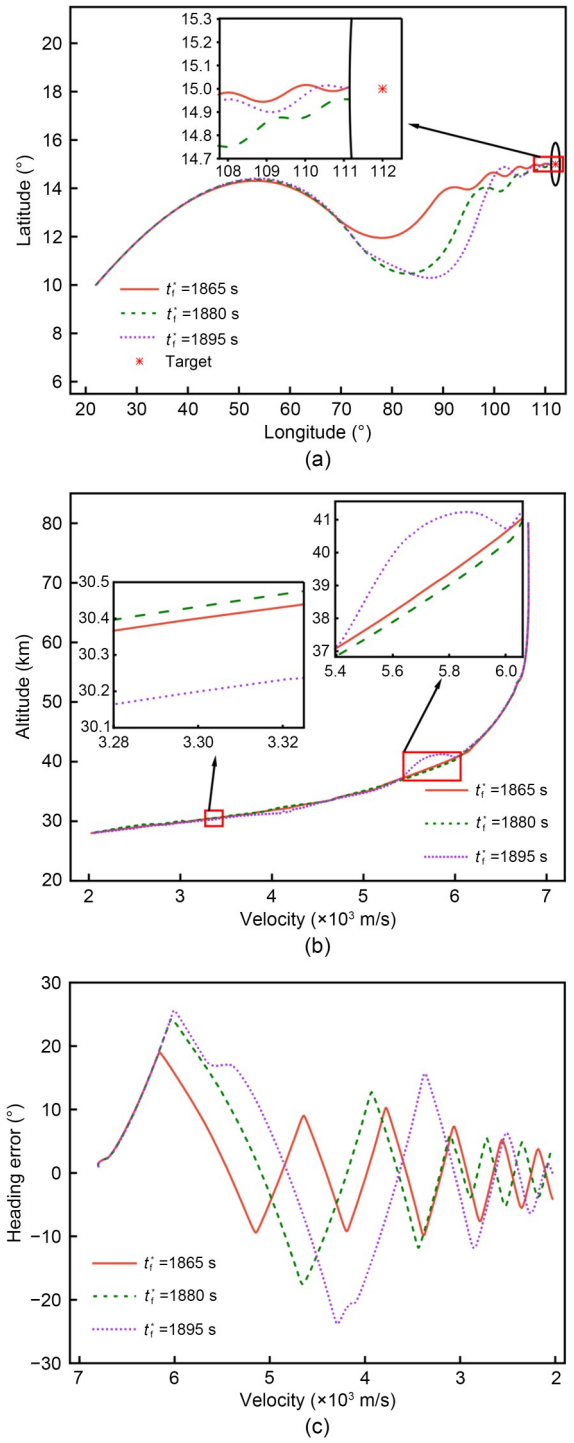


Fig. 8 Simulation results of flight states for different desired terminal times: (a) ground tracking; (b) altitude; (c) heading error

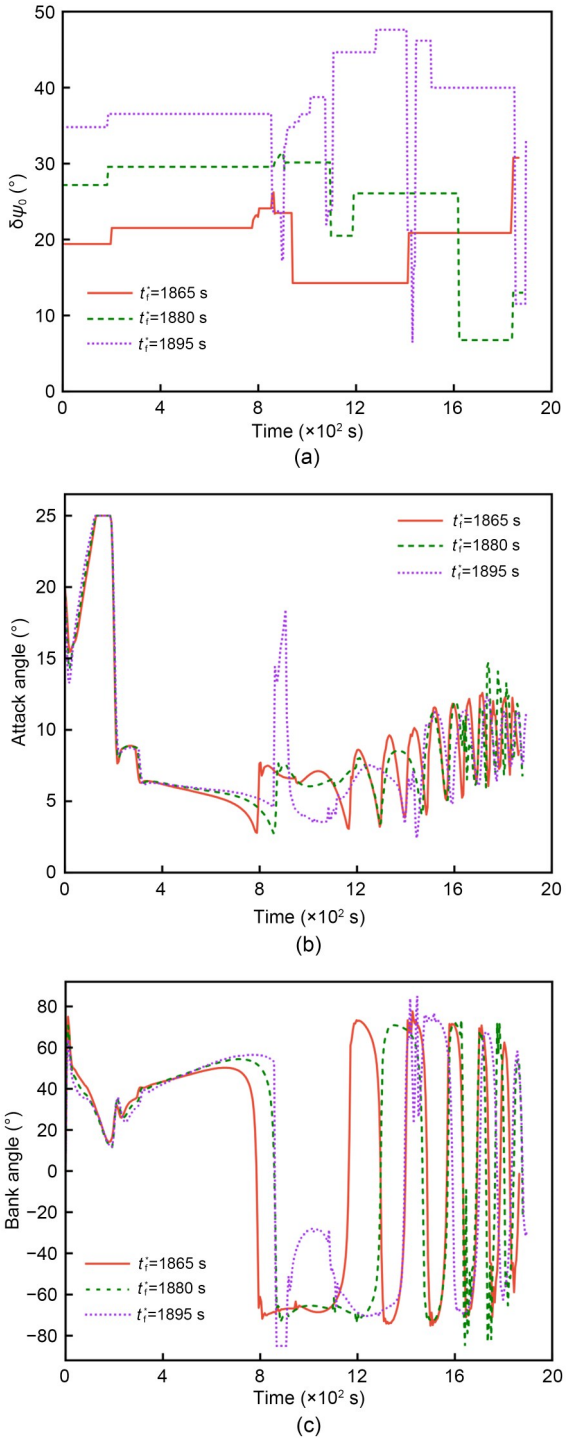


Fig. 9 Simulation results of control variables for different desired terminal times: (a) $\delta\psi_0$; (b) attack angle; (c) bank angle

terminal time, the smaller the lateral maneuvering range of the corresponding flight trajectory. The variation curve of $\delta\psi_0$ is shown in Fig. 9a. Due to the presence of remaining flight time prediction errors, continuous correction of $\delta\psi_0$ is required. Figs. 9b and 9c depict the variation curves of α and σ, respectively. It can be seen that when the sign of the bank angle changes, the attack angle varies accordingly. This is because, as the

bank angle crosses 0°, the component of the lift in the longitudinal plane $L\cos\sigma$ increases. Therefore, to maintain a constant magnitude of the lift component in the longitudinal plane, the attack angle command is correspondingly reduced.

Table 3 presents the terminal state errors under different terminal times. It can be observed that all terminal errors remain within the allowable range, thereby validating the effectiveness of the proposed method.

Table 3 Terminal state errors for different desired terminal times

t_f^* (s)	Δs_f (m)	Δh_f (m)	Δv_f (m/s)	$\Delta\psi_f$ (°)	Δt_f (s)
1865	−28.882	−6.537	28.583	−4.097	−0.26
1880	−4.895	30.281	32.748	3.603	−0.58
1895	−10.831	−26.738	23.415	−0.261	−0.22

4.3.2 Comparison results

To verify the comparative advantages of the proposed method in terms of guidance accuracy and real-time performance, the terminal TCEG methods based on an analytical predictor–corrector (Guo et al., 2020) and a numerical predictor–corrector (Li et al., 2019) are also tested. The desired terminal time is set to 1880 s, and the simulation results are shown in Fig. 10. It is noticeable that the terminal velocity error of the analytical predictor–corrector-based method is significantly larger than that of the other methods, which is attributed to the relaxation of the terminal velocity constraint that is applied in this method. Figs. 10c and 10d respectively depict the variation curves of the attack angle and the bank angle for the different methods.

Fig. 10a shows the simulation results of ground tracking for the proposed method and the two other methods. It is clear that all methods can successfully guide the HGV to the TAEM interface. The simulation results of altitude–velocity profiles for the different methods are presented in Fig. 10b; it can be seen that the altitude variations of the analytical predictor–corrector

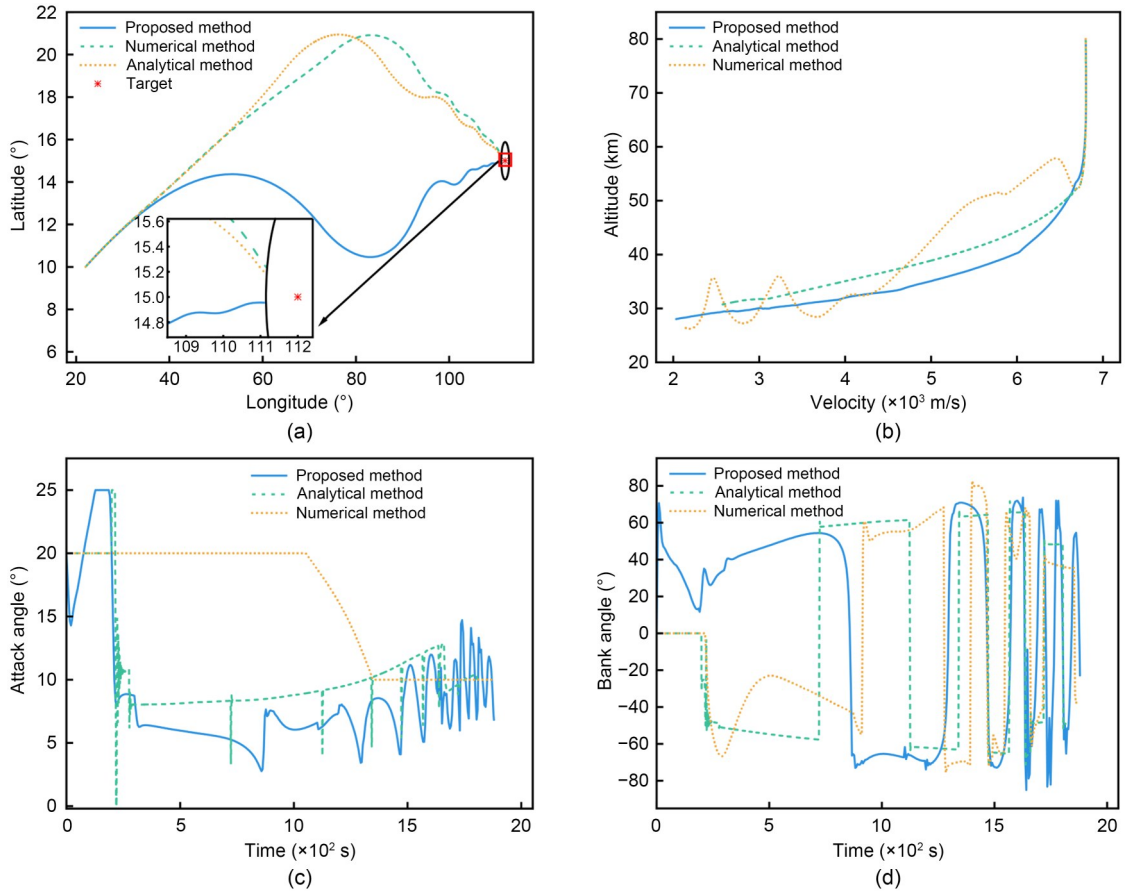


Fig. 10 Comparison of the simulation results for different methods: (a) ground tracking; (b) altitude; (c) attack angle; (d) bank angle

method and the proposed method are minimal (showing smooth profiles), while the numerical predictor–corrector guidance method exhibits trajectory jumps.

Tables 4 and 5 present the statistical results of the command computation time and terminal states for different methods, respectively. Although the analytical predictor–corrector-based method has the lowest command computation time, its corresponding terminal state error is significantly larger than that of the other two methods. In contrast, while the numerical predictor–corrector guidance method achieves smaller terminal state errors, its command computation time increases significantly. However, the proposed method achieves the highest terminal state accuracy and also shows no significant increase in command computation time compared with the analytical method; thus, it maintains excellent real-time performance.

Table 4 Statistical results of the command computation time for the different methods

Method	Command computation time (s)		
	Mean	Minimum	Maximum
Proposed method	0.055	0.019	0.478
Analytical method	0.003	0.001	0.021
Numerical method	0.917	0.073	5.050

Table 5 Comparison of the terminal state errors for the simulation results

Method	Δs_f (m)	Δh_f (m)	Δv_f (m/s)	$\Delta\psi_f$ (°)	Δt_f (s)
Proposed method	−4.895	30.281	32.748	3.603	−0.580
Analytical method	−33.404	2744.676	573.004	6.956	−64.580
Numerical method	−14.038	−1412.726	118.382	5.747	−7.540

4.3.3 Multi-HGV time coordination simulation results

To further validate the effectiveness and robustness of the proposed TCEG method, a multi-HGV time-coordinated entry mission was simulated. The initial conditions of the mission are shown in Table 6, with the target location specified as: ϕ_f^*=25° and λ_f^*=25°.

By setting $\delta\psi_0$ to a minimum value of 5° and a maximum value of 45°, as well as the safety factor k_t=0.25, the feasible ranges of terminal times for all HGVs are as follows: [1870 s, 1917 s], [1877 s, 1924 s], and [1868 s, 1914 s].

Table 6 Initial conditions for the multi-HGV simulation

Vehicle	h_0 (km)	λ_0 (°)	ϕ_0 (°)	v_0 (m/s)	γ_0 (°)	$\Delta\psi_0$ (°)
HGV-1	79.9	5.0	12.0	6770	−0.15	−0.05
HGV-2	80.1	8.0	5.0	6760	0.15	0.25
HGV-3	80.2	3.0	15.0	6810	−0.25	−0.10

HGV-i (i=1, 2, 3) is the i-th HGV

Then, based on Eqs. (26a)–(26c), we can obtain the feasible range for the coordinated terminal time as [1877 s, 1914 s]. Therefore, the coordination time is t_f^*=1890 s.

The simulation results under nominal conditions are shown in Figs. 11–13. Fig. 11 presents the three-dimensional flight trajectories of the multi-HGV. It can be observed that each HGV is able to fly smoothly to the TAEM interface. Fig. 12a illustrates the simulation results of the ground tracking, and Fig. 12b shows the resulting altitude–velocity profiles of the three HGVs; it can be clearly observed that all three HGVs transition smoothly from their initial states to the target states. Fig. 12c depicts the variation curves of the heading error, and Fig. 13a shows the correction results of $\delta\psi_0$ for the three HGVs. These results validate the effectiveness of the proposed TCEG method. The variation curves of α and σ are illustrated in Figs. 13b and 13c, respectively. One can see that the control inputs for each HGV vary smoothly, fully satisfying the constraints of the flight control system. The heading error and the variation curve of the heading error corridor for HGV-3 are presented in Fig. 14. It can be observed

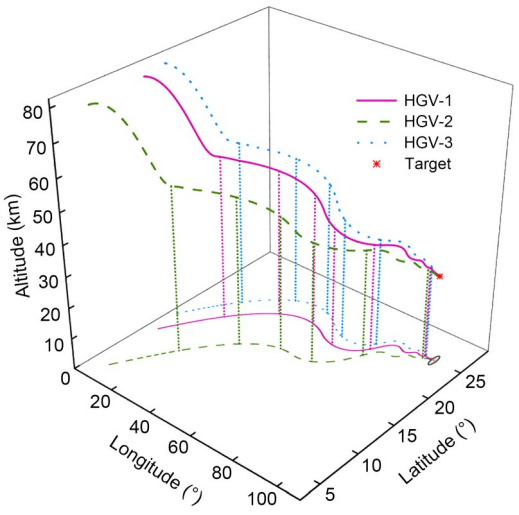
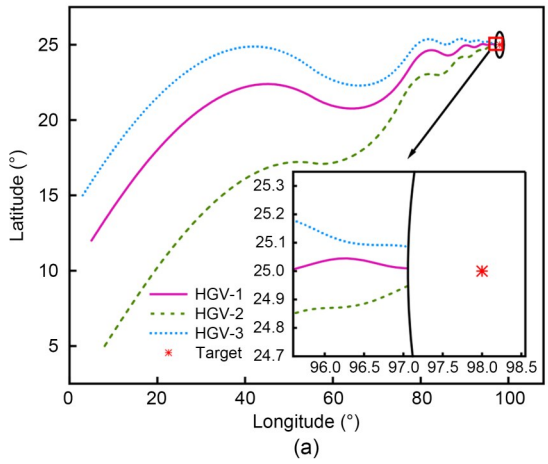
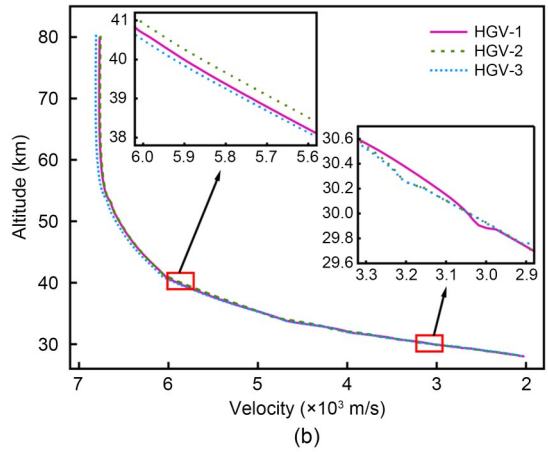


Fig. 11 Three-dimensional flight trajectories for the multi-HGV time coordination





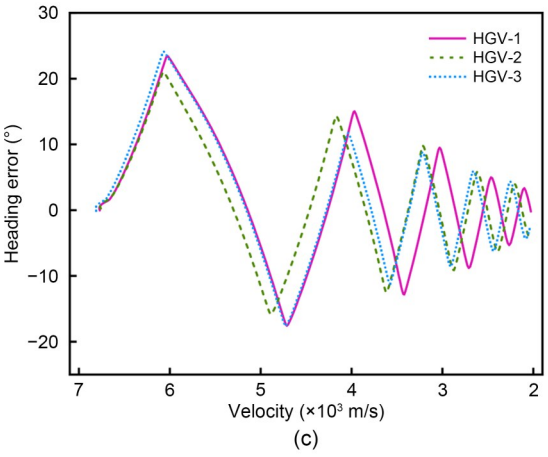
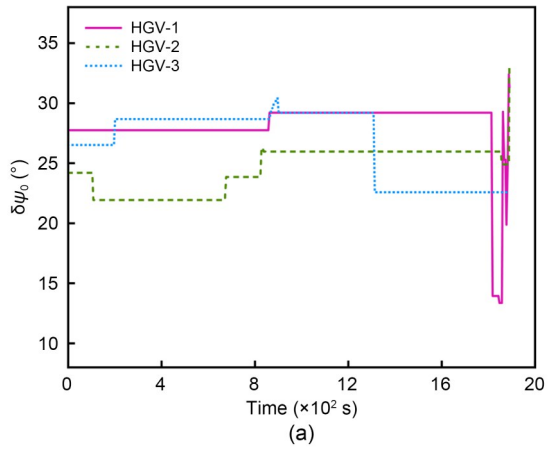
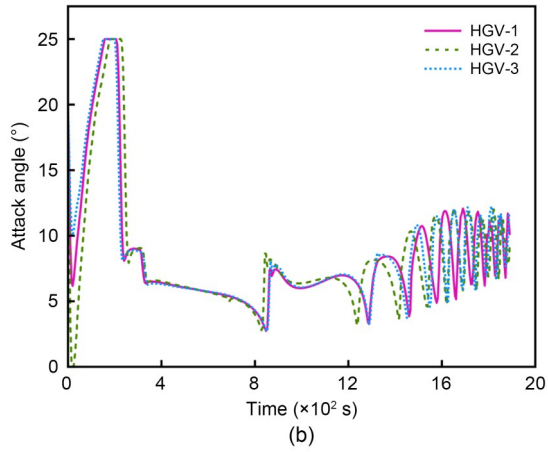


Fig. 12 Simulation results of flight states for multi-HGV time coordination: (a) ground tracking; (b) altitude; (c) heading error





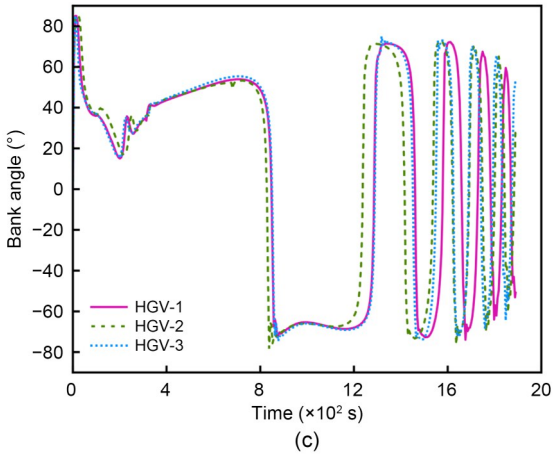


Fig. 13 Simulation results of control variables for multi-HGV time coordination: (a) $\delta\psi_0$; (b) attack angle; (c) bank angle

that the proposed method, which is based on adjusting the threshold of the heading error, is highly effective. Additionally, it is evident that the adoption of the event-triggered mechanism successfully avoids frequent corrections of the heading error corridor.

Furthermore, Table 7 presents the terminal state errors of the HGVs under nominal conditions. It is evident that the terminal state errors for each HGV remain within the acceptable range, with all terminal errors being less than 1 s. This indicates that the proposed

algorithm can achieve highly precise entry guidance under terminal time constraints.

To validate the robustness of the proposed method, 500 MC simulations were conducted under dispersed cases. The ranges of perturbations for various parameters are shown in Table 8.

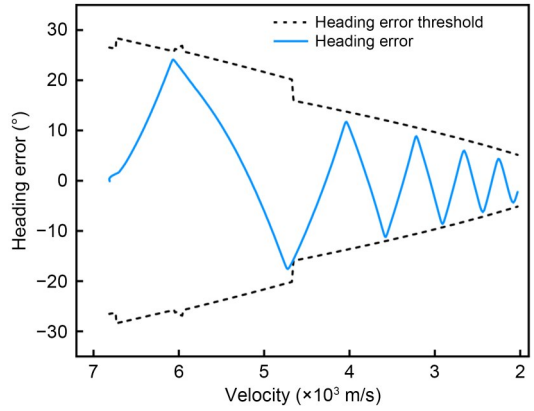


Fig. 14 Heading error and corresponding threshold for HGV-3

Table 7 Terminal state errors for multi-HGV time coordination in the nominal case

Vehicle	Δs_f (m)	Δh_f (m)	Δv_f (m/s)	$\Delta \psi_f$ (°)	Δt_f (s)
HGV-1	−10.232	17.006	26.868	−0.207	0.12
HGV-2	−32.533	−2.259	30.985	−3.971	−0.06
HGV-3	−2.382	8.261	29.964	−2.255	−0.40

Table 8 Parameter dispersion settings

Parameter	Deviation range	Parameter	Deviation range
h_0	[−1 km, 1 km]	v_0	[−30 m/s, 30 m/s]
λ_0	[−0.5°, 0.5°]	ϕ_0	[−0.5°, 0.5°]
γ_0	[−0.1°, 0.1°]	ψ_0	[−0.3°, 0.3°]
α_0	[−3°, 3°]	σ_0	[−3°, 3°]
C_L	[−10%, 10%]	C_D	[−10%, 10%]
ρ	[−10%, 10%]	m	[−3%, 3%]

C_L and C_D are the aerodynamic lift coefficient and aerodynamic drag coefficient, respectively

The MC simulation results are shown in Figs. 15 and 16. Fig. 15 shows the ground tracking results in the dispersed cases, from which it is clear that all HGVs can fly smoothly to the TAEM interface. Across all trajectories, Fig. 16a illustrates the distribution of terminal velocity and altitude errors, while Fig. 16b presents the statistical results of the terminal time errors. It can be seen that the Δh_f values for all trajectories

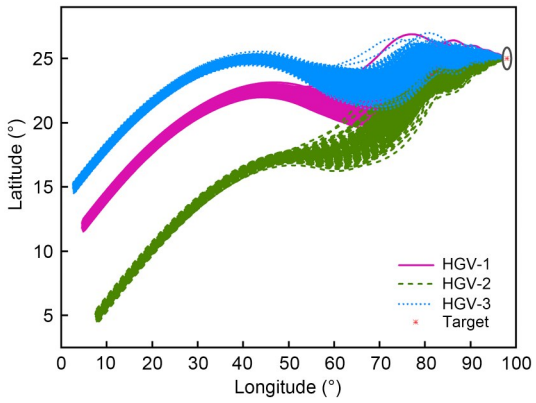


Fig. 15 Ground tracking results of the MC simulation

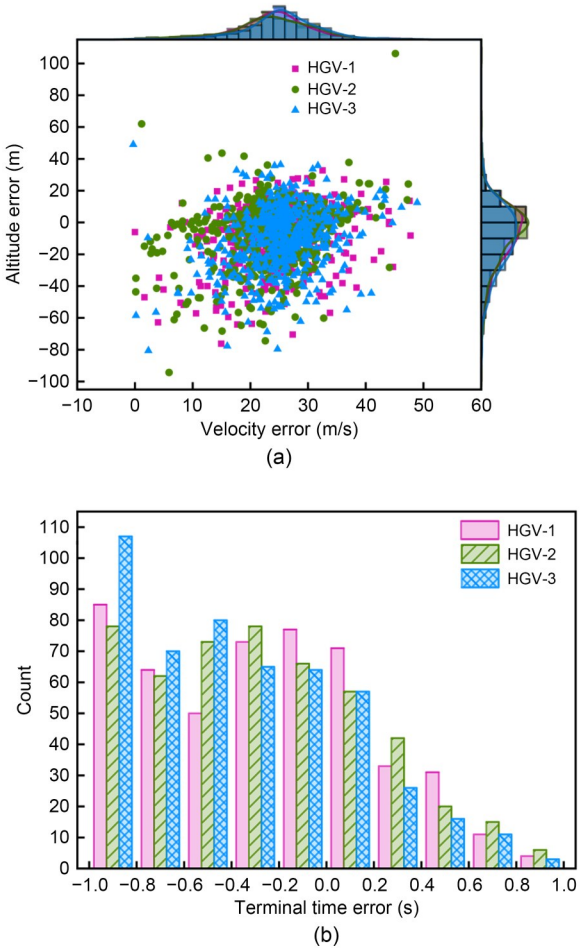


Fig. 16 MC simulation results: (a) terminal velocity and altitude error distributions; (b) histogram of terminal time errors

are within [−100 m, 100 m], the Δv_f are within [0 m/s, 50 m/s], and the Δt_f are within [−1 s, 1 s]. These results demonstrate that the proposed method is robust and maintains high guidance accuracy even under dispersed

Table 9 MC simulation results for a multi-HGV scenario under dispersed cases

Vehicle	Δs_f (m)			Δh_f (m)			Δv_f (m/s)			Δt_f (s)		
	Mean	Std.	Worst	Mean	Std.	Worst	Mean	Std.	Worst	Mean	Std.	Worst
HGV-1	−19.68	19.58	337.40	−8.32	18.52	−76.21	24.35	6.71	47.76	−0.27	0.47	−0.94
HGV-2	−19.55	20.86	368.33	−7.07	20.25	106.17	23.68	7.67	47.37	−0.27	0.48	−1.00
HGV-3	−19.70	16.12	211.33	−10.36	20.13	−80.62	25.53	6.57	48.91	−0.37	0.45	−0.98

Std. is the standard deviation

conditions. Table 9 presents the statistical results of the terminal state errors derived from the MC simulations for the three HGVs. The data clearly show that the proposed TCEG method achieves high precision and reliability, enabling the successful execution of time-coordinated entry missions for multiple HGVs.

While explicit modeling of the measurement errors associated with the HGV's onboard navigation system was not performed in this study, in practice, the parameter perturbations presented in Table 9 can be regarded as the ultimate effects of the navigation system errors. The high guidance accuracy achieved in these simulations indicates that the well-trained agent is highly robust to trajectory perturbations, which can, to a certain extent, verify the robustness of the method to navigation errors.

5 Conclusions

We proposed a TCEG method for HGVs based on DRL. An agent is trained to control the HGV to fly smoothly toward the designated TAEM interface, and the complex process of designing a reference flight profile is avoided. Moreover, generating the guidance commands only requires forward computations by the neural network, thereby satisfying the requirements of high accuracy and real-time performance. A novel method for estimating the remaining flight time was also proposed, which leverages a combination of a DNN and an analytical predictive formula to jointly estimate the remaining flight time. This method significantly improves the accuracy of the remaining flight time estimation. A theoretical investigation into the factors affecting entry flight time was then conducted, leading to the development of a TCEG strategy based on adaptive modulation of the heading error threshold. This method can rapidly estimate the flight time range of each vehicle. Theoretical derivations and simulations confirm that the proposed TCEG method achieves high-precision guidance with robust performance. Additionally, the generation of guidance commands avoids the complexity of trajectory predictions and iterative corrections, thus meeting the requirements for rapid response and practical application.

In terms of practical implementation of the proposed method, the well-trained agent and residual prediction network are well-suited for real-world HGVs. The generation of guidance commands only requires forward computation from the neural network, which is a computationally efficient process. This presents a significant advantage over traditional predictor–corrector methods, which demand substantial computational resources for iterative trajectory prediction and correction. Furthermore, the output of the agent can be directly fed into the HGV's inner-loop attitude control system, ensuring seamless integration with the existing flight control system. However, prior to actual deployment, one must consider how quickly to take over with manual control when safety constraints are violated in order to avoid serious consequences. Additionally, any trained strategy must undergo rigorous validation with high-fidelity hardware-in-the-loop simulations to ensure its reliability and performance across the entire operational range.

Future work will proceed in two main directions. First, a coupled longitudinal–lateral flight time control method could be designed to achieve a wider adjustable range of flight times. This would involve modulating the lateral heading error corridor and adjusting the longitudinal reference profile in real time, allowing the DRL agent to trade potential energy for kinetic energy more dynamically and meet a broader set of terminal time constraints. Second, future work could consider the attitude dynamics of the HGV by establishing a six-DOF dynamic model and developing an integrated guidance and control method based on DRL.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 62103432) and the Open Fund of Key

Laboratory of Cross-Domain Flight Interdisciplinary Technology (No. 2024-KYKF-4004), China.

Author contributions

Zhenyu LIU: investigation, methodology, and writing–original draft. Gang LEI: conceptualization, supervision, and writing–review & editing. Yong XIAN: project administration and writing–review & editing. Leliang REN: validation and data curation. Shaopeng LI: funding acquisition and writing–review & editing. Daqiao ZHANG: writing–review & editing.

Conflict of interest

Zhenyu LIU, Gang LEI, Yong XIAN, Leliang REN, Shaopeng LI, and Daqiao ZHANG declare that they have no conflict of interest.

References

Bao CY, Li XC, Xu WL, et al., 2025. Coordinated reentry guidance with A* and deep reinforcement learning for hypersonic morphing vehicles under multiple no-fly zones. *Aerospace*, 12(7):591. https://doi.org/10.3390/aerospace12070591

Brunner CW, Lu P, 2012. Comparison of fully numerical predictor-corrector and Apollo skip entry guidance algorithms. *The Journal of the Astronautical Sciences*, 59(3): 517-540. https://doi.org/10.1007/s40295-014-0005-1

Chai RQ, Tsourdos A, Savvaris A, et al., 2021. Review of advanced guidance and control algorithms for space/aerospace vehicles. *Progress in Aerospace Sciences*, 122:100696. https://doi.org/10.1016/j.paerosci.2021.100696

Cheng L, Jiang FH, Wang ZB, et al., 2021. Multiconstrained real-time entry guidance using deep neural networks. *IEEE Transactions on Aerospace and Electronic Systems*, 57(1): 325-340. https://doi.org/10.1109/TAES.2020.3015321

Chung J, Gulcehre C, Cho K, et al., 2015. Gated feedback recurrent neural networks. Proceedings of the 32nd International Conference on International Conference on Machine Learning, p.2067-2075. https://proceedings.mlr.press/v37/chung15.html

Gao Y, Zhou R, Chen JY, 2024. Integrated entry guidance with no-fly zone constraint using reinforcement learning and predictor-corrector technique. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 238(7):728-741. https://doi.org/10.1177/09544100241236995

Gaudet B, Drozd K, Furfaro R, 2022. Adaptive approach phase guidance for a hypersonic glider via reinforcement meta learning. AIAA SCITECH 2022 Forum, p.1-19. https://doi.org/10.2514/6.2022-2214

Guo YH, Li X, Zhang HJ, et al., 2020. Entry guidance with terminal time control based on quasi-equilibrium glide condition. *IEEE Transactions on Aerospace and Electronic Systems*, 56(2):887-896. https://doi.org/10.1109/TAES.2019.2921213

Harpold JC, Gavert DE, 1983. Space shuttle entry guidance performance results. *Journal of Guidance, Control, and Dynamics*, 6(6):442-447. https://doi.org/10.2514/3.8523

Hu QL, Cao RH, Han T, et al., 2021. Field-of-view limited guidance with impact angle constraint and feasibility analysis. *Aerospace Science and Technology*, 114:106753. https://doi.org/10.1016/j.ast.2021.106753

Hu YD, Gao CS, Li JL, et al., 2022. A novel adaptive lateral reentry guidance algorithm with complex distributed no-fly zones constraints. *Chinese Journal of Aeronautics*, 35(7): 128-143. https://doi.org/10.1016/j.cja.2021.06.016

Kim HG, Lee JY, Kim HJ, et al., 2020. Look-angle-shaping guidance law for impact angle and time control with field-of-view constraint. *IEEE Transactions on Aerospace and Electronic Systems*, 56(2):1602-1612. https://doi.org/10.1109/taes.2019.2924175

Lee S, Lee Y, Kim Y, et al., 2023. Impact angle control guidance considering seeker's field-of-view limit based on reinforcement learning. *Journal of Guidance, Control, and Dynamics*, 46(11):2168-2182. https://doi.org/10.2514/1.G007715

Li JQ, Zhang GQ, Shan QH, et al., 2023. A novel cooperative design for USV–UAV systems: 3-D mapping guidance and adaptive fuzzy control. *IEEE Transactions on Control of Network Systems*, 10(2):564-574. https://doi.org/10.1109/TCNS.2022.3220705

Li ZH, Hu C, Ding CB, et al., 2018. Stochastic gradient particle swarm optimization based entry trajectory rapid planning for hypersonic glide vehicles. *Aerospace Science and Technology*, 76:176-186. https://doi.org/10.1016/j.ast.2018.01.033

Li ZH, He B, Wang MH, et al., 2019. Time-coordination entry guidance for multi-hypersonic vehicles. *Aerospace Science and Technology*, 89:123-135. https://doi.org/10.1016/j.ast.2019.03.056

Liang ZX, Li QD, Ren Z, 2017. Virtual terminal-based adaptive predictor–corrector entry guidance. *Journal of Aerospace Engineering*, 30(4):04017013. https://doi.org/10.1061/(asce)as.1943-5525.0000716

Liang ZX, Lv C, Zhu SY, 2023. Lateral entry guidance with terminal time constraint. *IEEE Transactions on Aerospace and Electronic Systems*, 59(3):2544-2553. https://doi.org/10.1109/TAES.2022.3215554

Liu X, Li X, Zhang HJ, et al., 2025. Entry guidance with terminal time constraint based on reduced-order dynamics. *IEEE Transactions on Aerospace and Electronic Systems*, 61(3):5949-5961. https://doi.org/10.1109/TAES.2024.3524200

Lu P, 1997. Entry guidance and trajectory control for reusable launch vehicle. *Journal of Guidance, Control, and Dynamics*, 20(1):143-149. https://doi.org/10.2514/2.4008

Lu P, 2014. Entry guidance: a unified method. *Journal of Guidance, Control, and Dynamics*, 37(3):713-728. https://doi.org/10.2514/1.62605

Phillips TH, 2003. A common aero vehicle (CAV) model, description, and employment guide. *Schafer Corporation for AFRL and AFSPC*, 27:1-9.

Qiu XQ, Lai P, Gao CS, et al., 2024. Recorded recurrent deep reinforcement learning guidance laws for intercepting endo-atmospheric maneuvering missiles. *Defence Technology*, 31:457-470. https://doi.org/10.1016/j.dt.2023.02.016

Ren LL, Xian Y, Li SP, et al., 2023. Robust depletion shutdown guidance algorithm for long-range vehicles with a solid divert control system in large deviation conditions. *Advances in Space Research*, 72(9):3818-3841. https://doi.org/10.1016/j.asr.2023.07.049

Ren LL, Guo WL, Xian Y, et al., 2025. Deep reinforcement learning based integrated evasion and impact hierarchical intelligent policy of exo-atmospheric vehicles. *Chinese Journal of Aeronautics*, 38(1):103193. https://doi.org/10.1016/j.cja.2024.08.024

Schulman J, Wolski F, Dhariwal P, et al., 2017. Proximal policy optimization algorithms. arXiv:1707.06347. https://doi.org/10.48550/arXiv.1707.06347

Shen ZJ, Lu P, 2003. Onboard generation of three-dimensional constrained entry trajectories. *Journal of Guidance, Control, and Dynamics*, 26(1):111-121. https://doi.org/10.2514/2.5021

Suresh M, Swar SC, Shyam S, 2023. Autonomous cooperative guidance strategies for unmanned aerial vehicles during onboard emergency. *Journal of Aerospace Information Systems*, 20(2):102-113. https://doi.org/10.2514/1.I011095

Wang CY, Wang WL, Dong W, et al., 2024. Multiple-stage spatial–temporal cooperative guidance without time-to-go estimation. *Chinese Journal of Aeronautics*, 37(9):399-416. https://doi.org/10.1016/j.cja.2024.05.026

Wang HN, Guo J, Wang X, et al., 2022. Time-coordination entry guidance using a range-determined strategy. *Aerospace Science and Technology*, 129:107842. https://doi.org/10.1016/j.ast.2022.107842

Wang NY, Wang XG, Cui NG, et al., 2022. Deep reinforcement learning-based impact time control guidance law with constraints on the field-of-view. *Aerospace Science and Technology*, 128:107765. https://doi.org/10.1016/j.ast.2022.107765

Xue SB, Lu P, 2010. Constrained predictor–corrector entry guidance. *Journal of Guidance, Control, and Dynamics*, 33(4):1273-1281. https://doi.org/10.2514/1.49557

Yang HD, Liang HZ, Liu JQ, et al., 2024. Analytical time-coordinated entry guidance for multi-hypersonic vehicles within three-dimensional corridor. *Aerospace Science and Technology*, 155:109639. https://doi.org/10.1016/j.ast.2024.109639

Yang HW, Hu JC, Li S, et al., 2024. Reinforcement-learning-based robust guidance for asteroid approaching. *Journal of Guidance, Control, and Dynamics*, 47(10):2058-2072. https://doi.org/10.2514/1.G008085

Yu WB, Chen WC, Jiang ZG, et al., 2019. Analytical entry guidance for coordinated flight with multiple no-fly-zone constraints. *Aerospace Science and Technology*, 84:273-290. https://doi.org/10.1016/j.ast.2018.10.013

Zeng L, Zhang HB, Zheng W, 2018. A three-dimensional predictor–corrector entry guidance based on reduced-order motion equations. *Aerospace Science and Technology*, 73:223-231. https://doi.org/10.1016/j.ast.2017.12.009

Electronic supplementary materials

Sections S1–S3, Tables S1–S3, Figs. S1–S6, Eqs. (S1)–(S17)