Prescribed-Performance Heading Control for Unmanned Surface Vessels via Fixed-Time Recursive Design and SAC-Based Residual Compensation

Yingjie Deng[1,2], Yong Li[2], Chenfeng Huang[3], Fangcheng Liu[2], Namkyun Im[4] and Yifei Xu[2,*]

[1]State Key Laboratory of Crane Technology, Yanshan University, Qinhuangdao, 066004, China

[2]School of Mechanical Engineering, Yanshan University, Qinhuangdao, 066004, China

[3]Navigation College, Dalian Maritime University, Dalian, 116026, China

[4]Division of Navigation, Mokpo National Maritime University, Mokpo, 530729, Republic of Korea

Abstract: This paper proposes a prescribed-performance fixed-time recursive control framework with Soft Actor-Critic (SAC) residual compensation for unmanned surface vessel heading tracking under model uncertainties, external disturbances, and actuator saturation. A prescribed performance error transformation is adopted to strictly constrain tracking errors within predefined transient and steady-state bounds. Combined with a double-power recursive control law, fixed-time error convergence independent of initial conditions is guaranteed. A bounded lightweight SAC module is further embedded to compensate for lumped uncertainties without exact disturbance reconstruction, improving tracking performance while retaining the inherent stability of the model-based control scheme. Lyapunov analysis verifies the bounded closed-loop signal behavior, guaranteed prescribed performance, and fixed-time convergence. Comparative simulations confirm that the proposed method delivers higher tracking accuracy, smoother control responses, and stronger robustness than existing baseline and learning-based compensation strategies.

Keywords: Unmanned surface vessel, heading tracking, prescribed-performance control, fixed-time recursive design, reinforcement learning, residual disturbance compensation.

1. INTRODUCTION

Unmanned surface vessels (USVs), which encompass both sail-driven and conventional actuation configurations, have attracted significant research attention in recent years. Endowed with advantages such as long endurance, low energy consumption, and strong adaptability, USVs are widely regarded as promising platforms for marine environmental monitoring, autonomous navigation, and ocean observation missions [1]. As a core capability for autonomous operation, motion control directly dictates the navigation safety, tracking precision, and mission reliability of USVs [2]. However, in real-world marine environments, control system design is confronted with inherent strong nonlinear dynamics, parametric uncertainties, unknown external disturbances, and actuator saturation constraints. These complex and coupled factors collectively present considerable challenges to the development of high-performance motion control strategies [3, 4].

Among various motion control problems, heading regulation is recognized as one of the most fundamental and practically significant tasks for USVs. On the one hand, heading control serves as the cornerstone of high-level guidance and path-following frameworks, where the desired course generated by the guidance layer is ultimately implemented through the low-level heading control channel [5-7]. On the other hand, the heading subsystem is highly sensitive to hydrodynamic uncertainties, environmental disturbances, and actuator saturation, which can significantly degrade tracking performance if not properly compensated [2, 3]. As a conventional heading control approach, proportional-integral-derivative (PID) control has been widely adopted due to its simplicity and effectiveness [8, 9]. However, it is generally constrained by the inherent uncertainties and nonlinear dynamics of unmanned surface vessels systems. To mitigate these drawbacks, fuzzy logic control and backstepping design have been successfully applied to the heading control of unmanned sailboats [10-12]. Nevertheless, under complex operating conditions, purely model-based control strategies still face an inherent trade-off between tracking accuracy, robustness, and practical implementation complexity.

To explicitly characterize and regulate both transient and steady-state tracking behaviors, prescribed performance control (PPC) has emerged as a robust paradigm for confining tracking errors within a set of predefined bounded envelopes [13, 14]. Concurrently, fixed-time control and its associated double-power control strategies exhibit remarkable superiority in accelerating the convergence rate and enhancing the robustness of uncertain nonlinear systems [15, 16]. Nevertheless, when such

*Address correspondence to this author at the School of Mechanical Engineering, Yanshan University, Qinhuangdao, 066004, China;
E-mail: xyf@ysu.edu.cn

Received on 26-05-2026 Accepted on 29-06-2026 Published on 06-07-2026

model-based control schemes are implemented in practical USV heading control scenarios, residual modeling errors and complex external disturbances inevitably deteriorate the achievable tracking precision. This challenge becomes particularly pronounced under the simultaneous existence of actuator saturation and lumped uncertainties [17].

Over the past few decades, rapid advancements in artificial intelligence (AI) have provided numerous novel insights and paradigms for the control community. Consequently, a variety of neural network-based learning algorithms have been successfully incorporated into the design of intelligent controllers. Among these approaches, reinforcement learning (RL) has recently emerged as a particularly promising technique to further enhance the control performance of uncertain nonlinear systems [18]. Specifically, [19] demonstrated that RL-based methods exhibit considerable potential in marine path-following and motion control tasks, as they can achieve satisfactory control performance without relying on an accurate system model. [20] developed an end-to-end path-following controller using the deep deterministic policy gradient (DDPG) algorithm, which enables direct execution of tracking commands and validates the overall effectiveness of the proposed scheme through numerical simulations and practical experiments. [21] proposed a DDPG-based path-following control scheme utilizing a composite state space and a dynamic threshold reward function. By fusing current and historical states to satisfy the Markov property and adopting a hierarchical dynamic reward design to accelerate training convergence, the proposed method achieves high-precision, strong delay robustness, and fast-stable path tracking for USVs. [22] presented a fully data-driven motion controller based on deep reinforcement learning (DRL). By training a prediction model using recorded input-output data, the proposed controller is capable of accomplishing trajectory tracking tasks with high efficiency. [23] proposed a USV path planning method based on the conventional Q-learning algorithm, which can generate near-optimal and physically feasible paths under the nonholonomic motion constraints of marine vehicles. Furthermore, in disturbance-affected control systems, learning-enhanced compensation mechanisms are commonly adopted to mitigate the adverse effects of unknown dynamics and external perturbations. [24] developed a traditional PID controller optimized by the soft actor-critic (SAC) algorithm, which can achieve adaptive control and effectively suppress external disturbances. [25] employed the TD3 reinforcement learning algorithm to construct a disturbance observation module, which estimates the unknown external disturbances and unmodeled dynamics of an unmanned sailboat online and utilizes the estimated values for feedforward compensation; by integrating backstepping sliding mode control with dynamic dual-mode compensation, the method ultimately achieves high-precision, strong robustness, low overshoot, and smooth rudder angle control in the heading tracking of the unmanned sailboat. [26] argued that the performance of learning modules should be evaluated from the perspective of the resulting closed-loop behavior, robustness, and control smoothness, rather than by disturbance-reconstruction accuracy alone. This viewpoint is particularly relevant for marine heading control, where the ultimate objective is not to reproduce the unknown disturbance exactly, but to maintain satisfactory closed-loop tracking performance under model uncertainty, environmental perturbations, and actuator limitations.

Motivated by the aforementioned observations, this paper develops a heading control framework for unmanned surface vessels (USVs) by integrating a prescribed-performance recursive controller with a bounded Soft Actor-Critic (SAC)-based residual compensation mechanism. First, a prescribed-performance error transformation is introduced to convert the constrained heading-tracking problem into an unconstrained transformed-error stabilization problem. Subsequently, a double-power recursive control law is constructed to provide a fixed-time-inspired convergence structure for both the transformed heading error and the yaw-rate tracking error. On this basis, an SAC module is embedded into the control framework as an auxiliary residual compensator. Specifically, the SAC module designed in this work generates bounded and smooth residual compensation signals, thereby effectively mitigating the adverse impacts of model uncertainties and environmental disturbances on practical heading-tracking performance.

Compared with existing marine control studies that primarily focus on guidance-law improvement, path-following design, or end-to-end reinforcement learning (RL) policies [27], the present study concentrates on a more compact and practically implementable single-heading control framework. Furthermore, although several recent works have explored the integration of prescribed performance and RL mechanisms, or employed RL-based disturbance compensators within composite control architectures [19], the present work differs in that the RL module is intentionally restricted to a bounded residual-compensation role. In this regard, the model-based PPC/fixed-time-style controller retains the dominant closed-loop stability structure, while the SAC module serves solely as a lightweight performance-enhancement channel. This design renders the overall method more amenable to both rigorous Lyapunov stability analysis and practical engineering implementation.

The main contributions of this paper are summarized as follows.

- A prescribed-performance fixed-time recursive heading controller is developed for the reduced heading subsystem of a USV. By combining the prescribed-performance error transformation with a double-power recursive design, the proposed controller constrains the heading-tracking error within a predefined performance envelope and establishes a fixed-time convergence structure for the transformed heading error and the yaw-rate tracking error.

- A bounded SAC-based residual compensation mechanism is incorporated into the analytical controller. Unlike disturbance-observer-like learning modules, the proposed SAC compensator does not aim to reconstruct the total lumped disturbance exactly. Instead, it generates a bounded residual signal to improve closed-loop tracking performance while preserving the dominant stability structure of the model-based controller.

- A Lyapunov-based stability analysis is provided for the closed-loop system under model mismatch, bounded disturbances, actuator saturation, and bounded residual compensation. Theoretical results show that all closed-loop signals remain bounded, the prescribed-performance constraint is preserved, and the transformed tracking errors converge in fixed time under the proposed control law.

The remainder of this paper is organized as follows. Section 2 presents the heading-subsystem model and formulates the control problem. Section 3 develops the prescribed-performance recursive controller and the SAC-based residual compensation mechanism. Section 4 provides the closed-loop stability analysis. Section 5 reports simulation results to verify the effectiveness of the proposed method. Finally, Section 6 concludes the paper.

2. PROBLEM FORMULATION

2.1. Coordinate Systems and Original Vessel Model

To describe the vessel motion and clarify the physical background of the heading subsystem used later, the earth-fixed frame $O\text{-}XY$ and the body-fixed frame $o\text{-}xy$ are introduced, as shown in Figure **1**. The earth-fixed frame is used to describe the vessel position in the inertial plane, whereas the body-fixed frame is attached to the vessel hull and is used to define the body velocities and control-related variables.

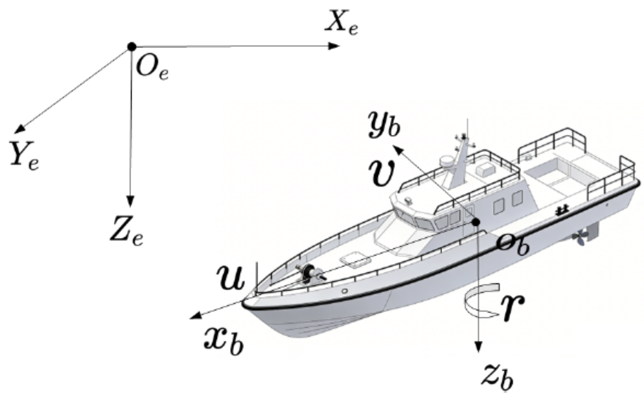


Figure 1: Earth-fixed and body-fixed coordinate frames of the unmanned surface vessel. The earth-fixed frame is used to describe the vessel position (x,y), while the body-fixed frame is used to define the surge velocity u, sway velocity v, and yaw rate r. The heading angle is denoted by ψ.

Let

$$\eta = [x, y, \psi]^{\mathrm{T}}, \qquad \upsilon = [u, v, r]^{\mathrm{T}} \tag{1}$$

The planar vessel motion can be described in the standard form

$$\begin{cases} \dot{\eta} = R(\psi)\upsilon \\ M\dot{\upsilon} + C(\upsilon)\upsilon + D(\upsilon)\upsilon = \tau_{\upsilon} + \tau_{d} \end{cases} \tag{2}$$

where

$$R(\psi) = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

M is the inertia matrix, $C(\upsilon)$ is the Coriolis and centripetal matrix, $D(\upsilon)$ is the hydrodynamic damping matrix, τ_{υ} denotes the control-related generalized force/moment vector, and τ_{d} denotes the environmental disturbance and unmodeled generalized force/moment vector.

(2) represents the original vessel model prior to heading-channel simplification. In this paper, the full vessel motion is acknowledged at the modeling level, while the controller is developed for the low-level single-heading control layer.

2.2. Reduced Heading-Channel Model

Since the objective of this paper is heading tracking, only the yaw channel is extracted for controller design. The resulting reduced-order heading subsystem is written as

$$\begin{cases} \dot{\psi} = r \\ \dot{r} = f(u,v,r) + b\tau + d \end{cases} \tag{4}$$

where $\tau \in \mathbb{R}$ is the heading control input, $f_d(u,v,r)$ denotes the dominant yaw-channel dynamics, $b > 0$ is the effective input gain, and d denotes the lumped residual uncertainty. For notational simplicity, $f(u,v,r)$ can be simplified as $f(r)$.

The reduced-order model (4) is obtained from the original vessel dynamics by retaining the dominant yaw motion and lumping the remaining coupling, uncertainty, and disturbance terms into d.

For controller design, the nominal part of the yaw-rate dynamics is denoted by

$$f_n(r) + b_n\tau \tag{5}$$

where $f_n(r)$ and $b_n > 0$ are the nominal yaw dynamics and nominal input gain, respectively. Accordingly, the actual yaw-rate dynamics can be decomposed as

$$\dot{r} = f_n(r) + b_n\tau + (f(r) - f_n(r)) + (b - b_n)\tau + d \tag{6}$$

To account for actuator saturation, define the actual control input as

$$\tau = sat(\tau^*, \tau_{\max}) = \begin{cases} \tau_{\max} & \tau^* > \tau_{\max} \\ \tau^* & |\tau^*| \le \tau_{\max} \\ -\tau_{\max} & \tau^* < -\tau_{\max} \end{cases} \tag{7}$$

where τ^* is the unsaturated control command and $\tau_{\max} > 0$ is the actuator limit. The corresponding saturation-induced mismatch is defined as

$$\Delta_s = b_n(\tau - \tau^*) \tag{8}$$

2.3. Prescribed-Performance Requirement

Let $\psi_d(t)$ denote the desired heading signal generated by an upper-level guidance or reference module. Define the heading-tracking error as

$$e_\psi = \psi - \psi_d \tag{9}$$

To explicitly shape the transient and steady-state behavior of e_ψ, the prescribed-performance function is chosen as

$$\rho(t) = \begin{cases} (\rho_0 - \rho_\infty)\left(1 - \dfrac{t}{T_\rho}\right)^{\ell} + \rho_\infty & 0 \le t < T_\rho \\ \rho_\infty & t \ge T_\rho \end{cases} \tag{10}$$

where $\rho_0 > \rho_\infty > 0$, $T_\rho > 0$, and $\ell > 0$ are design constants. The function $\rho(t)$ is positive and monotonically decreasing on $[0, T_\rho)$, and then remains equal to the residual bound ρ_∞.

The prescribed-performance requirement is

$$|e_\psi(t)| < \rho(t), \qquad \forall t \ge 0 \tag{11}$$

provided that the initial error lies inside the admissible performance set.

2.4. Control Objective

Based on (4)–(11), the objective is to design a prescribed-performance fixed-time heading controller such that the heading-tracking error remains within the prescribed-performance envelope, while the transformed heading error and the yaw-rate tracking error converge to zero in fixed time despite model mismatch, input-gain uncertainty, actuator saturation, and bounded auxiliary compensation from the learning module.

3. CONTROLLER DESIGN

In this section, a prescribed-performance fixed-time heading controller with SAC-based bounded auxiliary compensation is developed for the reduced-order heading subsystem. The design consists of three steps. First, the constrained heading-tracking error is transformed into an unconstrained variable by means of a prescribed-performance error transformation. Second, a virtual yaw-rate command is designed to impose a fixed-time convergence structure on the transformed error dynamics. Third, the actual control law is constructed by combining nominal model compensation, double-power feedback, a coupling-cancellation term, and a robust sign term. The SAC module is incorporated as an auxiliary bounded compensation channel and does not replace the model-based control backbone. The overall structure of the proposed controller is illustrated in Figure **2**.

3.1. Notation and Assumptions

Recall the heading-tracking error $e_\psi = \psi - \psi_d$ defined in Section 2, and introduce the normalized error

$$\xi = \frac{e_\psi}{\rho} \tag{12}$$

Then, define the prescribed-performance transformation

$$z_\psi = \frac{1}{2}\ln\left(\frac{1+\xi}{1-\xi}\right), \qquad |\xi| < 1 \tag{13}$$

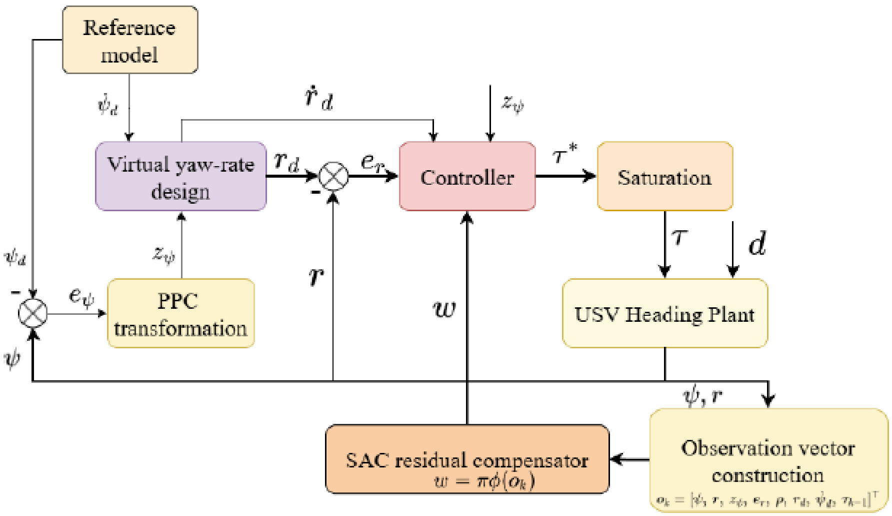


Figure 2: controller framework .

The second-layer tracking error is defined as

$$e_r = r - r_d \tag{14}$$

where r_d is a virtual yaw-rate command to be designed later.

For compact notation, define the signed-power function

$$sig(x)^a = |x|^a\, sign(x), \qquad a > 0 \tag{15}$$

To establish fixed-time convergence, choose the exponents as

$$0 < \alpha < 1, \qquad \beta > 1 \tag{16}$$

The controller design is based on the following assumptions.

Assumption 1. The desired heading signal ψ_d is twice continuously differentiable, and ψ_d, $\dot{\psi}_d$, and $\ddot{\psi}_d$ are bounded.

Assumption 2. The functions $f(r)$ and $f_n(r)$ are locally Lipschitz in r, and the mismatch $f(r) - f_n(r)$ is bounded on the operating domain of interest.

Assumption 3. The actual input gain satisfies $b > 0$, the nominal input gain satisfies $b_n > 0$, and the mismatch $b - b_n$ is bounded.

Assumption 4. The lumped uncertainty d is bounded, namely,

$$|d(t)| \le \bar{d} \tag{17}$$

where $\bar{d} > 0$ is an unknown finite constant.

Assumption 5. The SAC-generated auxiliary compensation signal ω is bounded by design, i.e.,

$$|\omega(t)| \le \varpi \tag{18}$$

where $\varpi > 0$ is a prescribed design constant.

Assumption 6. The initial heading-tracking error satisfies

$$|e_\psi(0)| < \rho(0) \tag{19}$$

so that the prescribed-performance transformation is initially well defined.

3.2. Prescribed-Performance Error Transformation

Differentiating (13) with respect to time gives

$$\dot{z}_\psi = \frac{1}{1-\xi^2}\dot{\xi} \tag{20}$$

Since

$$\dot{\xi} = \frac{\dot{e}_\psi}{\rho} - \frac{e_\psi \dot{\rho}}{\rho^2} \tag{21}$$

and

$$\dot{e}_\psi = \dot{\psi} - \dot{\psi}_d = r - \dot{\psi}_d \tag{22}$$

one obtains

$$\dot{z}_\psi = \frac{1}{\rho(1-\xi^2)}\left(r - \dot{\psi}_d - \dot{\rho}\xi\right) \tag{23}$$

For convenience, define

$$\chi_\psi = \frac{1}{\rho(1-\xi^2)} \tag{24}$$

Then, the transformed error dynamics become

$$\dot{z}_\psi = \chi_\psi\left(r - \dot{\psi}_d - \dot{\rho}\xi\right) \tag{25}$$

3.3. Step 1: Virtual Yaw-Rate Design

To impose a fixed-time convergence structure on the transformed heading error, choose the virtual yaw-rate command as

$$r_d = \dot{\psi}_d + \dot{\rho}\xi - \frac{1}{\chi_\psi}\left(k_{\psi 1} sig(z_\psi)^\alpha + k_{\psi 2} sig(z_\psi)^\beta\right) \quad (26)$$

where $k_{\psi 1} > 0$ and $k_{\psi 2} > 0$ are design gains.

Substituting (26) into (25) yields

$$\dot{z}_\psi = \chi_\psi (r - r_d) - k_{\psi 1} sig(z_\psi)^\alpha - k_{\psi 2} sig(z_\psi)^\beta \quad (27)$$

Using (14), one obtains

$$\dot{z}_\psi = \chi_\psi e_r - k_{\psi 1} sig(z_\psi)^\alpha - k_{\psi 2} sig(z_\psi)^\beta \quad (28)$$

(28) shows that once the second-layer error e_r is stabilized, the transformed heading error z_ψ admits a double-power fixed-time convergence structure.

3.4. Step 2: Actual Control Law Design

Using the nominal decomposition of the yaw-rate dynamics introduced in (6), the second-layer tracking error dynamics can be written as

$$\dot{e}_r = \dot{r} - \dot{r}_d \quad (29)$$

The actual control input is implemented through the saturation mapping introduced in (7), and the corresponding mismatch term is denoted by Δ_s as defined in (8).

The unsaturated control command is designed as

$$\begin{aligned} \tau^* = \frac{1}{b_n}[&-f_n(r) + \dot{r}_d - \omega - k_{r1} sig(e_r)^\alpha - \\ &k_{r2} sig(e_r)^\beta - c_r e_r - \chi_\psi z_\psi - k_s\, sign(e_r)] \end{aligned} \quad (30)$$

where $k_{r1} > 0$ and $k_{r2} > 0$ are the double-power feedback gains, $c_r > 0$ is the linear damping gain, and $k_s > 0$ is the robust switching gain.

Using (6)–(8), and (30) into (29) yields

$$\begin{aligned} \dot{e}_r &= (f(r) - f_n(r)) + (b - b_n)\tau + d - \omega + \Delta_s \\ &- k_{r1} sig(e_r)^\alpha - k_{r2} sig(e_r)^\beta \\ &- c_r e_r - \chi_\psi z_\psi - k_s\, sign(e_r) \end{aligned} \quad (31)$$

Define the total residual perturbation as

$$\Delta_r = (f(r) - f_n(r)) + (b - b_n)\tau + d - \omega + \Delta_s \quad (32)$$

Then, (31) can be rewritten as

$$\begin{aligned} \dot{e}_r &= -k_{r1} sig(e_r)^\alpha - k_{r2} sig(e_r)^\beta \\ &- c_r e_r - \chi_\psi z_\psi - k_s\, sign(e_r) + \Delta_r \end{aligned} \quad (33)$$

Equation (33) explicitly shows that the model mismatch, input-gain mismatch, SAC auxiliary compensation, actuator saturation effect, and lumped uncertainty are all embedded into the closed-loop error dynamics through the total residual term Δ_r.

3.5. SAC-Based Bounded Auxiliary Compensation

In this paper, the SAC module is not interpreted as a strict estimator of the external disturbance or the lumped uncertainty. Instead, it serves as an auxiliary bounded compensation channel. Let

$$o_k = \left[\psi_k, r_k, z_{\psi,k}, e_{r,k}, \rho_k, r_{d,k}, \dot{\psi}_{d,k}, \tau_{k-1}\right]^{\mathrm{T}} \quad (34)$$

The reward function is designed to penalize heading tracking error, yaw-rate tracking error, excessive control effort, and excessive residual action:

$$r_k = -q_\psi e_\psi^2 - q_r e_r^2 - q_\tau \tau^2 \quad (35)$$

Here, q_ψ =12, q_r =1.8, q_τ =0.04 are positive weighting coefficients. The first two terms encourage accurate heading and yaw-rate tracking, while the last terms prevent overly aggressive rudder input. o_k denote the observation vector at time step k. Based on o_k, the SAC actor generates a bounded compensation signal ω_k satisfying

$$|\omega_k| \le \bar{\omega} \quad (36)$$

The signal ω_k is injected into the control law through the term $-\omega$ in (30).

4. STABILITY ANALYSIS

This section establishes the fixed-time convergence of the proposed closed-loop system and proves that the prescribed-performance constraint is preserved for all time.

4.1. A Fixed-Time Stability Lemma

The following standard lemma is used in the subsequent analysis.

Lemma 1 Let $V(t)$ be a positive definite and continuously differentiable function satisfying

$$\dot{V}(t) \le -aV^\mu(t) - bV^\nu(t) \quad (37)$$

where $a > 0$, $b > 0$, $0 < \mu < 1$, and $\nu > 1$. Then, the origin of the corresponding system is fixed-time stable. Moreover, the settling time is uniformly bounded by

$$T^* \le \frac{1}{a(1-\mu)} + \frac{1}{b(\nu - 1)} \quad (38)$$

4.2. Main Result

Under Assumptions 2–5 and the bounded actuator input (7), the residual term Δ_r defined in (32) is bounded. Hence, there exists a finite constant $\bar{\Delta}_r > 0$ such that

$$|\Delta_r(t)| \le \bar{\Delta}_r, \qquad \forall t \ge 0 \tag{39}$$

The main result is stated as follows.

Theorem 1 Consider the reduced-order heading subsystem (4) under the prescribed-performance transformation (12)–(13), the virtual yaw-rate command (26), the actual control law (30)–(33), and the SAC-based bounded auxiliary compensation (34)–(36). Suppose that Assumptions 1–6 hold and that the robust gain is chosen such that

$$k_s > \bar{\Delta}_r \tag{40}$$

Then, all closed-loop signals remain bounded, the transformed heading error z_ψ and the yaw-rate tracking error e_r converge to zero in fixed time, and the heading-tracking error satisfies

$$|e_\psi(t)| < \rho(t), \qquad \forall t \ge 0 \tag{41}$$

and converges to zero in fixed time as well.

Proof. Choose the composite Lyapunov function

$$V = \frac{1}{2}z_\psi^2 + \frac{1}{2}e_r^2 \tag{42}$$

Clearly, V is positive definite with respect to (z_ψ, e_r).

Differentiating (42) along (28) and (33) gives

$$\begin{aligned}
\dot{V} &= z_\psi \dot{z}_\psi + e_r \dot{e}_r \\
&= z_\psi \left(\chi_\psi e_r - k_{\psi 1}\, sig(z_\psi)^\alpha - k_{\psi 2}\, sig(z_\psi)^\beta\right) \\
&\quad + e_r \left(-k_{r1}\, sig(e_r)^\alpha - k_{r2}\, sig(e_r)^\beta - c_r e_r - \chi_\psi z_\psi - k_s\, sign(e_r) + \Delta_r\right)
\end{aligned}$$

$$\begin{aligned}
&= -k_{\psi 1}|z_\psi|^{1+\alpha} - k_{\psi 2}|z_\psi|^{1+\beta} - k_{r1}|e_r|^{1+\alpha} \\
&\quad - k_{r2}|e_r|^{1+\beta} - c_r e_r^2 - k_s|e_r| + e_r\Delta_r
\end{aligned} \tag{43}$$

Using (39), one has

$$e_r\Delta_r \le |e_r||\Delta_r| \le \bar{\Delta}_r |e_r| \tag{44}$$

Substituting (44) into (43) yields

$$\begin{aligned}
\dot{V} &\le - k_{\psi 1}|z_\psi|^{1+\alpha} - k_{\psi 2}|z_\psi|^{1+\beta} \\
&\quad - k_{r1}|e_r|^{1+\alpha} - k_{r2}|e_r|^{1+\beta} - c_r e_r^2 - (k_s - \bar{\Delta}_r)|e_r|
\end{aligned} \tag{45}$$

By the gain condition (40), the last term is nonpositive. Therefore,

$$\dot{V} \le -a_1(|z_\psi|^{1+\alpha} + |e_r|^{1+\alpha}) - a_2(|z_\psi|^{1+\beta} + |e_r|^{1+\beta}) \tag{46}$$

where

$$a_1 = \min\{k_{\psi 1}, k_{r1}\}, \qquad a_2 = \min\{k_{\psi 2}, k_{r2}\} \tag{47}$$

Since (42) is a positive definite quadratic form, it is equivalent to the squared Euclidean norm of the vector $[z_\psi, e_r]^{\mathrm{T}}$. Hence, by equivalence of norms in finite-dimensional space, there exist positive constants λ_1 and λ_2 such that

$$|z_\psi|^{1+\alpha} + |e_r|^{1+\alpha} \ge \lambda_1 V^{\frac{1+\alpha}{2}} \tag{48}$$

$$|z_\psi|^{1+\beta} + |e_r|^{1+\beta} \ge \lambda_2 V^{\frac{1+\beta}{2}} \tag{49}$$

Substituting (48)–(49) into (46) gives

$$\dot{V} \le -\bar{\lambda}_1 V^\mu - \bar{\lambda}_2 V^\vartheta \tag{50}$$

where

$$\mu = \frac{1+\alpha}{2} \in (0,1), \qquad \vartheta = \frac{1+\beta}{2} > 1 \tag{51}$$

and

$$\bar{\lambda}_1 := a_1\lambda_1 > 0, \qquad \bar{\lambda}_2 := a_2\lambda_2 > 0 \tag{52}$$

Since (50) is in the form of Lemma 1, it follows that the Lyapunov function $V(t)$ converges to zero in fixed time. Moreover, the settling time is uniformly bounded by

$$T^* \le \frac{1}{\bar{\lambda}_1(1-\mu)} + \frac{1}{\bar{\lambda}_2(\vartheta - 1)} \tag{53}$$

Hence,

$$V(t) = 0, \qquad \forall t \ge T^* \tag{54}$$

Since

$$V = \frac{1}{2}z_\psi^2 + \frac{1}{2}e_r^2$$

is positive definite with respect to (z_ψ, e_r), one has

$$z_\psi(t) = 0, \qquad e_r(t) = 0, \qquad \forall t \ge T^* \tag{55}$$

Next, by inverting the prescribed-performance transformation (13), one obtains

$$\xi = \tanh(z_\psi), \qquad e_\psi = \rho(t)\tanh(z_\psi) \tag{56}$$

Since $z_\psi(t)$ is well defined and bounded for all $t \ge 0$, one has $|\tanh(z_\psi(t))| < 1$, and therefore

$$|e_{\psi}(t)| = \rho(t)\,|\tanh(z_{\psi}(t))| < \rho(t), \qquad \forall t \geq 0 \tag{57}$$

Thus, the prescribed-performance constraint is preserved throughout the entire closed-loop evolution.

Moreover, from (55) and (56), one obtains

$$e_{\psi}(t) = 0, \qquad \forall t \geq T^* \tag{58}$$

Hence, the heading-tracking error also converges to zero in fixed time.

Finally, since z_{ψ}, e_r, τ, and ω are bounded, and all exogenous and nominal terms are bounded by assumption, all other closed-loop signals remain bounded as well. This completes the proof.

Remark 1 The proposed fixed-time result does not rely on exact reconstruction of the true lumped uncertainty by the SAC module. The SAC output ω is treated only as a bounded auxiliary compensation signal, and its effect is explicitly incorporated into the total residual term Δ_r.

Remark 2 Compared with the existing research, the present proof is built on a single composite Lyapunov function and a standard fixed-time stability lemma. This avoids repetitive subsystem-wise arguments and leads to a shorter, clearer, and logically closed proof chain.

5. SIMULATION RESULTS

In this section, numerical simulations are conducted to verify the effectiveness of the proposed prescribed-performance fixed-time heading controller integrated with SAC-based residual compensation. Two representative heading-tracking scenarios are considered to comprehensively evaluate the controller's performance. The first scenario is a single-step heading-tracking task, where the desired heading changes from 0° to 40°. This case is designed to examine the basic transient tracking performance of the proposed controller. The second scenario is a multi-reference heading-tracking task, where the desired heading ψ_d changes stepwise as $0^\circ \to 20^\circ \to -10^\circ \to 25^\circ \to 0^\circ$. This case is intended to further assess the robustness and adaptability of the controller under multiple maneuvering stages, which is more consistent with practical marine operating conditions.

The proposed PFT-SAC controller is compared with five baseline methods, namely PFT-NC, PFT-TD, PFT-DDPG, PFT-TD3, and PFT-PPO. The detailed definitions of these methods are given in Section 5.2. It should be noted that PFT-TD assumes that the injected disturbance is perfectly known and is therefore used only as a noncausal ideal benchmark. For all RL-based controllers, including PFT-DDPG, PFT-TD3, PFT-PPO, and PFT-SAC, the learned policy directly generates a deployable residual compensation signal without relying on ideal disturbance information during online evaluation.

5.1. Simulation Setup

In the simulation, the true yaw dynamics are represented as

$$\dot{r} = a_1 r + a_3 r^3 + b\tau + d(t) \tag{59}$$

The controller is designed based on the nominal yaw model

$$\dot{r} = a_{1n} r + a_{3n} r^3 + b_n \tau \tag{60}$$

The true plant parameters and the nominal parameters used by the controller are intentionally selected to be different, as shown in Table **1**.

The sampling interval is set as $\Delta t = 0.02$ s. The rudder input is constrained by the amplitude limit $|\tau| < \tau_{\max}$. where $\tau_{\max} = 35^\circ$.

The prescribed-performance boundary is selected as $\rho(t)$. where $\rho_0 = 40^\circ$ and $\rho_\infty = 10^\circ$ denote the initial and final allowable tracking error bounds, respectively. The PPC boundary is used to constrain the heading tracking error during the transient and steady-state processes.

5.1.1. Disturbance and Model Uncertainty Settings

To evaluate the robustness and residual compensation capability of the proposed controller, an external time-varying disturbance is injected into the yaw-rate dynamics. The disturbance is designed as a bounded composite signal:

$$d = k_d \begin{bmatrix} 21400\sin\left(\dfrac{t}{2}\right) + 386\sin\left(\dfrac{t}{2}\right) \\ \psi + 614\sin\left(\dfrac{t}{4}\right) + 264\sin\left(\dfrac{t}{7.5}\right) r \end{bmatrix} \tag{61}$$

Table 1: True and Nominal Model Parameters

Parameter	True plant	Nominal model	Relative mismatch
a_1	–5.428	–5.000	7.89%
a_3	–0.364	–0.300	17.58%
b	2.156	2.000	7.24%

where $k_d = 3e-5$ is the disturbance scaling factor.

5.2. Compared Methods and Reinforcement Learning Settings

Six methods are compared in the simulations:

- **PFT-NC**: the prescribed-performance fixed-time heading controller without compensation;
- **PFT-TD**: the same analytical controller with ideal disturbance compensation;
- **PFT-DDPG**: the analytical controller augmented with a DDPG-based residual compensator;
- **PFT-TD3**: the analytical controller augmented with a TD3-based residual compensator;
- **PFT-PPO**: the analytical controller augmented with a PPO-based residual compensator;
- **PFT-SAC**: the proposed controller with SAC-based residual compensation.

For the RL-based residual compensation methods, the actor receives the observation vector o_k. The main RL training parameters are summarized in Table **2**.

5.3. Performance Indices

The following indices are used to evaluate the heading tracking performance:

$$Mean\,|e_\psi| = \frac{1}{T}\int_0^T |e_\psi(t)|\,dt, \quad RMS(\tau) = \sqrt{\frac{1}{T}\int_0^T \tau^2(t)dt} \tag{62}$$

The maximum tracking error and the maximum rudder input are also reported.

5.4. Case 1: Single-step Heading Tracking

In the first case, the desired heading changes from 0° to 40°. This scenario is used to evaluate the basic transient response, steady-state accuracy, and control smoothness of the proposed method under a single large heading maneuver.

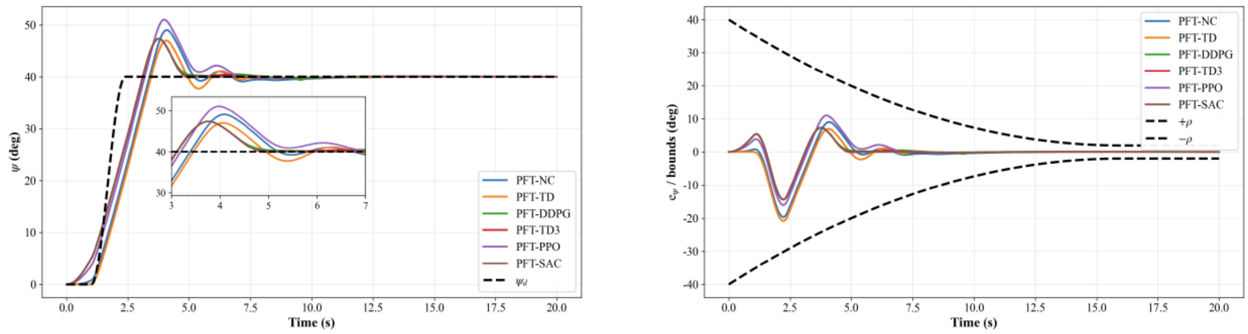
Figure **3** shows the heading response and the tracking error under the PPC constraint. The PFT-NC controller can track the desired heading, but its transient error is larger due to the absence of compensation. The ideal PFT-TD controller provides a benchmark when the true disturbance is available. The RL-based controllers reduce the tracking error to different degrees by generating deployable residual compensation signals.

Figure **4** shows the corresponding control input and training return curves. The rudder commands remain within the prescribed actuator bound. The training curves show the learning behavior of different RL algorithms in the residual compensation task.

Table **3** summarizes the quantitative results of Case 1. Compared with PFT-NC, the RL-based controllers reduce the mean heading tracking error to different degrees. In particular, PFT-SAC achieves the smallest mean tracking error of 1.2481°, while PFT-TD3 obtains a comparable value of 1.2577°. PFT-DDPG also improves the tracking performance, reducing the mean error from 1.7907° to 1.3055°. These results indicate that the learned residual compensation can effectively

Table 2: Main Reinforcement Learning Training Parameters

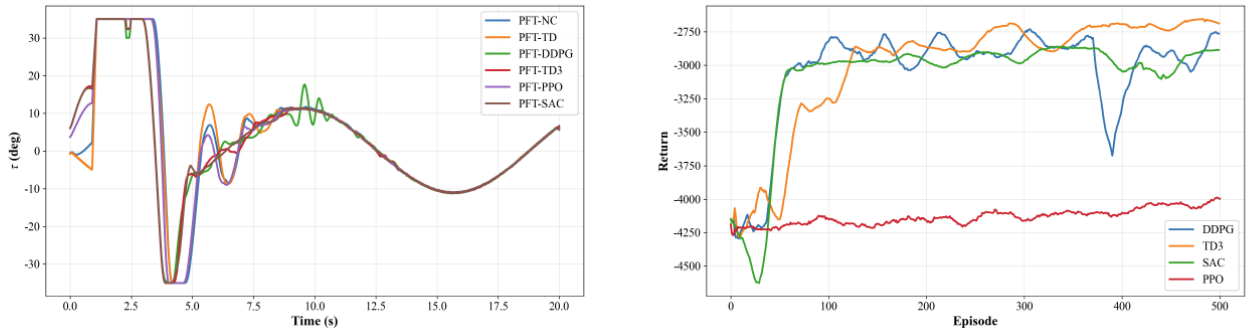
Parameter	Value	Description
Policy network	MLP	Actor and critic networks
Hidden layers	$[128,128]$	Network architecture
Activation function	ReLU	Hidden-layer activation
Replay buffer size	3×10^5	Off-policy replay buffer
Batch size	512	Mini-batch size
Discount factor γ	0.995	Reward discount factor
Learning rate	1×10^{-4}	Actor/critic learning rate
SAC entropy coefficient	Auto	Entropy tuning strategy
TD3 policy delay	3	Delayed actor update
TD3 target noise	0.02	Target policy smoothing noise
TD3 noise clip	0.05	Target noise clipping bound
DDPG exploration noise	0.03	Gaussian action noise
TD3 exploration noise	0.03	Gaussian action noise
PPO rollout steps	2048	Rollout length
PPO clip range	0.2	Clipped surrogate objective



(a) Heading response

(b) Tracking error and PPC bounds

Figure 3: Tracking performance in Case 1.



(a) Control input in Case 1.

(b) Training return comparison in Case 1.

Figure 4: Control input and training return curves in Case 1.

Table 3: Performance Comparison in Case 1

Mode	Mean $\lvert e_\psi \rvert$ (deg)	Steady $\lvert e_\psi \rvert$ (deg)	Max $\lvert e_\psi \rvert$ (deg)	Max $\lvert \tau \rvert$ (deg)	RMS (τ) (deg)
PFT-NC	1.7907	0.0145	19.6004	35.0000	16.3483
PFT-TD	1.7278	0.0029	20.8141	35.0000	16.1947
PFT-DDPG	1.3055	0.0093	14.3220	35.0000	15.2566
PFT-TD3	1.2577	0.0127	14.3346	35.0000	15.3485
PFT-PPO	1.7335	0.0266	16.0109	35.0000	16.3878
PFT-SAC	1.2481	0.0222	14.3963	35.0000	15.3748

improve the transient tracking accuracy under the single-step heading command.

5.5. Case 2: Multi-reference Heading Tracking

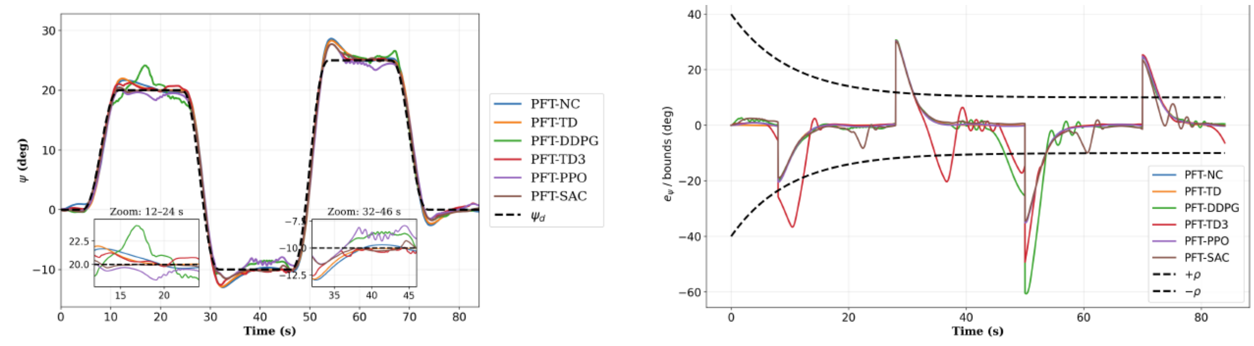
In the second case, a more challenging multi-reference heading command is considered: ψ_d changes stepwise as $0^\circ \to 20^\circ \to -10^\circ \to 25^\circ \to 0^\circ$.

This scenario contains both positive and negative heading transitions and is used to evaluate the adaptability of the proposed method under multiple maneuvering stages.

Figure **5** presents the heading tracking response and the PPC-constrained tracking error. All controllers maintain the heading tracking error within the prescribed-performance envelope. Compared with PFT-NC, the RL-based compensation methods reduce the transient deviation and improve the overall tracking performance. Among the RL-based methods, PFT-SAC achieves a favorable balance between tracking accuracy and control smoothness. It should be noted that the deviation of e_ψ beyond the prescribed bounds is acceptable since ψ_d is discontinuous.

Figure **6** gives the control input and training return comparison. The off-policy RL algorithms generally show faster convergence than PPO in the continuous residual compensation task. SAC exhibits stable learning behavior and achieves competitive performance during online evaluation.

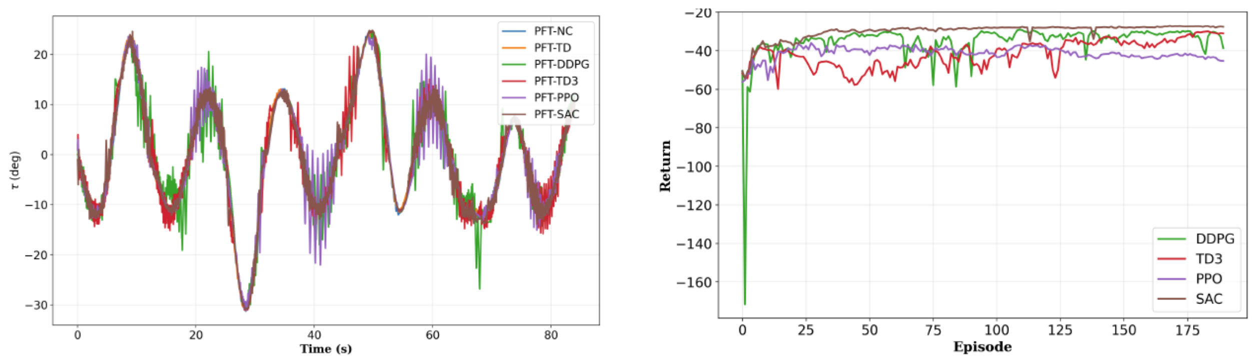
The quantitative results of Case 2 are reported in Table **4**. Compared with PFT-NC, the compensation-based methods generally improve the heading-tracking performance. The proposed



(a) Heading response

(b) Tracking error and PPC bounds

Figure 5: Tracking performance in Case 2.



(a) Control input

(b) Training return

Figure 6: Control input and training return curves in Case 2.

Table 4: Performance Comparison in Case 2: Multi-Reference Heading Tracking

Mode	Mean $\lvert e_\psi \rvert$ (deg)	Steady $\lvert e_\psi \rvert$ (deg)	Max $\lvert e_\psi \rvert$ (deg)	Max $\lvert \tau \rvert$ (deg)	RMS (τ) (deg)
PFT-NC	1.1262	1.1088	3.7843	31.2921	11.5189
PFT-TD	0.9993	1.1129	3.7030	31.2110	11.5477
PFT-DDPG	1.0515	0.7576	4.1497	31.0211	11.4266
PFT-TD3	0.7756	0.8130	3.1061	31.0908	11.6414
PFT-PPO	1.0455	0.9262	3.5959	31.1807	11.3966
PFT-SAC	0.6891	0.8260	3.1232	30.8924	11.3358

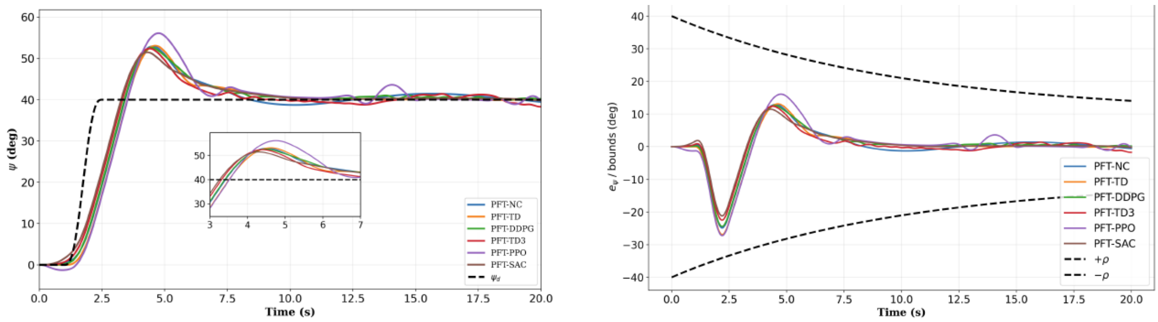
PFT-SAC achieves the smallest mean tracking error of 0.6891°, reducing the mean error by approximately 38.8% compared with PFT-NC. In addition, PFT-SAC yields the lowest maximum rudder amplitude and the smallest RMS rudder input, indicating reduced control effort in the multi-reference scenario. Although PFT-DDPG obtains the smallest steady-state error and PFT-TD3 gives the smallest maximum tracking error, PFT-SAC provides a more balanced performance in terms of average tracking accuracy, control effort, and prescribed-performance satisfaction.

5.6. Case 3: Additional Tests Under Increased Uncertainty and Stronger Disturbance

To further verify the transient convergence behavior and the implementation cost of the proposed method, an additional single-heading transition case is conducted. This condition sets stronger disturbances and model uncertainties based on the reference heading transition of Case 1.

Table 5: True and Nominal Model Parameters

Parameter	True plant	Nominal model	Relative mismatch
a_1	−5.428	−4.428	18.42%
a_3	−0.364	−0.270	25.82%
b	2.156	1.700	21.15%



(a) Heading response in Case 3.

(b) Tracking error in Case 3.

Figure 7: Tracking performance of different methods in Case 3.

Table 6: Performance Comparison in Case 3 under Increased Uncertainty and Stronger Disturbance

Mode	Mean $\lvert e_\psi \rvert$(deg)	Steady $\lvert e_\psi \rvert$(deg)	Max $\lvert e_\psi \rvert$ (deg)	Ts (s)	RMS(τ) (deg)
PFT-NC	1.6424	0.6618	7.4906	17.2600	18.0513
PFT-TD	1.1871	0.0629	9.4220	5.6800	18.3474
PFT-DDPG	1.2269	0.3671	7.1789	8.5800	18.2957
PFT-TD3	1.5077	0.9419	7.1504	8.0100	18.4934
PFT-PPO	1.9409	0.2960	10.4610	19.0800	20.3556
PFT-SAC	0.9307	0.2290	6.4471	5.7000	17.6410

The controller uses the nominal parameters listed in Table **5**. The disturbance scale is increased to $k_d = 4 \times 10^{-5}$. Therefore, this case provides a more challenging condition for examining the robustness, convergence behavior, and computational cost of the proposed PFT-SAC method.

The simulation results are shown in Figure **7**, and the quantitative comparison is summarized in Table **6**. It can be observed that all methods can track the desired heading after the transient response, while the proposed PFTSAC method achieves a better overall balance between transient recovery, steady-state accuracy, and control effort. Compared with PFT-NC, PFT-SAC reduces the mean heading tracking error from 1.6424° to 0.9307°, corresponding to an improvement of approximately 43.3%. The steady-state error is reduced from 0.6618° to 0.2290°, and the maximum tracking error is reduced from 7.4906° to 6.4471°. Moreover, the practical settling time under the 1° error threshold is shortened from 17.26 s to 5.70 s.

To reduce the effects of random initialization and exploration noise during reinforcement learning training, ten independent experiments with different random seeds were conducted for each learning-based residual compensation method. For each method, the training and evaluation procedure is repeated under the same Case 3 condition, and the mean and standard deviation of the main performance indices are reported. This statistical comparison provides a more reliable assessment of the tracking accuracy, convergence behavior, and control smoothness of different RL-based residual compensators.

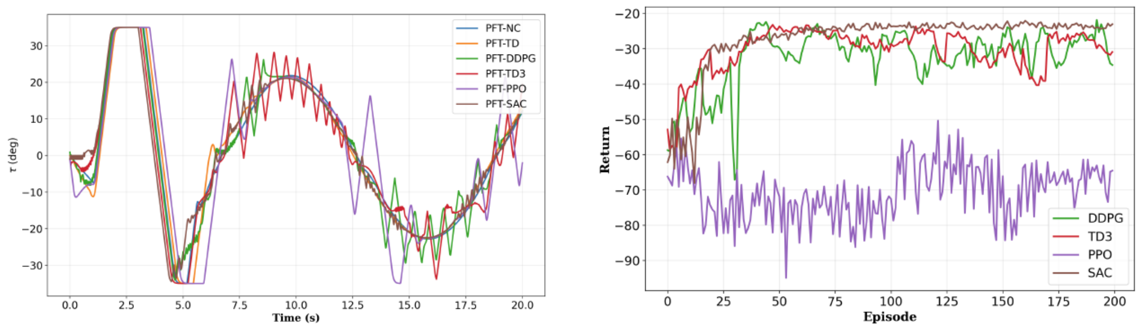
To further examine the influence of reward-weight selection, a sensitivity analysis is conducted under the Case 3 condition. The baseline reward setting is compared with several alternative settings by varying the heading-error weight q_ψ and the control-effort weight q_τ. The yaw-rate error weight q_r is kept unchanged because it mainly regulates the consistency between the actual yaw rate and the virtual yaw-rate command. The results are summarized in Table 8. The sensitivity results indicate that the reward weights affect the trade-off between tracking accuracy and control smoothness. A smaller q_ψ weakens the heading-error penalty and may lead to larger tracking

Table 7: Statistical Comparison of Learning-Based Residual Compensators Over Ten Independent Runs in Case 3

Method	Mean $\lvert e_\psi \rvert$ (deg)	Max $\lvert e_\psi \rvert$ (deg)	Ts (s)	RMS(τ) (deg)
PFT-DDPG	1.2516 ± 0.1955	8.0691 ± 1.2890	9.4920 ± 2.7798	18.2119 ± 0.4270
PFT-TD3	1.5366 ± 0.1196	7.8109 ± 1.3550	8.7560 ± 2.9103	18.1977 ± 0.3528
PFT-PPO	2.0319 ± 0.3233	9.8122 ± 1.6806	17.5080 ± 2.3849	19.8065 ± 0.8304
PFT-SAC	0.9874 ± 0.2827	7.2294 ± 1.2063	6.9080 ± 1.1784	17.6986 ± 0.2855

Table 8: Reward-Weight Sensitivity Analysis of PFT-SAC in Case 3

Setting	q_ψ	q_r	q_τ	Mean $\|e_\psi\|$ (deg)	Max $\|e_\psi\|$ (deg)	Ts (s)	RMS(τ) (deg)
Lower heading weight	8	1.8	0.04	1.6051	8.9446	6.4132	18.1152
Baseline	12	1.8	0.04	0.9307	6.4471	5.7000	17.6410
Higher heading weight	16	1.8	0.04	1.2178	6.7513	8.9831	17.9160
Lower effort weight	12	1.8	0.02	1.3587	6.8424	7.6648	18.0671
Higher effort weight	12	1.8	0.08	1.1340	7.2094	5.8645	17.4826



(a) Control input in Case 3. (b) Training return comparison of different RL algorithms.

Figure 8: Control effort and training performance in Case 3.

Table 9: Computational Cost Comparison in Case 3

Mode	Offline training time (min)	Peak memory (MB)	Online time per step (ms)
PFT-NC	—	—	0.0591
PFT-TD	—	—	0.1362
PFT-DDPG	12.5200	338.5200	0.5433
PFT-TD3	12.1000	349.0100	0.3177
PFT-PPO	1.4900	352.5500	0.3050
PFT-SAC	20.1500	352.4300	0.3533

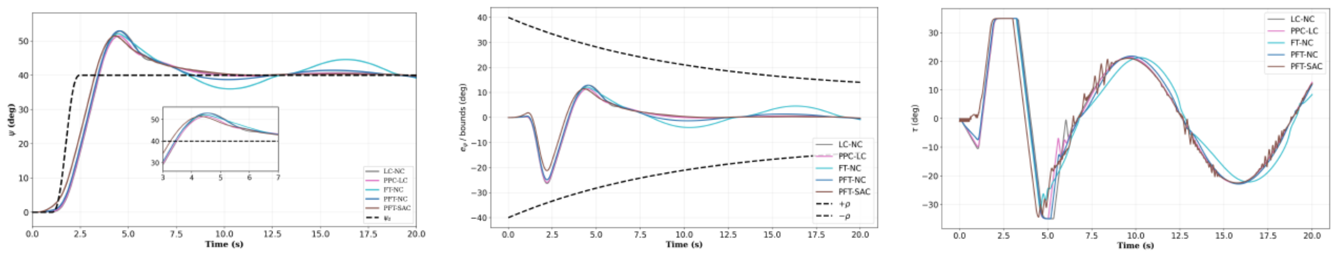
errors. A larger q_τ increases the penalty on rudder usage and may improve control smoothness, but it can also reduce the residual compensation capability and degrade tracking accuracy. The baseline setting provides a balanced performance in terms of mean tracking error, practical settling time, and RMS rudder input.

The training-return curves of different reinforcement learning residual compensators are shown in Figure **8**. The online computational burden is further summarized in Table **9**. The results show that the proposed PFT-SAC introduces additional offline training cost, but its online deployment cost remains lightweight because only the trained actor network is used during execution. The online computational time of PFT-SAC is approximately 0.3533 ms per simulation step, which is acceptable for the considered heading-control task.

5.7. Case 4: Ablation Study

To further clarify the individual contributions of the prescribed-performance constraint, the fixed-time recursive feedback, and the SAC-based residual compensation, an ablation study is conducted under the same single-heading transition condition as Case 3. Five controllers are compared: LC-NC, PPC-LC, FT-NC, PFT-NC, and PFT-SAC. Here, LC-NC denotes the baseline linear controller without PPC, fixed-time terms, or residual compensation; PPCLC adds the prescribed-performance transformation to the linear controller; FT-NC uses the fixed-time recursive terms without the complete PPC-SAC structure; PFT-NC denotes the proposed PFT backbone without residual compensation; and PFT-SAC denotes the complete proposed method.

The ablation results are presented in Figure **9** and Table **10**. Compared with LC-NC, PPC-LC reduces the mean tracking error from 1.3301° to 1.1932°, indicating that the prescribed-performance transformation improves the baseline transient regulation. The standalone fixed-time feedback case, FT-NC, produces a larger mean error of 2.6171°, which indicates that the fixed-time terms may lead to a more aggressive



(a) Heading response. (b) Tracking error and PPC bounds. (c) Rudder input.

Figure 9: Ablation results for PPC, fixed-time recursive feedback, and SAC residual compensation.

Table 10: Ablation Study in Case 4

Mode	Mean $\lvert e_\psi \rvert$ (deg)	Steady $\lvert e_\psi \rvert$ (deg)	Max $\lvert e_\psi \rvert$ (deg)	Ts (s)	RMS(τ) (deg)
LC-NC	1.3301	0.3769	8.7747	5.6400	18.1122
PPC-LC	1.1932	0.3008	8.3266	5.5000	17.6923
FT-NC	2.6171	2.4173	7.2598	18.6600	17.4493
PFT-NC	1.6424	0.6618	7.4906	17.2600	18.0513
PFT-SAC	0.9307	0.2290	6.4471	5.7000	17.6410

transient response when they are not coordinated with the complete PPC transformation and residual compensation mechanism.

When the complete PFT backbone is combined with the SAC residual compensator, PFT-SAC achieves the best overall ablation performance. Compared with PFT-NC, PFT-SAC reduces the mean tracking error from 1.6424° to 0.9307°, the maximum tracking error from 7.4906° to 6.4471°, and the practical settling time from 17.26 s to 5.70 s. Therefore, the ablation study verifies that the improvement is mainly obtained through the coordinated integration of PPC, fixed-time recursive feedback, and SAC-based residual compensation, rather than by any isolated module alone.

6. DISCUSSION

From the two simulation cases, several observations can be obtained. First, PFT-NC can maintain stable heading tracking due to the prescribed-performance fixed-time control backbone, but its tracking accuracy is limited in the presence of model uncertainties and external disturbances. Second, PFT-TD provides an ideal compensation benchmark by using true disturbance information; however, it is noncausal and cannot be directly implemented in practical applications. Third, the RL-based residual compensators improve the tracking performance without using ideal disturbance information during online evaluation.

Among the deployable RL-based methods, PFT-SAC achieves the smallest mean tracking error in both simulation cases. In Case 1, PFT-SAC provides the best average tracking accuracy, while PFT-TD3 and PFT-DDPG also show competitive transient responses. In Case 2, although PFT-DDPG obtains the smallest steady-state error and PFT-TD3 achieves the smallest maximum tracking error, PFT-SAC yields the lowest mean tracking error and the smallest RMS control input. These results indicate that the proposed SAC-based residual compensation provides a favorable tradeoff among tracking accuracy, control effort, and robustness.

Among the deployable RL-based methods, PFT-SAC generally provides the most favorable balance between tracking accuracy, convergence behavior, and control effort. In Case 2, PFT-SAC achieves the lowest mean tracking error and the smallest RMS rudder input in the multi-reference heading-tracking task. In Case 3, under increased uncertainty and stronger disturbance, PFT-SAC reduces the mean tracking error from 1.6424° to 0.9307° compared with PFT-NC, and shortens the practical settling time from 17.26 s to 5.70 s. The multi-seed statistical results further show that PFT-SAC achieves the lowest averaged mean tracking error, maximum tracking error, practical settling time, and RMS rudder input among the learning-based residual compensators.

The reward-sensitivity results indicate that the selected reward weights provide a balanced trade-off between tracking accuracy and rudder usage. The ablation study further confirms that the performance improvement is mainly obtained through the coordinated integration of prescribed-performance control, fixed-time recursive feedback, and SAC-based residual compensation, rather than by any isolated module alone. Overall, the proposed PFT-SAC controller enhances the practical heading-tracking performance of the prescribed-performance fixed-time controller while preserving the bounded residual-compensation structure.

7. CONCLUSION

This paper investigated the heading-tracking control problem of an unmanned surface vessel under model uncertainty, external disturbance, and actuator saturation. A prescribed-performance heading-control framework was developed by combining a prescribed-performance error transformation, a fixed-time-inspired recursive controller, and a bounded SAC-based residual compensation mechanism. In the proposed design, the model-based controller preserves the dominant closed-loop convergence structure, while the SAC module acts as an auxiliary residual compensator to improve practical tracking performance rather than to reconstruct the total lumped disturbance exactly.

A Lyapunov-based analysis was provided to show the boundedness of the closed-loop signals, the preservation of the prescribed-performance constraint, and the fixed-time convergence property under the stated bounded-operation assumptions.

Comparative simulations demonstrate the effectiveness of the proposed PFT-SAC method. In the multi-reference heading-tracking case, PFT-SAC reduces the mean heading tracking error from 1.1262° to 0.6891° compared with PFT-NC, corresponding to an improvement of approximately 38.8%. It also achieves the lowest RMS rudder input of 11.3358°, indicating a favorable balance between tracking accuracy and control effort. Under the increased uncertainty and stronger disturbance condition, PFT-SAC reduces the mean tracking error from 1.6424° to 0.9307°, reduces the steady-state error from 0.6618° to 0.2290°, and shortens the practical settling time from 17.26 s to 5.70 s. The ablation study further confirms that the improvement is mainly obtained from the coordinated integration of prescribed-performance control, fixed-time recursive feedback, and SAC-based residual compensation. In addition, the online execution time of PFT-SAC is approximately 0.3533 ms per step, showing that the proposed method remains computationally feasible for real-time heading-control implementation.

Future work will consider extending the proposed framework to more comprehensive path-following scenarios by combining the heading controller with LOS-based guidance laws and more strongly coupled surface vessel dynamics. Experimental validation on physical surface vessel platforms and automatic reward-weight tuning for broader operating conditions will also be pursued.

CONFLICT OF INTEREST:

There is no conflict of interest.

AUTHOR CONTRIBUTIONS

Deng Yingjie: conceptualization; Li Yong: writing-original draft; Huang Chenfeng: data curation; Liu Fangcheng: writing–review&editing; Im Namkyun: resources; Xu Yifei: supervision.

REFERENCES

[1] Tipsuwan Y, Sanposh P, Techajaroonjit N. Overview and control strategies of autonomous sailboats-a survey. Ocean Engineering 2023; 281: 114879. https://doi.org/10.1016/j.oceaneng.2023.114879

[2] Xiao L, Jouffroy J. Modeling and nonlinear heading control of sailing yachts. IEEE Journal of Oceanic Engineering 2014; 39(2): 256-268. https://doi.org/10.1109/JOE.2013.2247276

[3] Wen G, Ge SS, Chen CLP, Tu F, Wang S. Adaptive tracking control of surface vessel using optimized backstepping technique. IEEE Transactions on Cybernetics 2019; 49(9): 3420-3431. https://doi.org/10.1109/TCYB.2018.2844177

[4] Song B, Fan Y, Mu D, Wang G, Liu P, Wang Y. Global fixed-time path-following control with prescribed performance for an unmanned surface vehicle under lumped uncertainties and input saturation: Simulations and experiments. Ocean Engineering 2026; 351: 124024. https://doi.org/10.1016/j.oceaneng.2025.124024

[5] Deng Y, Zhang X, Zhang G. Line-of-sight-based guidance and adaptive neural path-following control for sailboats. IEEE Journal of Oceanic Engineering 2020; 45(4). https://doi.org/10.1109/JOE.2019.2923502

[6] Wang Y, Cao J, Sun J, Zou X, Sun C. Path following control for unmanned surface vehicles: A reinforcement learning-based method with experimental validation. IEEE Transactions on Neural Networks and Learning Systems 2024; 35(12): 18237-18250. https://doi.org/10.1109/TNNLS.2023.3313312

[7] Zhang Z, Liu M, Zhou J, Zheng R, Wei P. Hierarchical motion control framework based on slm-guided and learning-enhanced nmpc for autonomous underwater vehicles. Expert Systems with Applications 2026; 303: 130665. https://doi.org/10.1016/j.eswa.2025.130665

[8] Santos D, Silva Junior AG, Negreiros A, Vilas Boas J, Alvarez J, Araujo A, Aroca RV, Goncalves LMG. Design and implementation of a control system for a sailboat robot. Robotics 2016; 5(1): 5. https://doi.org/10.3390/robotics5010005

[9] Cruz NA, Alves JC. Auto-heading controller for an autonomous sailboat, in: Oceans'10 IEEE Sydney 2010; pp. 1-6. https://doi.org/10.1109/OCEANSSYD.2010.5603882

[10] Stelzer R, Proll T, John RI. Fuzzy logic control system for autonomous sailboats. in: 2007 IEEE International Fuzzy Systems Conference 2007; pp. 1-6. https://doi.org/10.1109/FUZZY.2007.4295347

[11] Deng Y, Zhang X, Zhang Q, Hu Y. Event-triggered composite adaptive fuzzy control of sailboat with heeling constraint. Ocean Engineering 2020; 211: 107627. https://doi.org/10.1016/j.oceaneng.2020.107627

[12] Xiao L, Jouffroy J. Modeling and nonlinear heading control of sailing yachts. IEEE Journal of Oceanic Engineering 2014; 39(2): 256-268. https://doi.org/10.1109/JOE.2013.2247276

[13] Deng Y, Zhang Z, Gong M, Ni T. Event-triggered asymptotic tracking control of underactuated ships with prescribed performance, IEEE Transactions on Intelligent Transportation Systems 2023; 24(1): 645-656. https://doi.org/10.1109/TITS.2022.3216808

[14] Cao L, Qin Y, Pan Y, Liang H. Prescribed performance-based optimal formation control for usvs with position constraints and yaw angle time-varying partial constraints. IEEE Transactions on Intelligent Transportation Systems 2025; 26(3): 4109-4121. https://doi.org/10.1109/TITS.2024.3520328

[15] Cao L, Cen Y, Li T, Liang H, Pan Y, Optimal fixed-time control for human-in-the-loop multiagent systems with actuator faults. IEEE Transactions on Cybernetics 2026; 56(3): 1773-1784. https://doi.org/10.1109/TCYB.2025.3632473

[16] Yang W, Wang Q-G, Liu J, Yu J. Actor-critic-based predefined-time fuzzy adaptive optimal control for uncertain nonlinear systems with input saturation. IEEE Transactions on Fuzzy Systems 2024; 32(4): 2448- 2457. https://doi.org/10.1109/TFUZZ.2024.3355129

[17] Zhang C, Guo G. Prescribed performance fault-tolerant control of nonlinear systems via actuator switching. IEEE Transactions On Fuzzy Systems 2024; 32(3): 1013-1022. https://doi.org/10.1109/TFUZZ.2023.3317017

[18] Wen G, Hao W, Feng W, Gao K. Optimized backstepping tracking control using reinforcement learning for quadrotor unmanned aerial vehicle system. IEEE Transactions on Systems, Man, and Cybernetics: Systems 2022; 52(8): 5004-5015. https://doi.org/10.1109/TSMC.2021.3112688

[19] Liu J, Xiao C, Yuan H, Li C, Li Q. Unmanned surface vessels path following control using improved transferring reinforcement learning: Simulation and experiment. Ocean Engineering 2025; 320: 120346. https://doi.org/10.1016/j.oceaneng.2025.120346

[20] Woo J, Yu C, Kim N. Deep reinforcement learning-based controller for path following of an unmanned surface vehicle. Ocean Engineering 2019; 183: 155-166. https://doi.org/10.1016/j.oceaneng.2019.04.099

[21] Zhong W, Li H, Meng Y, Yang X, Feng Y, Ye H, Liu W. Usv path following controller based on ddpg with composite state-space and dynamic reward function. Ocean Engineering 2022; 266: 112449. https://doi.org/10.1016/j.oceaneng.2022.112449

[22] Peng Z, Liu E, Pan C, Wang H, Wang D, Liu L. Model-based deep reinforcement learning for data-driven motion control of an under-actuated unmanned surface vehicle: Path following and trajectory tracking. Journal of the Franklin Institute 2023; 360(6): 4399-4426. https://doi.org/10.1016/j.jfranklin.2022.10.020

[23] Yoo B, Kim J. Path optimization for marine vehicles in ocean currents using reinforcement learning. Journal of Marine Science and Technology (Japan) 2016; 21(2): 334-343. https://doi.org/10.1007/s00773-015-0355-9

[24] Zheng Y, Tao J, Sun Q, Sun H, Chen Z, Sun M, Xie G. Soft actor-critic based active disturbance rejection path following control for unmanned surface vessel under wind and wave disturbances. Ocean Engineering 2022; 247: 110631. https://doi.org/10.1016/j.oceaneng.2022.110631

[25] Li C, Zhang B, Cao Y, Yin B. Reinforcement learning-based backstepping sliding mode heading control for unmanned sailboats. Ocean Engineering 2025; 327: 120936. https://doi.org/10.1016/j.oceaneng.2025.120936

[26] Fan Y, Dong H, Zhao X, Denissenko P. Path-following control of unmanned underwater vehicle based on an improved td3 deep reinforcement learning. IEEE Transactions on Control Systems Technology 2024; 32(5): 1904-1919. https://doi.org/10.1109/TCST.2024.3377876

[27] Qin H, Xu P, Zhang F, Xue Y. An improved los guidance and sail-rudder cooperative controller for the path tracking of unmanned sailboats. IEEE Transactions on Intelligent Vehicles 2024; 9(10): 6610-6623. https://doi.org/10.1109/TIV.2024.3386964