Journal of Marine Science and Engineering

MDPI

Article

Event-Based Robust Reinforcement Learning Course Tracking for MSVs with Actuator Faults

Xiaoyang Gao [1], Lirong Kou [1], Dakuan Song [1,*], Quan Shi [2] and Xin Hu [3]

[1] School of Maritime Economics and Management, Dalian Maritime University, Dalian 116026, China; xiaoyang_gao@yeah.net (X.G.)
[2] Artificial Intelligence Lab, CASCO Signal Ltd., Shanghai 200072, China
[3] School of Mathematics and Statistics Science, Ludong University, Yantai 264025, China
* Correspondence: sdk1231@live.com

Abstract

This paper investigates the robust course control for marine surface vehicles (MSVs) by using reinforcement learning and a dynamic event-triggered mechanism under rudder faults. A fault-tolerant optimized tracking control scheme is designed based on reinforcement learning. Firstly, a general actuator fault model is established. Secondly, internal and external disturbances as well as fault dynamics are aggregated into lumped uncertainties, and a fault observer is constructed to estimate and compensate the lumped uncertainty term. Then, a tracking cost function for the MSV system is constructed, and a reinforcement learning framework with a single critic network is used to approximate the value function and generate a learning-based compensation input. Finally, a dynamic event-triggered mechanism is implemented between the control layer and the execution layer to reduce the triggering frequency of actuators. Lyapunov stability theory is utilized to prove that all signals within the closed-loop system are uniformly ultimately bounded. Simulation results verify that the proposed control scheme achieves favorable overall performance.

Keywords: marine surface vehicles; fault-tolerant control; reinforcement learning; dynamic event-triggered

1. Introduction

Marine surface vehicles (MSVs) are increasingly required to undertake navigation tasks with higher levels of autonomy in long-distance transportation, environmental monitoring, maritime patrol, and offshore operations [1–3]. Among the basic motion-control functions of an MSV, automatic course control directly determines whether the vessel can maintain or follow a prescribed heading command in the presence of continuously changing operating conditions. Ship-steering dynamics are characterized by large inertia, pronounced nonlinearities, uncertain hydrodynamic coefficients, and persistent environmental loads induced by wind, waves, and currents. Moreover, long-term operation of steering gears may introduce loss of effectiveness, bias, or other actuator abnormalities, which can further deteriorate course-tracking accuracy and even threaten navigational safety. Meanwhile, frequent rudder actions increase steering-gear activity, mechanical wear, and unnecessary control expenditure. Therefore, a course-control system should not only preserve tracking accuracy and robustness against uncertainties and actuator faults but also account for control performance and actuator utilization [4,5].

A variety of control techniques have been developed for ship autopilots. Owing to their transparent structure and convenient implementation, PID controllers and their

Check for updates

Academic Editors: Graciliano Nicolás Marichal and Weicheng Cui

Received: 26 August 2026
Revised: 19 October 2026
Accepted: 28 October 2026
Published: 5 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

variants remain attractive for ship-steering applications. Moving beyond linear feedback, adaptive and nonlinear control techniques have been introduced to cope with uncertain ship dynamics. Skjetne et al. [6] developed an adaptive maneuvering control framework and experimentally validated it using a model ship, demonstrating the capability of nonlinear adaptive control in marine applications. Zhang et al. [7] combined backstepping with neural networks for nonlinear ship course-keeping, where neural approximators and adaptive laws were employed to compensate for unknown nonlinearities and environmental disturbances. These studies substantially improve robustness and tracking performance. Nevertheless, most conventional control strategies are constructed primarily from stability and tracking-error considerations, while the compromise between tracking accuracy and control effort is not explicitly characterized by a performance criterion [8,9].

In the modern day, a wide range of optimization methods have been applied to ship systems [10–13]. Reinforcement learning (RL) provides an alternative route for nonlinear optimal control by approximating the solution of the Hamilton–Jacobi–Bellman (HJB) equation without requiring its explicit analytical solution [14,15]. Zhang et al. [16] subsequently developed a data-driven approximate optimal tracking framework in which the optimal tracking problem was addressed through RL. These developments stimulated the application of learning-based optimization to marine control systems. In particular, Wen et al. [17] incorporated optimization into the recursive backstepping procedure for vessel tracking and used actor-critic learning structures to approximate the optimal virtual and actual control policies. Chen et al. [18] further investigated adaptive optimal tracking of an underactuated surface vessel by integrating backstepping and actor-critic reinforcement learning in the presence of uncertain dynamics and external disturbances. Although the aforementioned methods verified the application potential of RL in improving tracking performance, adaptively adjusting the actor-critic network simultaneously introduces additional network weights and update rules, which increases structural complexity. More critically, if actuator performance degradation is taken into account, the difficulty of optimal heading control for MSVs will rise significantly.

Actuator faults constitute another important issue in autonomous marine control [19–21]. Zhu et al. [22] studied tracking control of MSVs subject to actuator faults, input saturation, parameter uncertainties, and environmental disturbances, and developed an appointed-time fault-tolerant tracking strategy. In addition to fault accommodation, reducing unnecessary control updates has become increasingly relevant because continuously refreshed control commands may result in excessive actuator activity. Event-triggered control addresses this issue by updating the transmitted control signal only when a prescribed condition is satisfied. Zhu et al. [23] developed an event-triggered adaptive neural fault-tolerant controller for underactuated MSVs with actuator faults and saturation, thereby reducing actuator execution while maintaining tracking stability. More recently, Shen et al. [24] combined ADP with a dynamic event-triggered mechanism for constrained optimal tracking of MSVs. These results provide important progress toward resource-efficient intelligent marine control. However, fault compensation, optimized course regulation, and dynamic event-triggered actuator operation are often treated separately or require relatively complicated learning architectures. In particular, simultaneously accounting for rudder faults, lumped internal and external uncertainties, optimal course-tracking performance, and actuator execution frequency using a compact single-critic learning architecture remains a problem worthy of further investigation. As summarized in Table 1, unlike representative existing methods that address only subsets of performance awareness, fault tolerance, disturbance rejection, and dynamic event-triggered implementation, the proposed scheme integrates all four capabilities within a unified framework, highlighting its methodological novelty and comprehensive control performance.

Table 1. Different control performance comparison.

Control Method	Performance-Aware	Fault-Tolerant	Disturbance Rejection	Dynamic ET
[17]	✓			
[14]	✓		✓	
[19]		✓	✓	✓
[23]		✓		✓
[24]	✓			✓
Proposed Method	✓	✓	✓	✓

✓ indicates that the corresponding control method possesses the corresponding property.

Motivated by the above observations, this paper develops an event-based optimized course-tracking scheme for MSVs subject to actuator faults. The main contributions are summarized as follows.

1. A fault-tolerant course-tracking controller is developed under the RL framework. In contrast to conventional actor-critic implementations requiring separate policy and value networks [14,16–18], only a critic network is employed to approximate the value function, from which the approximate optimal control input is obtained. The compensation policy is constructed from the estimated value-function gradient, avoiding the need for a separate actor network. This provides a compact learning structure for the proposed control scheme.
2. A dynamic event-triggered mechanism is inserted between the control layer and the actuator execution layer. The continuously generated optimized command is therefore transmitted to the actuator only when the triggering condition is activated, reducing unnecessary rudder updates and execution frequency.

2. Problem Formulations and Preliminaries

2.1. Problem Formulation

The model of the MSV course-control system is shown as follows [25]:

$$\begin{cases} \dot{\varphi} = r \\ \dot{r} = -a_1 r - a_2 r^3 - a_3 r^5 + g\delta_f + \Delta \end{cases} \tag{1}$$

where φ is the heading angle, and r is the yaw velocity. $g = K_0/T_0$ is the gain of the control input. K_0 is the ship turning index and T_0 is the following index. δ_f is the faulty rudder. Δ is the unknown ocean disturbance affect on the rudder. a_i is the unknown coefficient of r. φ_r is the desired signal.

2.2. Fault Rudder Model of the MSV

The fault model of the rudder can be written as [26]

$$\delta_f = \kappa(t)\delta + \rho(t) \tag{2}$$

where $0 \leq \kappa(t) \leq 1$ and $\rho(t)$ denote the rudder loss efficiency factor and bounded additive fault signal. It can be assumed that $|\dot{\kappa}(t)| \leq \kappa_m$ is bounded and $\rho(t)$ is an unknown varying constant. Based on $\kappa(t)$ and $\rho(t)$, there are three conditions:

- Fault-free: When $\kappa(t) = 1$ and $\rho(t) = 0$, there exists $\delta_f = \delta$.
- Loss-of-effectiveness fault: When $0 < \kappa(t) < 1$ and $\rho(t) = 0$, there exists $\delta_f = \kappa(t)\delta$.
- Bias fault: When $\rho(t) \neq 0$, there exists $\delta_f = \kappa(t)\delta + \rho(t)$.

Lemma 1. *For any given continuous function $\sigma(\chi)$ with $\sigma(0) = 0$, defined in a compact set $\Omega_\chi \subseteq \mathcal{R}$, $\sigma(\chi)$ can be approximated by the NN with arbitrary precision [17], namely,*

$$\sigma(\chi) = \theta^\top \vartheta(\chi) + \epsilon \tag{3}$$

where θ denotes the weight vector of the neural networks. $\vartheta(\chi)$ is the basis function. ϵ is the bounded approximation error.

3. Event-Based RL Course-Tracking Design

3.1. Fault-Tolerant Tracking Control Design

Establish the following error surfaces as

$$\varepsilon_\varphi = \varphi - \varphi_r \tag{4}$$

$$\varepsilon_r = r - \alpha \tag{5}$$

where α denotes the virtual optimal course signal, which can be defined as $\alpha = \alpha_a + \alpha_b$.

Step 1: Differentiating the first error surface (4) and using $r = \varepsilon_r + \alpha$ gives

$$\begin{aligned} \dot{\varepsilon}_\varphi &= \dot{\varphi} - \dot{\varphi}_r \\ &= \varepsilon_r + \alpha_a + \alpha_b - \dot{\varphi}_r \end{aligned} \tag{6}$$

Select the candidate for the Lyapunov function in the following:

$$V_\varphi = \frac{1}{2}\varepsilon_\varphi^2 \tag{7}$$

Taking the time derivative of (7) gives

$$\dot{V}_\varphi = \varepsilon_\varphi(\varepsilon_r + \alpha_a + \alpha_b - \dot{\varphi}_r) \tag{8}$$

Design the virtual control as

$$\alpha_a = -c_\varphi \varepsilon_\varphi + \dot{\varphi}_r \tag{9}$$

where c_φ is a positive constant to be designed.

Substituting (9) into (8) yields

$$\dot{V}_\varphi = -c_\varphi \varepsilon_\varphi^2 + \varepsilon_\varphi \varepsilon_r + \varepsilon_\varphi \alpha_b. \tag{10}$$

This expression shows that the nominal virtual control α_a introduces the stabilizing term $-c_\varphi \varepsilon_\varphi^2$, whereas α_b is retained as the RL-based compensation input.

Step 2: Taking the derivative of the second error surface (5) gives

$$\begin{aligned} \dot{\varepsilon}_r &= \dot{r} - \dot{\alpha} \\ &= f(r) + g(\kappa(t)\delta + \rho(t)) + \Delta - \dot{\alpha} \\ &= \delta + [f(r) - f(\alpha)] + f(\alpha) + (g\kappa - 1)\delta + g\rho + \Delta - \dot{\alpha} \\ &= \delta + o(\gamma) + \theta_a^\top \vartheta_a(\alpha) + \epsilon_a + (g\kappa - 1)\delta + g\rho + \Delta - \dot{\alpha} \end{aligned} \tag{11}$$

where $f(r) = -a_1 r - a_2 r^3 - a_3 r^5$. In the third line of (11), $f(\alpha)$ and δ are added and subtracted to separate the known input channel from the nonlinear uncertainty. Define $\gamma = [r, \alpha]^\top$, $o(\gamma) = f(r) - f(\alpha)$, and $f(\alpha) = \theta_a^\top \vartheta_a(\alpha) + \epsilon_a$, where θ_a is the ideal NN weight vector, $\vartheta_a(\alpha)$ is the basis-function vector, and ϵ_a is the bounded approximation error. The composite uncertainty is then defined as

$$\Theta = \Delta + (g\kappa - 1)\delta + \epsilon_a + g\rho \tag{12}$$

The lumped term Θ is estimated online using the auxiliary observer state ν, as follows:

$$\begin{cases} \nu = \hat{\Theta} - b_1 r \\ \dot{\nu} = -b_1 \nu - b_1 \left(\hat{f}(r) + \delta + b_1 r \right) \end{cases} \tag{13}$$

where $b_1 > 0$ is the observer gain and $\hat{f}(r)$ denotes the available NN estimate of $f(r)$. From the first equation in (13), the lumped-uncertainty estimate is reconstructed as $\hat{\Theta} = \nu + b_1 r$.

The stability of the designed observer can be proved by [19]. Following the backstepping structure in Step 2, the stabilizing controller and its adaptation law are designed as

$$\delta_a = -c_r \varepsilon_r - \varepsilon_\varphi - \hat{\theta}_a^\top \vartheta_a(\alpha) + \dot{\alpha} - \hat{\Theta} \tag{14}$$

$$\dot{\hat{\theta}}_a = \varepsilon_r \vartheta_a(\alpha) - c_\theta \hat{\theta}_a \tag{15}$$

where $c_r > 0$ and $c_\theta > 0$ are design parameters. Define the NN weight error as $\tilde{\theta}_a = \theta_a - \hat{\theta}_a$.

Substituting $\delta = \delta_a + \delta_b$, (12), and (14) into (11) gives the explicit second-error dynamics

$$\dot{\varepsilon}_r = -c_r \varepsilon_r - \varepsilon_\varphi + \delta_b + o(\gamma) + \tilde{\theta}_a^\top \vartheta_a(\alpha). \tag{16}$$

This equation makes explicit the stabilizing feedback terms, the RL input δ_b, the NN weight error, and the residual nonlinear/observer errors.

Choose the augmented Lyapunov function

$$V_r = \frac{1}{2}\varepsilon_\varphi^2 + \frac{1}{2}\varepsilon_r^2 + \frac{1}{2}\tilde{\theta}_a^\top \tilde{\theta}_a \tag{17}$$

Taking the derivative of (17), using the Step-1 result, and substituting the second-error dynamics above yields

$$\begin{aligned} \dot{V}_r = & -c_\varphi \varepsilon_\varphi^2 - c_r \varepsilon_r^2 + \varepsilon_\varphi \alpha_b + \varepsilon_r \delta_b \\ & + \varepsilon_r o(\gamma) + \varepsilon_r \tilde{\theta}_a^\top \vartheta_a(\alpha) - \tilde{\theta}_a^\top \dot{\hat{\theta}}_a. \end{aligned} \tag{18}$$

Using (15), the two terms containing $\varepsilon_r \tilde{\theta}_a^\top \vartheta_a(\alpha)$ cancel. In addition,

$$c_\theta \tilde{\theta}_a^\top \hat{\theta}_a = c_\theta \tilde{\theta}_a^\top (\theta_a - \tilde{\theta}_a) \le \frac{c_\theta}{2}\|\theta_a\|^2 - \frac{c_\theta}{2}\|\tilde{\theta}_a\|^2. \tag{19}$$

The derivative of V_r therefore satisfies

$$\begin{aligned} \dot{V}_r \le & -c_\varphi \varepsilon_\varphi^2 - c_r \varepsilon_r^2 + \frac{1}{2} c_\theta \|\theta_a\|^2 - \frac{1}{2} c_\theta \|\tilde{\theta}_a\|^2 \\ & + \mathcal{E}^T \begin{bmatrix} 0 \\ O(\gamma) \end{bmatrix} + \mathcal{E}^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathcal{U}_b \end{aligned} \tag{20}$$

where $\mathcal{E} = [\varepsilon_\varphi, \varepsilon_r]^\top$ and $\mathcal{U}_b = [\alpha_b, \delta_b]^\top$. Equation (20) clearly separates the negative feedback terms, the bounded residual term, and the RL compensation input. According to [26], the remaining task is to design $\mathcal{U}_b$ so that the error system is stabilized while the prescribed performance index is minimized.

3.2. RL-Based Compensation Design

Based on the residual and compensation terms identified in (20), the subsystem optimized by RL is written compactly as

$$\dot{\mathcal{E}} = F(\mathcal{E}) + \mathcal{U}_b \tag{21}$$

where $F(\mathcal{E}) = [0, O(\gamma)]^\top$ represents the bounded residual channel after the nominal backstepping terms have been introduced. The cost function is defined as

$$\mathcal{J} = \int_t^\infty (\mathcal{E}^\top \mathcal{Q} \mathcal{E} + \mathcal{U}_b^\top \mathcal{R} \mathcal{U}_b) dt \tag{22}$$

where $\mathcal{Q}$ and $\mathcal{R}$ are positive definite matrices.

For an admissible initial control $\overline{\mathcal{U}}_b$, the Hamiltonian function is defined as

$$\mathcal{H} = \mathcal{E}^\top \mathcal{Q}\mathcal{E} + \overline{\mathcal{U}}_b^\top \mathcal{R}\overline{\mathcal{U}}_b + \left(\nabla_\mathcal{E}\mathcal{J}\right)^\top \left(F(\mathcal{E}) + \overline{\mathcal{U}}_b\right) \tag{23}$$

where $\nabla_\mathcal{E}\mathcal{J}$ denotes the column gradient of $\mathcal{J}$ with respect to $\mathcal{E}$. Applying the stationary condition $\partial\mathcal{H}/\partial\mathcal{U}_b = 0$ gives

$$\mathcal{U}_b^* = -\frac{1}{2}\mathcal{R}^{-1}\nabla_\mathcal{E}\mathcal{J}^* \tag{24}$$

Substitution of (24) into the Hamiltonian yields the HJB equation

$$0 = \mathcal{E}^\top \mathcal{Q}\mathcal{E} + \left(\nabla_\mathcal{E}\mathcal{J}^*\right)^\top F(\mathcal{E}) - \frac{1}{4}\left(\nabla_\mathcal{E}\mathcal{J}^*\right)^\top \mathcal{R}^{-1}\nabla_\mathcal{E}\mathcal{J}^* \tag{25}$$

The NN is used to approximate the critic term

$$\begin{aligned}\mathcal{J} &= \theta_b^\top \vartheta_b(\mathcal{E}) + \epsilon_b \\ &= \begin{bmatrix} \theta_{b1}^\top & \cdots & 0_{1\times n} \\ \vdots & \ddots & \vdots \\ 0_{1\times n} & \cdots & \theta_{bn}^\top \end{bmatrix}^\top \begin{bmatrix} \vartheta_{b1} \\ \vdots \\ \vartheta_{bn} \end{bmatrix} + \begin{bmatrix} \epsilon_{b1} \\ \vdots \\ \epsilon_{bn} \end{bmatrix}\end{aligned} \tag{26}$$

where θ_b is the adaptive parameter of the critic term and ϵ_b is the approximation error.

Hence, the approximations of the optimal term are shown as

$$\hat{\mathcal{U}}_b = -\frac{1}{2}\mathcal{R}^{-1}\nabla\vartheta_\varepsilon^T\hat{\theta}_b \tag{27}$$

with the update principle as

$$\begin{aligned}\dot{\hat{\theta}}_b = -c_b &\begin{bmatrix} (\nabla\vartheta_{b\varepsilon}^\top)^\top \hat{F}(\mathcal{E}) \\ -\frac{1}{2}(\nabla\vartheta_{b\varepsilon}^\top)^\top \mathcal{R}^{-1}(\nabla\vartheta_{b\varepsilon}^\top)\hat{\theta}_b \end{bmatrix} \\ &\times \begin{bmatrix} \mathcal{E}^\top\mathcal{Q}\mathcal{E} + \hat{\theta}_b^\top(\nabla\vartheta_{b\varepsilon}^\top)^\top \hat{F}(\mathcal{E}) \\ -\frac{1}{4}\hat{\theta}_b^\top(\nabla\vartheta_{b\varepsilon}^\top)^\top \mathcal{R}^{-1}(\nabla\vartheta_{b\varepsilon}^\top)\hat{\theta}_b \end{bmatrix}\end{aligned} \tag{28}$$

where c_b is the NN learning rate parameter. The optimal fault-tolerant tracking design can be shown as $\delta = \delta_a + \hat{\delta}_b$.

3.3. Event-Based Mechanism

To reduce excessive wear of the rudder and save communication resources in the control–actuator link, this paper proposes a event-based mechanism to enhance the effect and scope of optimal control. The proposed mechanism is expressed as follows:

$$\begin{cases} E_\delta(t) = \hat{\delta}(t_k), \forall t \in [t_k, t_{k+1}) \\ t_{k+1} = \inf\{t > t_k \mid \lambda\zeta(t) \le |e_\delta(t)|\} \\ \dot{\zeta}(t) = -\lambda\zeta(t) + \varpi|e_\delta(t)|, \end{cases} \tag{29}$$

where $e_\delta(t) = \delta(t) - \hat{\delta}(t_k)$, $\zeta(0) > 0$ is the initial value of the internal dynamic variable, $\lambda > 0$ is the threshold attenuation coefficient, and $\varpi \ge 0$ is a design parameter. E_δ is the actuator input held between two consecutive triggering instants.

3.4. Stability Analysis

Theorem 1. *Consider the MSV course-tracking system subject to the actuator faults in (2). Under the fault-tolerant controller (14), the RL-based compensation controller (27), the critic updating law (28), and the dynamic event-triggered mechanism (29), suppose that the desired course signal and its required derivatives are bounded, the NN approximation errors are bounded, and the critic regressor*

satisfies the persistent excitation condition. Then, by appropriately selecting the controller gains and the critic learning gain, all signals in the resulting closed-loop system are uniformly ultimately bounded. Moreover, the triggering sequence generated by (29) has no finite accumulation point; therefore, Zeno behavior is excluded.

Proof. Choose the sum of the Lyapunov function candidate as

$$V_{all} = V_r + \frac{1}{2}\tilde{\theta}_b^\top \tilde{\theta}_b \tag{30}$$

According to the error transformation and controller (14), the derivative of V_r can be rearranged as

$$\begin{aligned} \dot{V}_r \leq & -c_\varphi \varepsilon_\varphi^2 - c_r \varepsilon_r^2 - \frac{1}{2} c_\theta \left\|\tilde{\theta}_a\right\|^2 + \frac{1}{2} c_\theta \|\theta_a\|^2 \\ & + \mathcal{E}^\top \left[F(\mathcal{E}) + \hat{\mathcal{U}}_b\right] + D_a \end{aligned} \tag{31}$$

D_a collects the bounded NN approximation error, residual fault-observation error and other bounded uncertainty terms.

From the ideal cost function

$$\mathcal{J}^*(\mathcal{E}) = \theta_b^{*\top} \vartheta_b(\mathcal{E}) + \epsilon_b(\mathcal{E}), \tag{32}$$

The ideal optimal compensation term can be obtained as

$$\mathcal{U}_b^* = -\frac{1}{2} \mathcal{R}^{-1} (\nabla \vartheta_b)^\top \theta_b^* + \epsilon_u \tag{33}$$

where $\epsilon_u = -\frac{1}{2}\mathcal{R}^{-1}\nabla\epsilon_b$ represents the error caused by the approximation of the critic function.

Consider the actual control input as (27). Therefore, it has

$$\hat{\mathcal{U}}_b - \mathcal{U}_b^* = -\frac{1}{2} \mathcal{R}^{-1} (\nabla \vartheta_b)^\top \tilde{\theta}_b - \epsilon_u \tag{34}$$

Based on the universal approximation theorem of neural networks and reference [24], we can assume that $\|\nabla\vartheta_b\| \leq \bar{\vartheta}_b$, $\|\epsilon_b\| \leqslant \epsilon_{bM}$, $\|\theta_b\| \leqslant \theta_{bM}$, and $\|\epsilon_u\| \leq \bar{\epsilon}_u$.

Define

$$l_b = \frac{1}{2} \left\|\mathcal{R}^{-1}\right\| \bar{\vartheta}_b \tag{35}$$

Then, we yield

$$\left\|\hat{\mathcal{U}}_b - \mathcal{U}_b^*\right\| \leq l_b \left\|\tilde{\theta}_b\right\| + \bar{\epsilon}_u \tag{36}$$

Hence,

$$\mathcal{E}^\top \left(\hat{\mathcal{U}}_b - \mathcal{U}_b^*\right) \leq l_b \left\|\tilde{\theta}_b\right\| \|\mathcal{E}\| + \bar{\epsilon}_u \|\mathcal{E}\| \tag{37}$$

Applying Young's inequality gives, for arbitrary constants,

$$l_b \left\|\tilde{\theta}_b\right\| \|\mathcal{E}\| \leq \frac{1}{2}\eta_1 \|\mathcal{E}\|^2 + \frac{l_b^2}{2\eta_1} \left\|\tilde{\theta}_b\right\|^2 \tag{38}$$

and

$$\bar{\epsilon}_u \|\mathcal{E}\| \leq \frac{1}{2}\eta_2 \|\mathcal{E}\|^2 + \frac{1}{2\eta_2}\bar{\epsilon}_u^2 \tag{39}$$

where $\eta_1, \quad \eta_2 > 0$.

In summary, it can be obtained that

$$\begin{aligned} \dot{V}_{all} \leq & -\left(c_0 - \frac{\eta_1 + \eta_2}{2}\right) \|\mathcal{E}\|^2 \\ & - \frac{c_\theta}{2} \|\bar{\theta}_a\|^2 - \frac{l_b^2}{2\eta_1} \|\bar{\theta}_b\|^2 + D_r \end{aligned} \tag{40}$$

where $D_r = D_a + \frac{1}{2\eta_2}\bar{\epsilon}_u^2$.

The derivative bound in Equation (40) can equivalently be written in the transpose-times-vector form as

$$\dot{V}_{\text{all}} \leq -\left(c_0 - \frac{\eta_1 + \eta_2}{2}\right)\mathcal{E}^\top\mathcal{E} - \frac{c_\theta}{2}\tilde{\theta}_a^\top\tilde{\theta}_a - \frac{l_b^2}{2\eta_1}\tilde{\theta}_b^\top\tilde{\theta}_b + D_r. \tag{41}$$

Since $V_{\text{all}} = \frac{1}{2}\mathcal{E}^\top\mathcal{E} + \frac{1}{2}\tilde{\theta}_a^\top\tilde{\theta}_a + \frac{1}{2}\tilde{\theta}_b^\top\tilde{\theta}_b$, the preceding inequality gives

$$\dot{V}_{\text{all}} \leq -\lambda V_{\text{all}} + D_r, \tag{42}$$

where $\lambda = \min\left\{2c_0 - (\eta_1 + \eta_2),\ c_\theta,\ \frac{l_b^2}{\eta_1}\right\} > 0$.

Thus, Equation (40) can also be written as

$$\dot{V}_{\text{all}} \leq -\lambda V_{\text{all}} + D_r. \tag{43}$$

where $\lambda = \min\{2c_0 - (\eta_1 + \eta_2), c_\theta, \frac{l_b^2}{\eta_1}\} > 0$.

Since the desired course signal and its derivatives are bounded, the system states, adaptive parameters, and optimized commands are bounded. Consequently, all signals in the closed-loop system are UUB.

It remains to prove that the triggering mechanism is free of Zeno behavior. From (29), $\varpi \geq 0$, and $|e_\delta(t)| \geq 0$, one has

$$\dot{\zeta}(t) \geq -\lambda\zeta(t), \tag{44}$$

which, together with $\zeta(0) > 0$, yields

$$\zeta(t) \geq \zeta(0)e^{-\lambda t} > 0. \tag{45}$$

According to the preceding boundedness result, the continuously generated rudder command $\delta(t)$ evolves in a compact set. Because the closed-loop vector field, the selected basis functions, and the desired signal are smooth with bounded required derivatives, there exists a constant $\bar{d}_\delta > 0$ such that $|\dot{\delta}(t)| \leq \bar{d}_\delta$. At each triggering instant, the newly computed command is transmitted to the actuator, so that $\hat{\delta}(t_k) = \delta(t_k)$ and $e_\delta(t_k) = 0$. Therefore, for all $t \in [t_k, t_{k+1})$,

$$|e_\delta(t)| = \left|\int_{t_k}^{t} \dot{\delta}(s)\,\mathrm{d}s\right| \leq \bar{d}_\delta(t - t_k). \tag{46}$$

Suppose, by contradiction, that infinitely many triggering instants accumulate at a finite time T_z. For every $t \in [0, T_z]$, the dynamic threshold satisfies

$$\lambda\zeta(t) \geq \lambda\zeta(0)e^{-\lambda T_z} =: \underline{\mu}_z > 0. \tag{47}$$

If a subsequent event occurs, the continuity of $e_\delta(t)$ and $\zeta(t)$ implies that the triggering boundary is reached at t_{k+1}, namely $|e_\delta(t_{k+1}^-)| = \lambda\zeta(t_{k+1})$. It follows that

$$t_{k+1} - t_k \geq \frac{\underline{\mu}_z}{\bar{d}_\delta} > 0. \tag{48}$$

Thus, all inter-event times before T_z possess a strictly positive lower bound, contradicting the assumed finite accumulation of triggering instants. If no subsequent event occurs, Zeno behavior is trivially impossible. Hence, the proposed dynamic event-triggered control scheme is Zeno-free. □

4. Simulation Results and Discussion

4.1. Simulation Results

To validate the proposed fault-tolerant RL control strategy, numerical studies are conducted using the sea-going training vessel *Yulong* as the simulation platform, with its parameters found in [25].

In the simulations, the initial states are chosen as $\psi_0 = 5^\circ$ and $r_0 = 0.01^\circ/\mathrm{s}$; the observer state and parameter are designed as $\nu_0 = 0.1$ and $b_1 = 10$. For the optimal compensation term design, select $\mathcal{Q} = 5$ and $\mathcal{R} = 2$. The parameters are selected as $c_\varphi = 1.5$ and $c_r = 5$. The initial adaptive parameter and weight are chosen randomly in $(0, 1)$. The adaptive parameter is set as $c_\theta = 0.02$ and the learning rate is $c_b = 0.01$. Gaussian basis functions are selected as the basis functions of the neural network. In addition, we take into account the practical requirement that the rudder angle shall not exceed $\pm 35^\circ$, and impose constraints on the control input.

To support persistent excitation (PE), the excitation signal added to the rudder-command channel is specified as

$$p(t) = s_e(t)[0.5\sin(0.05t) + 0.3\sin(0.13t) + 0.2\sin(0.23t)]^\circ \tag{49}$$

with $s_e(t) = \begin{cases} [1 - \cos(\pi t/10)]/2, 0 \le t < 10\ \mathrm{s} \\ 1, t \ge 10\ \mathrm{s} \end{cases}$. The minimum value within $minp(t)$ in 0–1200 s of the computation window is -0.714°. The eigenvalues, i.e., the window energy, can be expressed by the following formula:

$$G_p(t) = \int_t^{t+T} p^2(\tau)\,\mathrm{d}\tau.$$

Window length is $T = 250$ s; we can thus calculate $\lambda_{min}(G_p(t))$.

We select autopilot tracking signals that reflect specific actual performance requirements as follows:

$$\ddot{\varphi}_m(t) + 0.1\dot{\varphi}_m(t) + 0.0025\varphi_m(t) = 0.0025\varphi_r(t). \tag{50}$$

Among them, φ_m represents the ideal system performance characterizing the heading; φ_r is the command input signal, whose value ranges from 0° to 30°.

The external disturbance is chosen as

$$\Delta = 0.025\sin(0.05t) + 0.015\cos(0.03t). \tag{51}$$

To verify the effectiveness of the algorithm proposed in this paper, two cases are used for specific demonstration. Among them, Case 1 and Case 2 demonstrate the performance under moderate magnitude faults and higher magnitude faults, respectively. This is shown as

$$\delta_{f1} = \begin{cases} \delta, & 0 < t \le 600\ \mathrm{s} \\ 0.6|\sin(0.02t)| \times \delta + 5\sin(0.1t), & t \ge 600\ \mathrm{s} \end{cases} \tag{52}$$

$$\delta_{f2} = \begin{cases} \delta, & 0 < t \le 600\ \mathrm{s} \\ 0.4|\sin(0.02t)| \times \delta + 7\sin(0.1t). & t \ge 600\ \mathrm{s} \end{cases} \tag{53}$$

It is particularly pointed out that in order to ensure the validity of the algorithm comparison, the parameters and initial values in these control experiments are designed to be basically consistent.

Figures 1–8 and Figures 9–16 present the simulation results for Cases 1 and 2, respectively. In Case 1, the actual course closely follows the desired profile, while the yaw-rate and tracking-error responses exhibit only short transients near the command changes and

remain close to zero during the steady portions. The constrained rudder input remains within the prescribed range, and the estimated lumped uncertainty nearly overlaps its actual value over the simulated interval. The event-triggering instants are nonuniform, and the accumulated number of updates is lower than that of the time-triggered implementation; meanwhile, the adaptive parameters and critic weights remain bounded. In Case 2, the course is also tracked over most of the simulation, but a visibly larger transient occurs at approximately 1000 s under the higher-magnitude fault. This transient is accompanied by larger excursions in the yaw rate, tracking error, rudder demand, and adaptive parameter, after which these responses decrease toward their previous levels. The observer still follows the dominant variation of the lumped uncertainty, and the event-triggered updates remain intermittent. Therefore, the two cases indicate bounded closed-loop responses and reduced update activity under the tested conditions, while the Case 2 results also show that the transient tracking accuracy deteriorates when the fault magnitude is increased. Figure 17 shows an overall inward-winding trend of the error trajectories toward a neighborhood of the origin in both cases, with larger intermediate excursions in Case 2.

The simulation results for Case 2 are presented in Figures 9–16, following the same order as the results for Case 1.

4.2. Discussion

Theoretical guarantees. The Lyapunov analysis establishes uniform ultimate boundedness of the closed-loop signals, while the event-triggered analysis excludes Zeno behavior by deriving a positive lower bound on the inter-event time. These are analytical results under the assumptions stated in this manuscript, including bounded signals, approximation errors, and sufficient control authority.

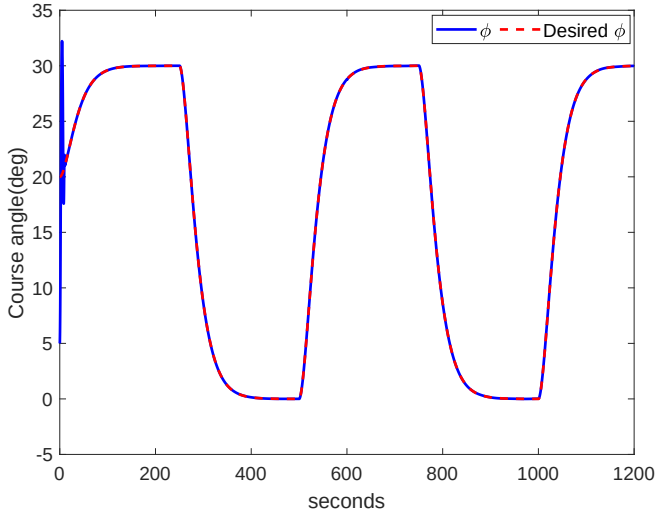


Figure 1. Course-tracking performance of the RL fault-tolerant control scheme in Case 1.

Numerical performance and trade-offs. The two simulation cases show that the course is tracked under the tested fault conditions, but also reveal a limitation: the higher-magnitude fault in Case 2 produces visibly larger transient excursions around 1000 s. The benchmark table makes the tracking–actuation trade-off more explicit. For example, the course RMSE of the proposed controller is 0.386° in Case 1 and 3.613° in Case 2, compared with 0.215° and 3.401°, respectively, for the time-triggered controller. Thus, in these simulations, the event-triggered implementation reduces update activity but does not achieve the time-triggered controller's lower RMSE. The results also show favorable course RMSE and computation time relative to the reported method [14] under the same test settings (Table 2).

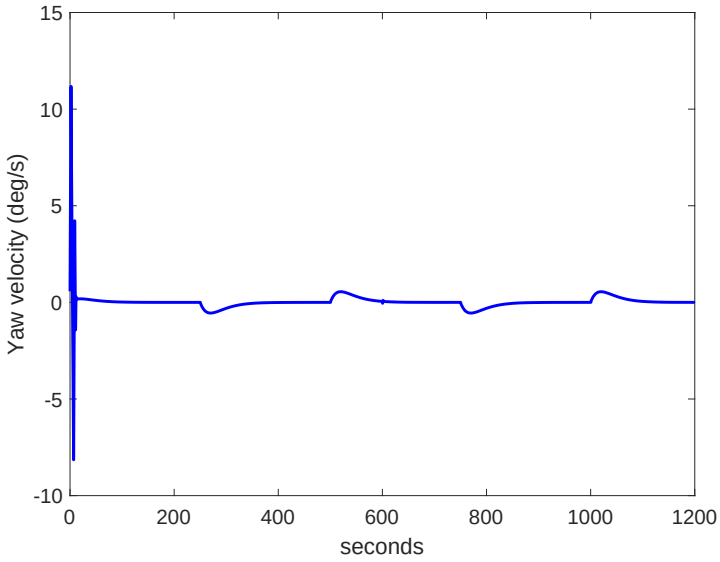


Figure 2. Yaw-rate response of the MSV in Case 1.

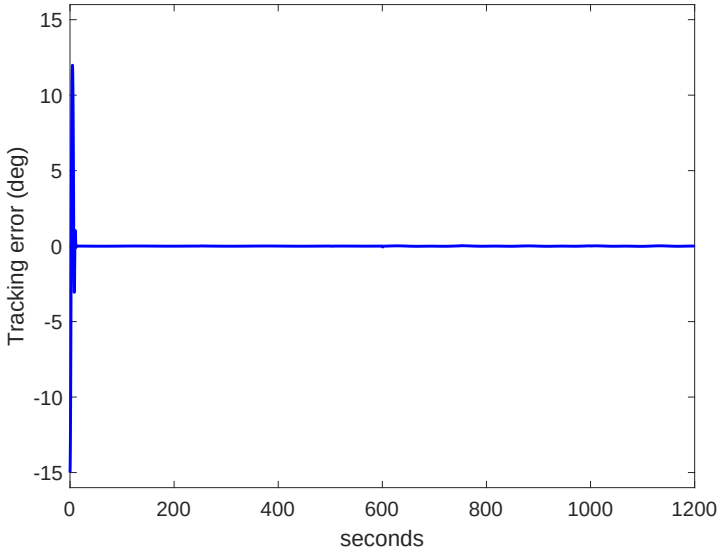


Figure 3. Course-tracking errors in Case 1.

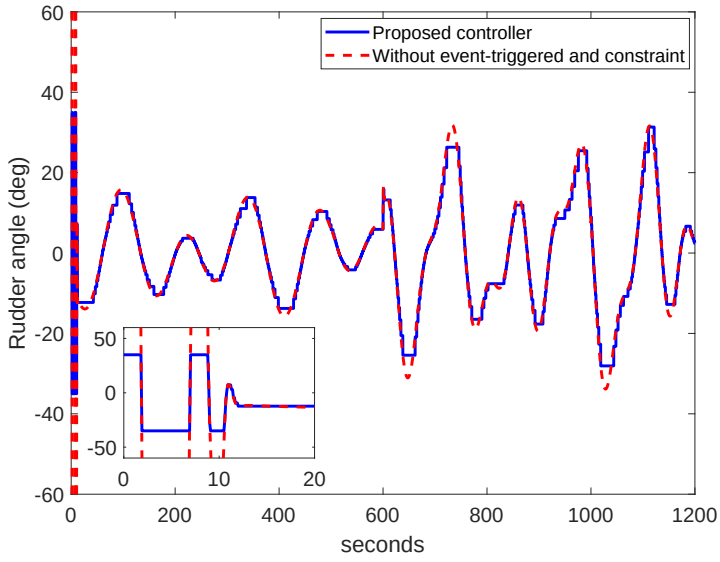


Figure 4. Rudder control input under the prescribed constraint in Case 1.

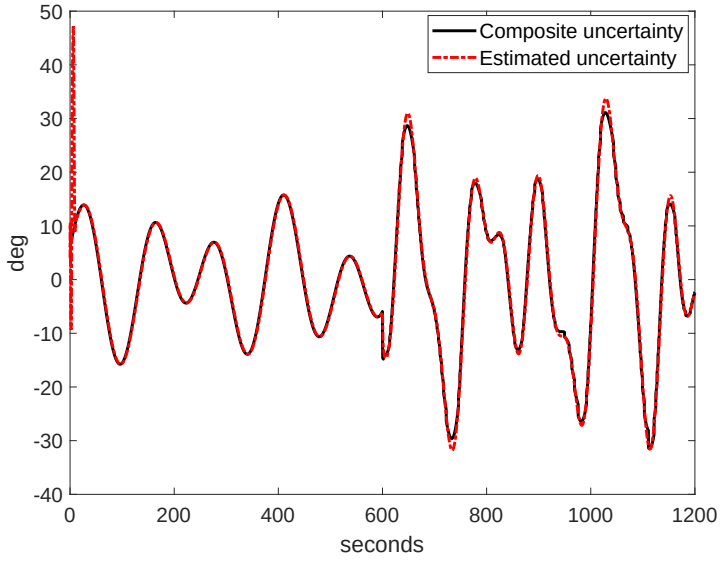


Figure 5. Lumped-uncertainty estimation performance of the observer in Case 1.

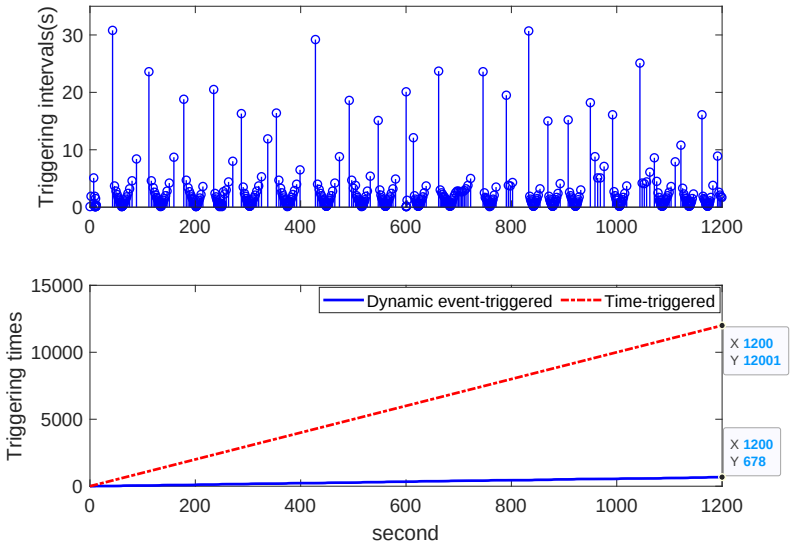


Figure 6. Performance of the dynamic event-triggered mechanism in Case 1.

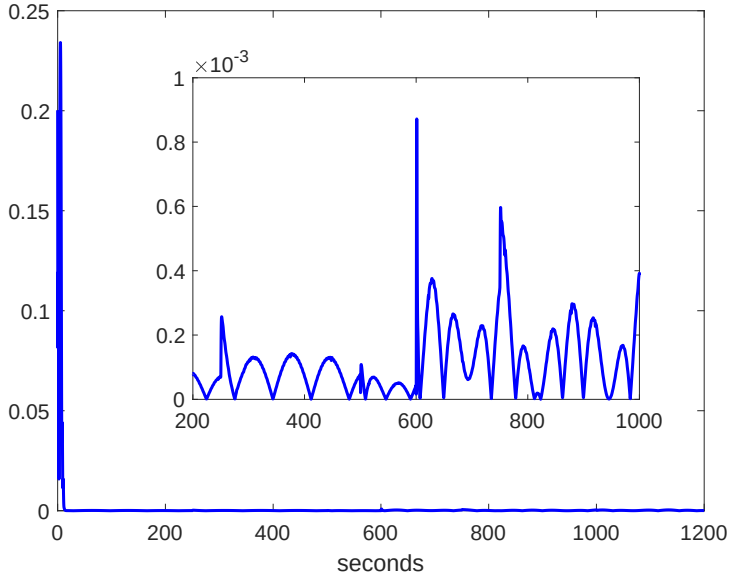


Figure 7. Convergence behavior of the adaptive parameters in Case 1.

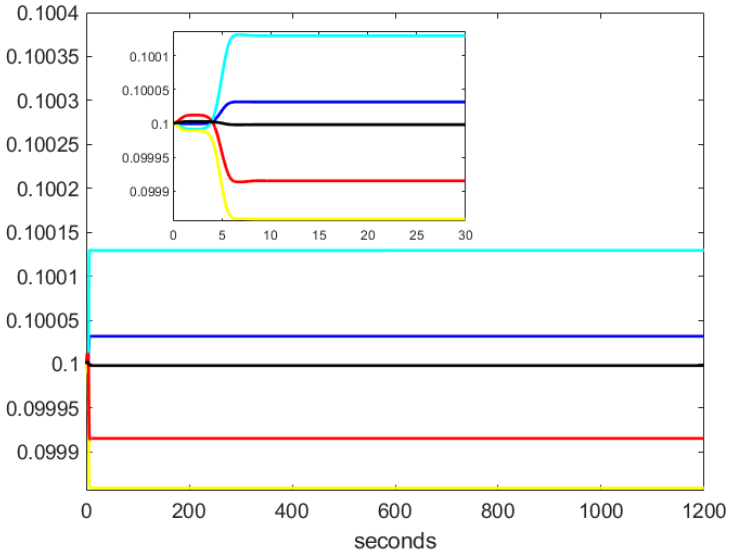


Figure 8. Convergence behavior of the critic-network weights in Case 1.

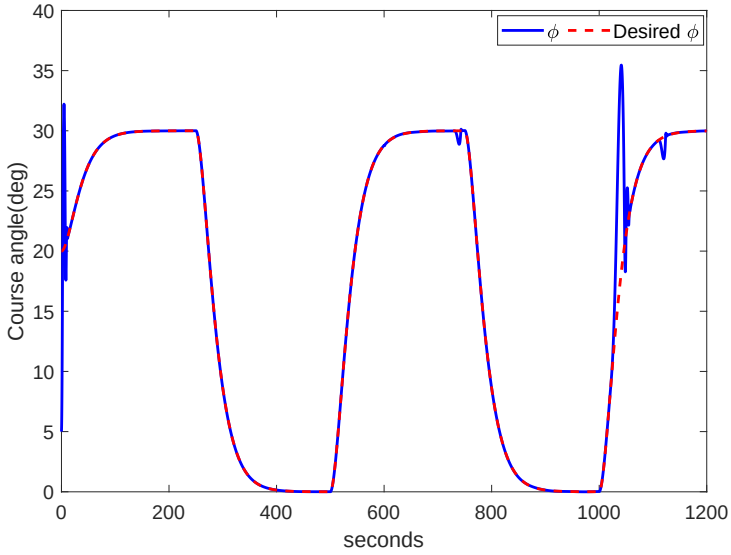


Figure 9. Course-tracking performance of the RL fault-tolerant control scheme in Case 2.

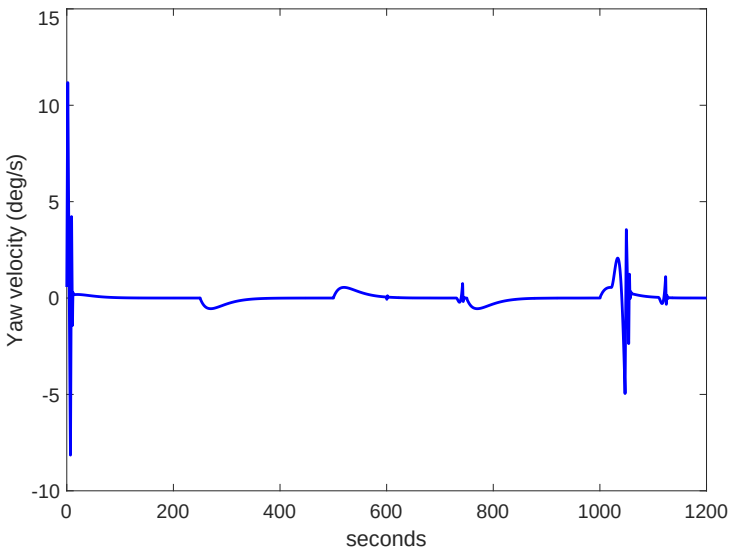


Figure 10. Yaw-rate response of the MSV in Case 2.

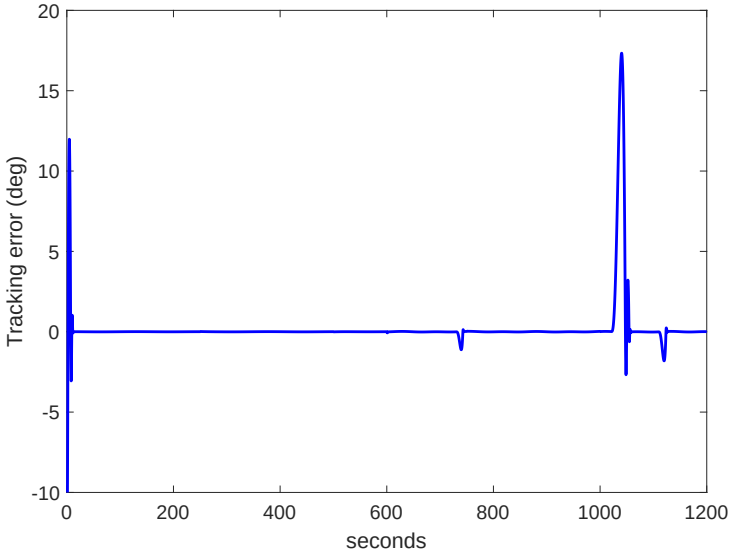


Figure 11. Course-tracking errors in Case 2.

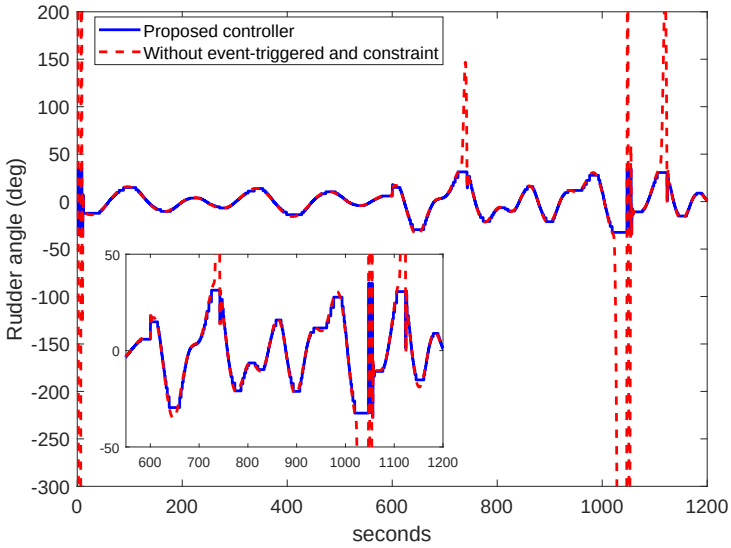


Figure 12. Rudder control input under the prescribed constraint in Case 2.

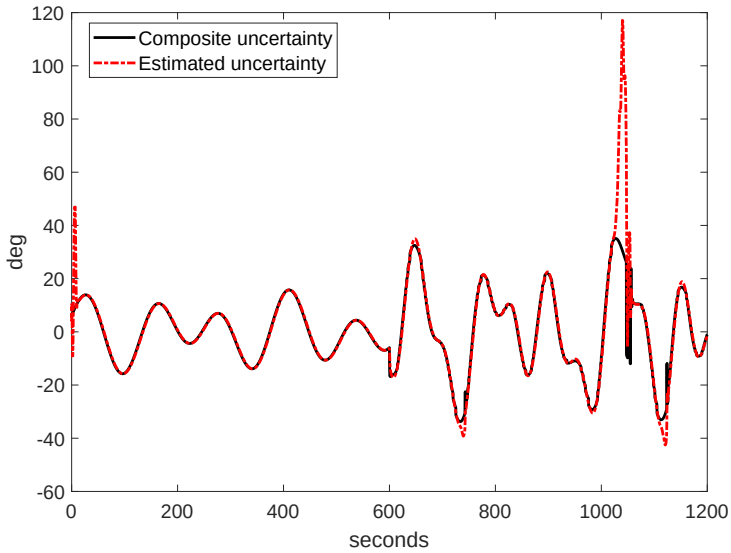


Figure 13. Lumped-uncertainty estimation performance of the observer in Case 2.

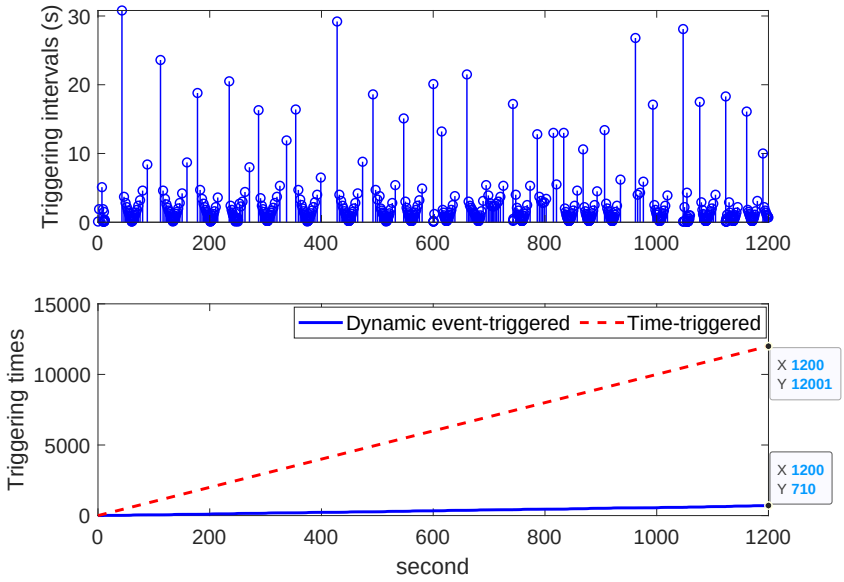


Figure 14. Performance of the dynamic event-triggered mechanism in Case 2.

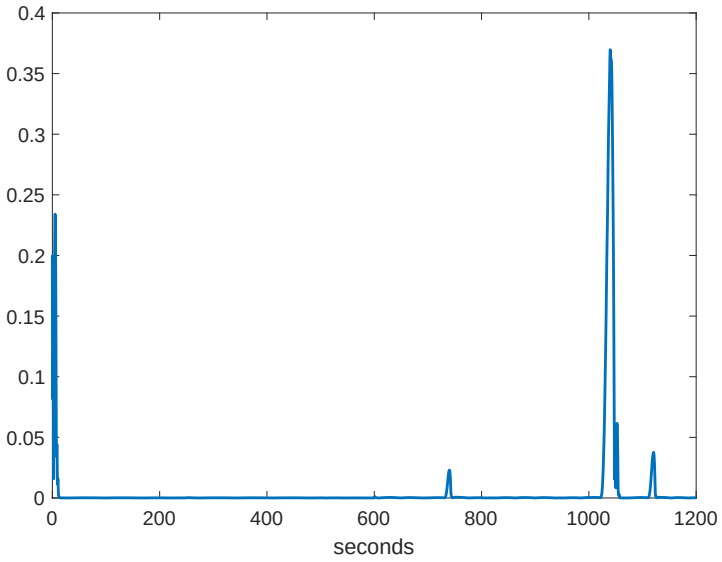


Figure 15. Convergence behavior of the adaptive parameters in Case 2.

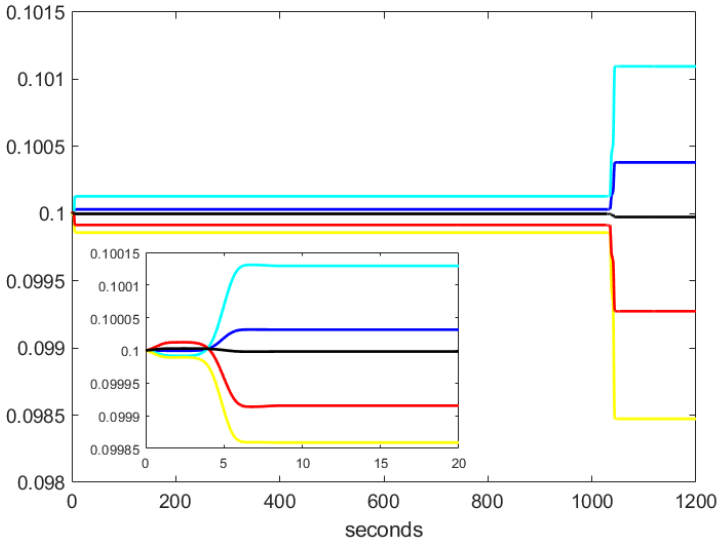


Figure 16. Convergence behavior of the critic-network weights in Case 2.

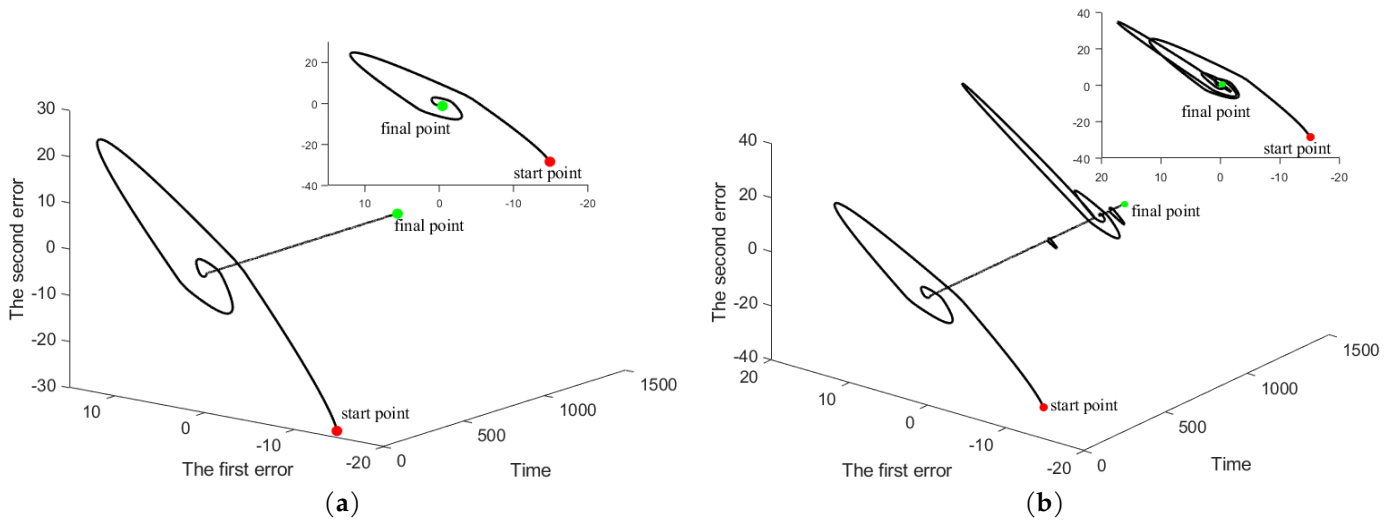


Figure 17. Inward-spiral convergence of error surface in two cases. (**a**) Case 1; (**b**) Case 2.

Practical applicability. The evidence remains numerical and is based on the specified Yulong heading-dynamics model and simulation settings. This manuscript does not claim experimental validation or establish performance across different vessels, operating conditions, or sea environments. The limitations and future-work discussion identify hardware-in-the-loop and vessel experiments, sensitivity analysis, and more detailed actuator, sensor, communication, and computational constraints as necessary next steps.

Table 2. Indicator comparison of different controllers.

Controller	Case 1			Case 2		
	Proposed Controller	Time-Triggered Controller	[14]	Proposed Controller	Time-Triggered Controller	[14]
$RMSE_{\varphi}$ (°)	0.386	0.215	0.505	3.613	3.401	7.324
$RMSE_{\theta}$ (°)	0.690	0.431	Not considered	5.922	4.647	Not considered
Time Consumption (s)	5.540	5.134	10.472	7.746	7.035	15.819
Triggered Times	678	12001	1121	710	12001	1338

Limitations. The present study is limited to theoretical analysis and numerical simulations using a single vessel heading-dynamics model. No hardware-in-the-loop, model-scale, or full-scale experimental validation was performed. The reported results therefore provide evidence of the controller behavior under the specified simulation conditions, rather than establishing its applicability across different vessel types, operating speeds, loading conditions, or sea environments. The parameter configuration was not subjected to a comprehensive sensitivity analysis, and the effects of broader variations in the controller gains, observer gain, learning parameters, network initialization, and triggering settings remain to be investigated. Although a rudder-angle constraint is included in the simulations, the effects of steering-gear rate limits, detailed actuator dynamics, sensor noise, communication delays, and real-time computational limitations are not explicitly validated. Furthermore, the analytical claims depend on the stated approximation, excitation, and fault assumptions. The present results do not establish comparable performance outside these assumptions, including complete actuator failure or fault variations beyond the assumed regularity and available control authority.

5. Conclusions

This paper developed an event-based robust course-tracking scheme for MSVs subject to uncertain nonlinear dynamics, environmental disturbances, and rudder faults. The internal uncertainty, external disturbance, and actuator-fault effects were collected into a lumped term and reconstructed by an observer. A tracking-oriented performance index and a single-critic RL structure were then used to generate the learning-based compensation input without introducing a separate actor network. In addition, a dynamic event-triggered mechanism was placed between the controller and rudder actuator so that the calculated command was transmitted only when the triggering condition was satisfied. Under the stated boundedness, approximation, excitation, and control-authority assumptions, the Lyapunov analysis established uniform ultimate boundedness of the closed-loop signals, and the derived positive lower bound on the inter-event time excluded Zeno behavior.

Future work will address the present limitations through systematic sensitivity analysis and parameter-selection studies; hardware-in-the-loop, model-scale, and full-scale experiments involving different vessel types, speeds, loading conditions, and sea environments; and the inclusion of steering-gear rate limits, detailed actuator dynamics, sensor noise, communication delays, and real-time computational constraints. Further research will also investigate less restrictive approximation and excitation conditions, fault scenarios approaching complete loss of actuator authority using suitable reconfiguration or redundancy, and direct experimental evaluation of energy consumption and mechanical wear.

Author Contributions: Conceptualization, X.G. and D.S.; methodology, X.G.; software, X.G.; validation, X.G., L.K. and Q.S.; formal analysis, X.G. and Q.S.; investigation, X.G.; resources, D.S. and X.H.; data curation, X.G.; writing—original draft preparation, X.G.; writing—review and editing, X.G., L.K., D.S., Q.S. and X.H.; visualization, X.G.; supervision, D.S. and X.H.; project administration, D.S.; funding acquisition, X.G., D.S. and X.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grants 52301418 and 52471376; in part by the Natural Science Foundation of Liaoning Province under Grants 2026-BS-0240; in part by the Liaoning Young Elite Scientists Sponsorship Program; in part by the Fundamental Research Funds for the Central Universities under Grants 3132026366; in part by the Dalian Science and Technology Talent Innovation Support Program (Outstanding Young Scientific and Technological Talents Project) under Grant No. 2025RJ06; and in part by the Liaoning Revitalization Talents Program (Youth Top-notch Talents Project) under Grant No. XLYC2503132.

Data Availability Statement: The original contributions presented in this study are included in this article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Quan Shi was employed by the company CASCO Signal Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Shu, Y.; Zhu, Y.; Xu, F.; Gan, L.; Lee, P.; Yin, J.; Chen, J. Path planning for ships assisted by the icebreaker in ice-covered waters in the northern sea route based on optimal control. *Ocean Eng.* **2023**, *267*, 113182. [CrossRef]
2. Yang, X.; Shi, Y.; Liu, W.; Ye, H.; Zhong, W.; Xiang, Z. Global path planning algorithm based on double DQN for multi-tasks amphibious unmanned surface vehicle. *Ocean Eng.* **2022**, *266*, 112809. [CrossRef]
3. Fu, J.; Sun, G.; Yao, W.; Wu, C.; Wu, L. Dubins path planning of heterogeneous UAV collaborative data collection for IoT network. *IEEE Trans. Intell. Transp. Syst.* **2026**, *27*, 2214–2224. [CrossRef]
4. Yang, Y.; Zhang, J.; Liu, K.; Li, T. Multitarget finite-time dynamic encirclement control for multi-ASV systems. *IEEE Trans. Syst. Man Cybern. Syst.* **2026**, *56*, 565–578. [CrossRef]
5. Liu, S.; Zuo, Y.; Li, T.; Wang, H.; Gao, X.; Xiao, Y. Adaptive fixed-time PID-based control of uncertain nonlinear systems and its application to unmanned surface vehicles. *Int. J. Syst. Sci.* **2024**, *55*, 2815–2824. [CrossRef]

6. Skjetne, R.; Fossen, T.I.; Kokotović, P.V. Adaptive maneuvering, with experiments, for a model ship in a marine control laboratory. *Automatica* **2005**, *41*, 289–298. [CrossRef]
7. Zhang, Q.; Jiang, N.; Hu, Y.; Pan, D. Design of course-keeping controller for a ship based on backstepping and neural networks. *Int. J. e-Navig. Marit. Econ.* **2017**, *7*, 34–41. [CrossRef]
8. Yan, Y.; Liang, H.; Li, T.; Li, H.; Huang, T. Tunable Funnel-Based Distributed Safety-Critical Formation Maneuvering of Air-Ground Swarms. *IEEE Trans. Autom. Control* **2026**, 1–16. . [CrossRef]
9. Gao, X.; Cui, E.; Yang, D.; Tan, Z.; Sun, J. Adaptive Displacement Constraint Control With Predefined Performance for Active Magnetic Bearings. *IEEE Trans. Autom. Sci. Eng.* **2025**, *22*, 883–894. [CrossRef]
10. Zhang, Y.; Xiao, Y.; Teng, F.; Zhang, W.; Li, T. Distributed Prescribed-Time Optimal Dispatch for Ship-Integrated Energy System With Environmental Awareness. *IEEE Trans. Netw. Sci. Eng.* **2026**, *13*, 2181–2196. [CrossRef]
11. Zhang, Y.; Xiao, Y.; Teng, F.; Li, T. Distributed Energy Management for Ship-Integrated Energy System: Secure and Green Sailing. *IEEE Trans. Transp. Electrif.* **2025**, *11*, 8189–8201. [CrossRef]
12. Wu, Z.J.; Hao, L.Y. Multivariate Gaussian process-based learning model predictive control with unscented Kalman filter for autonomous surface vehicles. *ISA Trans.* **2026**, *169*, 408–418. [CrossRef]
13. Ning, J.; Cui, X.; Gao, X.; Li, T.; Tong, S.; Chen, C.L.P. Fuzzy Event-Driven Optimal Quantized Control for USVs Under Thruster and Positioning Constraints: An Adaptive Dynamic Programming Method. *IEEE Trans. Fuzzy Syst.* **2026**, *34*, 2601–2613. [CrossRef]
14. Gao, X.; Hu, X.; Liu, J.; Li, T. Event-Driven Prescribed Optimal Disturbance Rejection for Dynamic Positioning of Ships via Reinforcement Learning. *IEEE Trans. Autom. Sci. Eng.* **2026**, *23*, 4615–4626. [CrossRef]
15. Huang, Z.; Li, T.; Liu, L.; Liang, H. Event-Based Human-in-the-Loop Optimal Bipartite Consensus Control for Multiagent Systems Under False Data Injection Attacks. *IEEE/CAA J. Autom. Sin.* **2026**, *13*, 2119–2131. [CrossRef]
16. Zhang, H.; Cui, L.; Zhang, X.; Luo, Y. Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method. *IEEE Trans. Neural Netw.* **2011**, *22*, 2226–2236. [CrossRef]
17. Wen, G.; Ge, S.S.; Chen, C.L.P.; Tu, F.; Wang, S. Adaptive tracking control of surface vessel using optimized backstepping technique. *IEEE Trans. Cybern.* **2019**, *49*, 3420–3431. [CrossRef]
18. Chen, L.; Dai, S.L.; Dong, C. Adaptive optimal tracking control of an underactuated surface vessel using actor–critic reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 7520–7533. [CrossRef]
19. Hu, X.; Gao, X.; Li, T.; Long, Y. Dynamic event-triggered prescribed performance disturbance rejection fault-tolerant trajectory tracking for surface vehicles. *Ocean Eng.* **2024**, *312*, 119063. [CrossRef]
20. Li, X.; Hu, Y. Time-varying disturbance observer-based adaptive neural network control for uncertain nonlinear systems. *Appl. Soft Comput.* **2026**, *202*, 115815. [CrossRef]
21. Yuan, L.E.; Zuo, Y.; Tong, S.; Li, T. Output-feedback optimal tracking control of multiple ASVs with intermittent actuator faults and unknown disturbances. *Ocean. Eng.* **2026**, *364*, 127033. [CrossRef]
22. Zhu, C.; Zhang, E.; Li, J.; Huang, B.; Su, Y. Approximation-free appointed-time tracking control for marine surface vessel with actuator faults and input saturation. *Ocean Eng.* **2022**, *245*, 110468. [CrossRef]
23. Zhu, G.; Ma, Y.; Li, Z.; Malekian, R.; Sotelo, M.A. Event-triggered adaptive neural fault-tolerant control of underactuated MSVs with input saturation. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 7045–7057. [CrossRef]
24. Shen, Z.; Zhou, L.; Yang, Y.; Liu, Y.; Yu, H. Dynamic event-triggered ADP optimized tracking control of marine surface vessels with asymmetric time-varying full-state constraint. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 7352–7362. [CrossRef]
25. Li, J.; Li, T.; Fan, Z.; Bu, R.; Li, Q.; Hu, J. Direct adaptive NN control of ship course autopilot with input saturation. In Proceedings of the Fourth International Workshop on Advanced Computational Intelligence, Wuhan, China, 19–21 October 2011; pp. 655–661.
26. Tong, S.; Sun, K.; Sui, S. Observer-based adaptive fuzzy decentralized optimal control design for strict-feedback nonlinear large-scale systems. *IEEE Trans. Fuzzy Syst.* **2018**, *26*, 569–584. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.