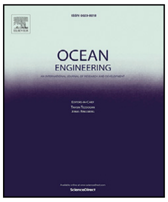
Ocean Engineering 358 (2026) 125757

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng



Research paper

Integrated A*-elliptical trust-region planning for fast USVs formation with control convergence

Zhong Weibo [a],*, Cai Shaohan [a], Gu Yang [b,a], Su Xin [a,b], Zhu Yazhou [a]

[a] Ocean college & Excellent Engineer College, Jiangsu University of Science and Technology, Zhenjiang, 212003, China
[b] Ocean college, Zhejiang University, Zhoushan, 316022, China

ARTICLE INFO

Keywords:
Formation generation
Elliptical Trust Region optimization
Path planning
Sliding mode control

ABSTRACT

This paper proposes a formation framework for USVs to achieve fast and synchronized assembly in complex environments. The method integrates A* path planning with an elliptical trust-region (ETR) optimizer that dynamically adjusts the formation center to minimize the maximum completion time. A sliding-mode controller, structured on remaining-time synchronization and augmented with a disturbance observer, ensures simultaneous arrival with common terminal velocity and heading. Simulations conducted under varying environmental complexity, fleet sizes, and formation geometries demonstrate that the proposed approach attains success rates of 75.3–97.8%, keeps terminal velocity errors below 6% and heading errors under 0.5°, and reduces speed and heading fluctuations by 16%–28%. Compared with conventional and learning-based benchmarks, the framework exhibits higher efficiency, faster convergence, and improved motion smoothness in simulation, providing an effective and computationally efficient approach for fast multi-USV formation generation.

1. Introduction

Unmanned Surface Vehicles (USVs) have emerged as pivotal platforms for a wide range of maritime missions. The operational demands of these complex, large-scale tasks often exceed the capabilities of a single vehicle, driving significant research interest in the coordination of multiple USVs (Zhong et al., 2025). Within this domain, formation control is a fundamental capability for enhancing operational efficiency, coverage, and system robustness.

USV formation control encompasses formation generation, transformation, and navigation (Sang et al., 2023). Recent years have witnessed remarkable progress in formation control, fueled by advances in multi-agent systems, nonlinear control, and artificial intelligence. Formation control for multi-USVs primarily follows three paradigms: behavior-based, virtual structure, and leader-follower (Er et al., 2023). Gu et al. (2024) investigated a leader–follower motion control strategy for cooperative USV–ROV operations in subsea cable inspection missions. The virtual structure and leader-follower methods are often used for navigation control of established formations. Yu et al. (2025) proposed a virtual–real integrated framework for validating USV formation control algorithms. The former models the formation as a rigid body, providing robustness at the expense of flexibility. The latter relies on a leading USV but often lacks sufficient feedback, leading to over-dependence on the leader's performance and potential inefficiencies. In contrast, the behavior-based approach focuses on formation shaping and dynamic adaptation through decentralized rules for individual agents in uncertain environments (Matarić, 1995).

The rapid generation of a USV formation involves two critical phases. First, an optimal formation configuration must be determined. This requires minimizing the total formation time by synthesizing mission requirements, individual USV dynamics, their initial states, and the operational environment. Subsequently, precise formation control is necessary to ensure consistency during convergence. This entails guiding all USVs to arrive at their designated positions simultaneously, with their final velocities and headings precisely aligned to meet the specified formation parameters while avoiding collisions. Liu et al. (2026) developed a distributed H_∞ time-varying formation tracking method for 3-DOF USVs with disturbance observer–based disturbance compensation. Xu et al. (2026) proposed a robust formation control strategy for unmanned surface vehicles to improve formation stability under external disturbances.

USVs formation generation is typically addressed through multi-agent path planning, where bio-inspired optimization algorithms and multi-task assignment strategies are essential (Xing et al., 2023). With advances in artificial intelligence, reinforcement learning (RL) has been widely adopted for USV formation and navigation control

---

* Corresponding author.
E-mail address: vebo@just.edu.cn (W. Zhong).

https://doi.org/10.1016/j.oceaneng.2026.125757
Received 21 January 2026; Received in revised form 19 April 2026; Accepted 23 April 2026

(Mou et al., 2021). To satisfy Markovian assumptions, redesigned state spaces help mitigate issues caused by system inertia and lag (Zhong et al., 2022). Within this field, Wang et al. (2021) proposed a data-driven, performance-prescribed RL control scheme using an actor-critic framework, ensuring both optimality and prescribed tracking accuracy. Wang et al. (2023) developed a speed-adaptive obstacle avoidance method based on deep RL, demonstrating robustness in large-scale uncertain environments. Xiaofei et al. (2022) proposed a DDQN-based global static path planning method to address path planning problems of amphibious unmanned surface vehicles in complex maritime environments. Furthermore, accounting for USV motion characteristics, Jin et al. (2022) introduced a distributed soft collision avoidance strategy and a multi-task training framework for leader-follower formation control.

However, despite their strengths in handling complex interactions, bio-inspired and machine learning approaches exhibit inherent limitations in fast formation generation. A primary concern is their heavy reliance on extensive training data, which is particularly challenging to obtain (Cui et al., 2024). Traditional optimization and model-based control methods remain essential in multi-USV formation generation.

Path planning plays a critical role in the performance of USVs formations. Commonly used algorithms such as A* and its variants have been enhanced through techniques like Octile-distance heuristics, bidirectional search, and path smoothing, which significantly improve search efficiency and path quality. These advances enable the rapid generation of feasible formation paths (Li et al., 2025), collectively establishing a reliable and computationally efficient foundation for formation.

Trust region methods are iterative algorithms that optimize an objective function by repeatedly constructing a local model — typically quadratic — within a bounded trust region (Waltz et al., 2006; Fang et al., 2025). A candidate step is generated by minimizing this model inside the region and is accepted based on how closely the actual function reduction matches the predicted reduction (Yu et al., 2024). Key advantages include strong convergence even for nonconvex problems, adaptability through dynamic adjustment of the region size, and suitability for derivative-free optimization via interpolation or regression from sampled values (Yu et al., 2024; Mohammadi et al., 2025). However, the approach requires solving a constrained quadratic subproblem, which can be computationally expensive in high dimensions. Performance also depends on the initial trust region size, update rules, and the accuracy of the local model, where poor approximations may lead to slow convergence or stagnation.

To achieve rapid USV formation, optimal positions must be determined based on each USV's current location, capabilities, environment, and formation shape. The process typically aims to minimize the total formation time by ensuring that the last-arriving USV completes its task as quickly as possible. A trust-region method with path planning can be applied for this purpose, which iteratively constructs a quadratic model, proposes updated formations by solving a trust-region subproblem, and dynamically adjusts the trust-region radius according to actual performance improvement. As a result, it enables real-time, collaborative, and derivative-free position optimization with theoretical convergence guarantees (Yuan, 2015).

To ensure that all USVs reach their designated positions simultaneously with matching velocity and heading along a planned trajectory, Sliding Mode Control (SMC) is well-suited to this task due to its straightforward design, rapid convergence, and intrinsic robustness against matched uncertainties and disturbances (Zhang et al., 2024). Recent research has focused on integrating SMC with complementary techniques to mitigate its primary practical limitation — control chattering — while enhancing overall performance (Meng et al., 2024). Representative integrations include combining SMC with genetic algorithm-based extended state observers for advanced disturbance estimation, and with prescribed performance control to guarantee convergence within a predefined time frame (Dong et al., 2024).

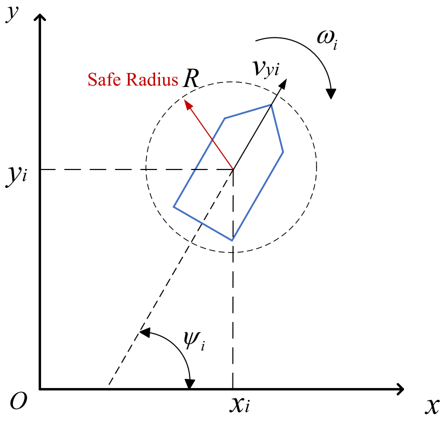


Fig. 1. Kinematic model.

This paper introduces a real-time formation-control framework that couples A* path planning with an Elliptical Trust-Region (ETR) optimizer. The main contributions are summarized as follows:

1. This paper reformulates the USVs formation problem as a minimax arrival-time optimization and proposes an integrated A*–ETR framework to determine the optimal formation center.
2. A residual-time-based sliding mode controller is developed, capable of enforcing arrival-time consistency and maintaining terminal speed and heading alignment, thereby achieving synchronized arrival under dynamic replanning.
3. A dedicated disturbance observer is designed to compensate for disturbances associated with A* replanning and environmental variations. A stability analysis is provided for the observer-assisted navigation dynamics over continuous-time intervals between successive replanning events.
4. Extensive simulations — spanning varying environmental complexities, fleet sizes, and formation geometries — demonstrate that our framework outperforms representative learning-based and model-based benchmarks. It achieves higher success rates, faster convergence, and a decisive edge in motion smoothness.

2. Problem statement

Multiple USVs formation is typically defined as a motion state in which all USVs maintain prescribed relative positions and orientations. Specifically, let the formation center $\mathbf{c}(t) = (x, y)$, and the desired target positions of the ith USV be (x_d^i, y_d^i). Moreover, all USVs should reach their designated target points at the same terminal time T_f, while also sharing a common terminal speed and heading angle:

$$\min(T_f), \quad and \quad v_i(T_f) = \hat{v}, \quad \psi_i(T_f) = \hat{\psi} \tag{1}$$

As illustrated in Fig. 1, the state vector of the ith USV is defined as follows:

$$\boldsymbol{\eta}_i = [x_i,\ y_i,\ v_i,\ \psi_i]^\top \tag{2}$$

where (x_i, y_i) denote the position of the USV in the earth-fixed coordinate frame, v_i represents the surge speed, and ψ_i denotes the heading angle. Correspondingly, the control input is defined as:

$$\mathbf{u}_i = [u_i,\ \omega_i]^\top \tag{3}$$

where u_i denotes the rate of change of the surge speed, and ω_i denotes the rate of change of the heading angle. Under this assumption, the continuous-time kinematic model of the USV can be expressed as:

$$\begin{cases} \dot{x}_i = v_i \cos\psi_i \\ \dot{y}_i = v_i \sin\psi_i \\ \dot{v}_i = u_i \\ \dot{\psi}_i = \omega_i \end{cases} \tag{4}$$

The environment contains a set of static obstacles $\mathcal{O}_s = \{O_1, O_2, \dots, O_m\}$, where each O_j is modeled as a region that no USV is permitted to enter. To represent moving maritime traffic and other time-varying interferers, we further define a set of dynamic obstacles $\mathcal{O}_d(t) = \{D_1(t), D_2(t), \dots, D_n(t)\}$. Each dynamic obstacle $D_k(t)$ is characterized by a time-varying center position $(x_k(t), y_k(t))$, a velocity vector (v_{xk}, v_{yk}), and a safety radius r_k. To guarantee collision avoidance, the following constraint must hold for every USV and for all obstacles:

$$\left\|(x_i(t), y_i(t)) - (x, y)\right\| \geq R + r_k, \quad \forall (x, y) \in O_j \cup D_k(t) \tag{5}$$

where R denotes the safety radius of the USV.

To simply represent the key features of multi-USV path planning and synchronized arrival in complex environments, we adopt the following simplifying assumptions:

Assumption 1. The motion of USVs is typically modeled in the plane, neglecting higher-order dynamics such as roll, pitch, heave, and simplifying the impact of the environment.

Assumption 2. The static and dynamic obstacles are known in advance via sensing or communication.

Assumption 3. USVs are connected through a stable communication network.

3. Proposed algorithm

This study presents a real-time formation control framework that integrates A* path planning with a trust-region optimization strategy. Under the constraints of a predefined formation geometry, the A* algorithm computes a feasible path length from each USV to every candidate target point. These lengths are then incorporated into the trust-region optimizer as time-based cost functions. By iteratively adjusting the global position of the formation, the proposed method minimizes the maximum transit time across all USVs while ensuring obstacle avoidance, thereby achieving a globally optimal assignment of target points. Furthermore, the optimization loop operates in real time, enabling adaptive recalculation of target points in response to dynamic environmental changes or external disturbances.

The controller design incorporates essential physical and safety constraints. Each USV and obstacle is assigned a safety radius R. The sailing safety is handled at the planning layer through distance-threshold monitoring. The primary control objective is global spatiotemporal coordination, requiring all USVs to arrive at their assigned target points simultaneously with a common terminal velocity and heading. To achieve this, a dual-phase control strategy is implemented, as illustrated in Fig. 2.

3.1. Formation position determination

This work focuses on achieving fast formation, where the key performance metric is not the shortest individual path, but rather the total time required for the entire team to reach the desired formation.

When multiple USVs start from different initial positions and navigate through an obstacle-populated environment toward their assigned formation locations, the overall completion time is dictated by the slowest USV. Therefore, the optimization objective is to minimize the team makespan — the maximum transit time among all USVs — by adjusting the global formation center in the plane. This adjustment indirectly influences each USV's target position and path length, thereby shortening the overall completion time.

Assume there are N USVs, each with a desired cruising speed v_i for the ith USV. The desired formation is defined by a given geometric shape, whose adjustable center is denoted by $\mathbf{c} = (x, y)$. In an obstacle-populated environment, a feasible path for the ith USV—from its current position (x_i, y_i) to a target position (x_d^i, y_d^i) determined by the formation shape and center $\mathbf{c}$—can be generated using the heuristic A* search algorithm. Let the corresponding path length be $L_i(c)$. The optimization objective is to minimize the team makespan, defined as the maximum travel time among all USVs:

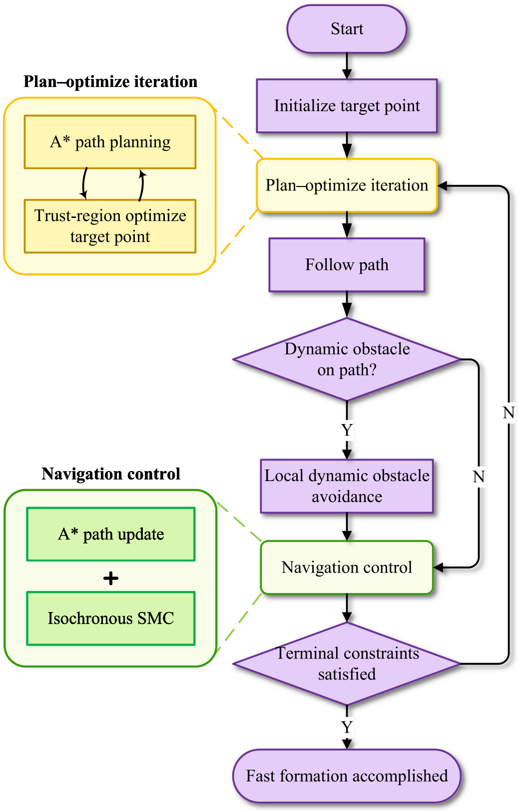


Fig. 2. Overall task flow.

$$\min_{c} \max_{i=1,\dots,N} \frac{L_i(c)}{v_i} \tag{6}$$

The trust-region method is a widely used non-linear optimization algorithm. Its core idea is to construct a quadratic approximation model at the current iterate and to restrict the search within a trust-region, within which the model is considered reliable. The classical trust-region method typically adopts a circular trust region, which can be formulated as:

$$m_k(s) = f(c_k) + g_k^\top s + \frac{1}{2} s^\top B_k s \tag{7}$$

$$\min_{s} \; m_k(s) \quad \text{s.t.} \quad \|s\|_2 \leq \Delta \tag{8}$$

where $s \in \mathbb{R}^2$ denotes the search step, c_k is the formation center at the kth iteration, $g_k = \nabla f(c_k)$ is the gradient, B_k is a quasi-Newton approximation of the local curvature rather than the exact Hessian, and Δ denotes the trust-region radius.

The curvature approximation B_k is updated using the BFGS scheme:

$$B_{k+1} = B_k - \frac{B_k \mathbf{s}_k \mathbf{s}_k^T B_k}{\mathbf{s}_k^T B_k \mathbf{s}_k} + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} \tag{9}$$

where

$$\mathbf{s}_k = \mathbf{c}_{k+1} - \mathbf{c}_k, \quad \mathbf{y}_k = \nabla f(\mathbf{c}_{k+1}) - \nabla f(\mathbf{c}_k)$$
$$\nabla f(\mathbf{c}_k) \approx \begin{bmatrix} \dfrac{f(x_k + h, y_k) - f(x_k, y_k)}{h} \\ \dfrac{f(x_k, y_k + h) - f(x_k, y_k)}{h} \end{bmatrix} \tag{10}$$

h is a small perturbation parameter, and $B_0 = \gamma I, \gamma > 0$.

This method, while straightforward and numerically stable, imposes an isotropic constraint that treats all search directions equally. Consequently, certain directions may be underexplored while others are oversearched, potentially lowering the overall convergence efficiency. To address this limitation inherent in the classical trust-region framework, the present work employs an Elliptical Trust Region (ETR) strategy. The core idea is to introduce a scaling matrix D that assigns distinct stretching or shrinking factors along different coordinate axes, thereby shaping the trust region into an ellipse:

$$\Omega_k = \left\{ \boldsymbol{s} \in \mathbb{R}^2 \,\middle|\, \boldsymbol{s}^\top D_k^{-1} \boldsymbol{s} \le \Delta_k^2 \right\} \tag{11}$$

where $\Delta_k > 0$ denotes the trust-region radius. Accordingly, the subproblem can be formulated as:

$$\min_{\boldsymbol{s}} \; m_k(\boldsymbol{s}) \quad \text{s.t.} \quad \boldsymbol{s}^\top D^{-1} \boldsymbol{s} \le \Delta_k^2 \tag{12}$$

The scaling matrix D_k is constructed using a diagonal curvature-based strategy with regularization:

$$D_k = \operatorname{diag}\left(\frac{1}{\sqrt{|B_{k,11}| + \varepsilon}}, \frac{1}{\sqrt{|B_{k,22}| + \varepsilon}} \right) \tag{13}$$

where $\varepsilon > 0$ is a small regularization parameter. The diagonal approximation is chosen for two reasons: first, computational simplicity without full eigenvalue decomposition; second, and more importantly, the minimax arrival time from A* costs is non-smooth and highly sensitive to obstacles. Off-diagonal entries from a finite-difference BFGS approximation may introduce unstable directional rotations. Hence, D_k is not intended to recover the full Hessian geometry, but rather serves as a conservative anisotropic scaling that prioritizes numerical stability over accurate second-order information.

To evaluate the predictive accuracy of the quadratic model $m_k(\boldsymbol{s})$ with respect to the true objective function, the ratio between the actual reduction and the predicted reduction (reduction ratio) is introduced as:

$$r_k = \frac{\mathrm{Ared}_k}{\mathrm{Pred}_k} = \frac{f(\boldsymbol{c}_k) - f(\tilde{\boldsymbol{c}}_{k+1})}{m_k(\boldsymbol{0}) - m_k(\boldsymbol{s}_k)} \tag{14}$$

This ratio reflects the reliability of the local quadratic model within the current trust region.

Based on the reduction ratio r_k, the acceptance criterion for the trial step is defined as:

$$\boldsymbol{c}_{k+1} = \begin{cases} \boldsymbol{c}_k + \boldsymbol{s}_k, & r_k > 0 \\ \boldsymbol{c}_k, & \text{otherwise} \end{cases} \tag{15}$$

Meanwhile, the trust-region radius Δ_k is adaptively updated according to the following rule:

$$\Delta_{k+1} = \begin{cases} a_2 \Delta_k, & r_k > b_2 \\ \Delta_k, & b_2 \ge r_k \ge b_1 \\ a_1 \Delta_k, & r_k < b_1 \end{cases} \tag{16}$$

where the parameters satisfy

$$0 < b_1 < b_2 < 1, \qquad 0 < a_1 < 1 < a_2 \tag{17}$$

This update mechanism ensures consistent progress toward convergence while allowing the elliptical trust region to dynamically adapt the search scope in response to local model fidelity. Specifically, the scaling matrix D_k shapes the direction and geometry of the trust region, whereas the radius Δ_k governs the overall step size. In contrast to a circular trust region, the elliptical formulation facilitates directional scaling, permitting larger steps along well-conditioned directions and finer adjustments along more sensitive ones, thereby enhancing the overall efficiency of the optimization search.

The A* algorithm and the Elliptical Trust Region method are integrated into a tightly coupled, closed-loop iterative framework. Given a current formation center $\boldsymbol{c}_k = (x_k, y_k)$, the A* algorithm computes an optimal obstacle-free path for each USV to its assigned target position, from which the path length $L_i(\boldsymbol{c}_k)$ and travel time $T_i(\boldsymbol{c}_k)$ are obtained. These times are aggregated into the objective function $f(\boldsymbol{c}_k) = maxT_i(\boldsymbol{c}_k)$, representing the maximum completion time (i.e., the makespan) under $\boldsymbol{c}_k$.

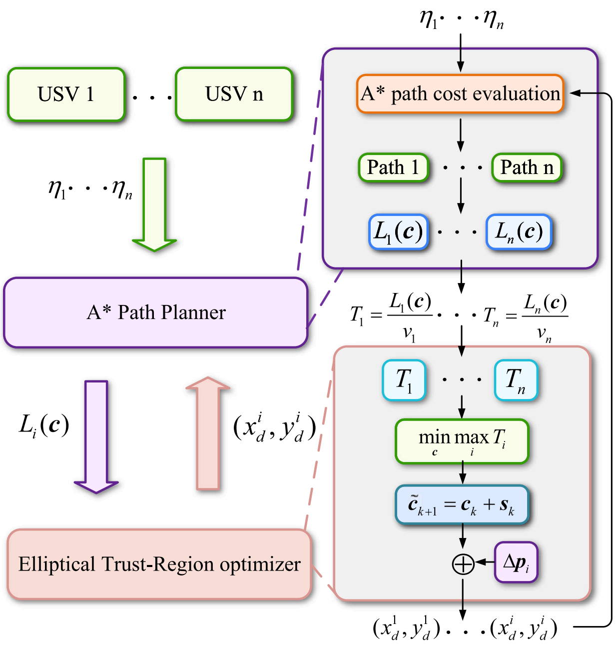


Fig. 3. A* planner with elliptical trust-region.

Subsequently, the ETR method constructs a local quadratic model of f around $\boldsymbol{c}_k$ and solves a constrained subproblem to update the center to $\boldsymbol{c}_{k+1} = (x_{k+1}, y_{k+1})$. Thus, A* provides the time-cost evaluation for a given center, while ETR adjusts the center to reduce the makespan.

The two modules operate in a mutually reinforcing loop: A* supplies feasible-path time feedback, and ETR uses this information to iteratively refine the formation center, progressively narrowing the feasible region. Throughout the iterations, the identity of the bottleneck USV(s) may change, yet the optimization objective remains consistently the minimization of the maximum travel time.

The loop converges to an optimal formation center $\boldsymbol{c}^* = (x^*, y^*)$, from which the corresponding target positions $(x_d^{(i)}, y_d^{(i)})$ for all USVs are determined, as illustrated in Fig. 3. The iterative process is outlined in Algorithm 1. To handle infeasible trial centers generated by ETR, a hierarchical backup mechanism is incorporated into the optimization loop, comprising trust-region shrinking(B1), a local feasible-center search(B2), and fallback to the previous feasible center(B3).

3.2. Formation navigation control

Once the optimal formation position is determined, all USVs must navigate to their assigned positions and arrive simultaneously. This paper adopts a sliding-mode control strategy (SMC) for formation navigation. For the ith unmanned surface vehicle (USV), let $L_{\mathrm{rem},i}$ denote the remaining path length along the A*-planned path and $v_i(t)$ denote the current surge speed. The remaining arrival time is defined as:

$$\tau_i(t) = \frac{L_{\mathrm{rem},i}(t)}{v_i(t)} \tag{18}$$

$$\bar{\tau}(t) = \max_i \tau_i(t) \tag{19}$$

To promote active acceleration of the USV that dominates the overall completion time, different priority-adjusted remaining-time references are constructed from the shared $\bar{\tau}(t)$. Let

$$i^\star = \arg\max_i \tau_i(t) \tag{20}$$

Algorithm 1: Pseudocode of A*–ETR with backup mechanism

```
Input: Initial states $\{\eta_i^0\}$, obstacle map $\mathcal{O}$, formation offsets $\{\Delta \boldsymbol{p}_i\}$,
       initial center $\boldsymbol{c}_0$, trust-region radius $\Delta_0$, parameters
       $(a_1, a_2, b_1, b_2)$, $K_{\max}$, $M$
Output: Optimized center $\boldsymbol{c}^\star$, target points $\{\boldsymbol{p}_i^\star\}$, A* paths $\{\mathcal{P}_i^\star\}$
Initialize $\boldsymbol{c} \leftarrow \boldsymbol{c}_0$, $\Delta \leftarrow \Delta_0$, $k \leftarrow 0$;
while $k < K_{\max}$ do
    Compute $\boldsymbol{p}_i(\boldsymbol{c})$, run A*, obtain $L_i(\boldsymbol{c})$;
    Evaluate $f(\boldsymbol{c})$;
    Solve ETR to obtain $\boldsymbol{s}_k$, set $\tilde{\boldsymbol{c}} = \boldsymbol{c} + \boldsymbol{s}_k$;
    if A* feasible for all $\boldsymbol{p}_i(\tilde{\boldsymbol{c}})$ then
        Evaluate $f(\tilde{\boldsymbol{c}})$, compute $r_k$;
    end if
    else
        for $m = 1$ to $M$ do
            $\Delta \leftarrow a_1 \Delta$, re-solve ETR, update $\tilde{\boldsymbol{c}}$;
            if A* feasible then
                break
            end if
        end for
        if A* infeasible then
            Search $\boldsymbol{c}'$ within $\|\boldsymbol{c}' - \boldsymbol{c}\| \le \Delta$;
            if found then
                $\tilde{\boldsymbol{c}} \leftarrow \boldsymbol{c}'$
            end if
        end if
        if A* still infeasible then
            $\tilde{\boldsymbol{c}} \leftarrow \boldsymbol{c}$, $\Delta \leftarrow 0.5\Delta$;
        end if
        Evaluate $f(\tilde{\boldsymbol{c}})$, compute $r_k$;
    end if
    if $r_k > 0$ then
        $\boldsymbol{c} \leftarrow \tilde{\boldsymbol{c}}$;
    end if
    if $r_k > b_2$ then
        $\Delta \leftarrow a_2 \Delta$;
    end if
    else if $r_k < b_1$ then
        $\Delta \leftarrow a_1 \Delta$;
    end if
    $k \leftarrow k + 1$;
end while
Set $\boldsymbol{c}^\star \leftarrow \boldsymbol{c}$, $\boldsymbol{p}_i^\star \leftarrow \boldsymbol{c}^\star + \Delta \boldsymbol{p}_i$, $\mathcal{P}_i^\star \leftarrow \mathcal{P}_i(\boldsymbol{c}^\star)$;
```

denote the index of the USV with the largest remaining arrival time. The reference remaining time for the ith USV is defined as:

$$\tau_i^{\mathrm{ref}}(t) = \begin{cases} \bar{\tau}(t) - \Delta_\tau, & i = i^\star \\ \bar{\tau}(t) - \delta_\tau, & i \ne i^\star \end{cases} \qquad \Delta_\tau > \delta_\tau > 0 \tag{21}$$

In the proposed cooperative control framework, δ_τ serves as an allowable synchronization margin to prevent other USVs from arriving too early. In contrast, Δ_τ defines the chasing intensity for the USV with the largest remaining arrival time by assigning it a smaller reference remaining time. The design condition $\Delta_\tau > \delta_\tau$ ensures that this slowest USV is granted higher acceleration priority, thereby preventing the overall formation efficiency from being compromised by a single lagging USV. Based on the reference remaining time introduced above, the remaining-time error and the heading error for the ith USV are respectively defined as:

$$e_{\tau,i}(t) = \tau_i(t) - \tau_i^{\mathrm{ref}}(t), \qquad e_{\psi,i}(t) = \psi_i(t) - \psi_i^{\mathrm{ref}}(t) \tag{22}$$

where $\psi_i^{\mathrm{ref}}(t)$ denotes the reference heading angle determined by the local tangential direction of the A* path.

The error vector is then constructed as:

$$\boldsymbol{e}_i(t) = \begin{bmatrix} e_{\tau,i}(t) \\ e_{\psi,i}(t) \end{bmatrix} \tag{23}$$

To ensure fast convergence and suppress steady-state bias, a vector-form PI-type sliding surface is defined as:

$$\boldsymbol{s}_i(t) = \boldsymbol{e}_i(t) + \Lambda \int_0^t \boldsymbol{e}_i(\zeta)\,\mathrm{d}\zeta, \qquad \Lambda = \mathrm{diag}(\lambda_\tau, \lambda_\psi), \;\; \lambda_\tau, \lambda_\psi > 0 \tag{24}$$

$$\dot{\boldsymbol{s}}_i = \dot{\boldsymbol{e}}_i + \Lambda \boldsymbol{e}_i = \boldsymbol{\phi}_i(\cdot) + \boldsymbol{\Psi}_i(\cdot)\mathbf{u}_i + \mathbf{d}_i(t) \tag{25}$$

where $\mathbf{u}_i = [u_i, \;\omega_i]^\top$ is the control input vector, and $\mathbf{d}_i(t)$ collects unmodeled dynamics and external disturbances.

For the USV model considered in this work, a diagonal input-gain matrix is adopted:

$$\boldsymbol{\Psi}_i(\cdot) = \begin{bmatrix} \psi_{\tau,i}(\cdot) & 0 \\ 0 & 1 \end{bmatrix} \tag{26}$$

where $\psi_{\tau,i}(\cdot)$ is the input gain associated with the remaining-time dynamics, while the heading dynamics satisfy $\dot{\psi}_i = \omega_i$ with a unit input gain.

To guarantee finite-time reachability of the sliding manifold and suppress control chattering, a component-wise vector finite-time reaching law is employed:

$$\dot{\boldsymbol{s}}_i = -K_1 |\boldsymbol{s}_i|^{\boldsymbol{\alpha}} \odot \mathrm{sgn}(\boldsymbol{s}_i) - K_2 \boldsymbol{s}_i \tag{27}$$

$$K_1 = \mathrm{diag}(k_{\tau,1}, k_{\psi,1}) \succ 0, K_2 = \mathrm{diag}(k_{\tau,2}, k_{\psi,2}) \succ 0 \tag{28}$$

$$\boldsymbol{\alpha} = [\alpha_\tau, \alpha_\psi]^\top, \quad 0 < \alpha_\tau, \alpha_\psi < 1 \tag{29}$$

where $\odot$ denotes element-wise multiplication.

Substituting the reaching law Eq. (27) into Eq. (25), the vector sliding-mode control input is obtained as:

$$\mathbf{u}_i^{\mathrm{smc}} = \boldsymbol{\Psi}_i^{-1}(\cdot)\Big(-K_1 |\boldsymbol{s}_i|^{\boldsymbol{\alpha}} \odot \mathrm{sgn}(\boldsymbol{s}_i) - K_2 \boldsymbol{s}_i - \boldsymbol{\phi}_i(\cdot)\Big) \tag{30}$$

The surge acceleration term u_i in $\mathbf{u}_i^{\mathrm{smc}}$ regulates the USV's speed to synchronize the remaining arrival times across the fleet, whereas the yaw-rate term ω_i compensates for the heading error and ensures stable tracking of the A*-planned path. Thus, the overall control input for each USV can be concisely expressed as:

$$\mathbf{u}_i = \begin{bmatrix} u_i \\ \omega_i \end{bmatrix} = \mathbf{u}_i^{\mathrm{smc}} \tag{31}$$

When the A* algorithm performs online path replanning, the remaining path length L_{rem} can change discontinuously, causing abrupt jumps in the estimated time of arrival $\tau = L_{\mathrm{rem}}/v$. Beyond the resulting temporal deviation, path replanning may also introduce sudden changes in the reference heading derived from the updated path.

To suppress these adverse effects, disturbance observers are designed to estimate the equivalent disturbances arising from A* replanning and to provide corresponding compensation in the control law. The time-disturbance term accounts for variations in the remaining path length and speed fluctuations, while the heading-disturbance term captures the impact of sudden changes in the reference heading. Differentiating τ with respect to time:

$$\dot{\tau}_i = g_{\tau,i}(t)\,u_i + d_{\tau,i}(t), \tag{32}$$

where

$$g_{\tau,i}(t) = -\frac{L_{\mathrm{rem},i}}{v_i^2}, \qquad d_{\tau,i}(t) = \frac{\dot{L}_{\mathrm{rem},i}}{v_i} \tag{33}$$

The term $d_{\tau,i}(t)$ represents the equivalent disturbance in the remaining-time channel due to path updates and other unmodeled effects. After a replanning instant, the reference path may switch abruptly, causing a finite jump in $L_{\mathrm{rem},i}$ and consequently in τ_i. Hence, the effect of replanning is modeled here as piecewise-smooth inter-event evolution combined with finite jumps at discrete replanning instants, rather than as a globally smooth continuous-time disturbance.

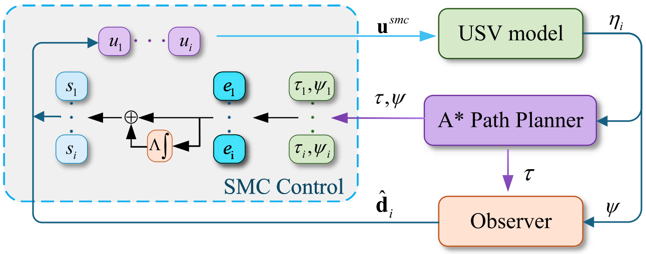


Fig. 4. A* planner with SMC.

To estimate the above disturbance online, an extended state observer (ESO) is constructed for the remaining-time channel as:

$$\begin{cases} \dot{\hat{\tau}}_i = \hat{z}_{\tau 2,i} + \ell_{\tau 1}(\tau_i - \hat{\tau}_i) \\ \dot{\hat{z}}_{\tau 2,i} = g_{\tau,i}(t)\,u_i + \hat{d}_{\tau,i} + \ell_{\tau 2}(\tau_i - \hat{\tau}_i) \\ \dot{\hat{d}}_{\tau,i} = \ell_{\tau 3}(\tau_i - \hat{\tau}_i) \end{cases} \tag{34}$$

where $\hat{\tau}_i$ denotes the estimate of the remaining arrival time, $\hat{z}_{\tau 2,i}$ denotes the estimate of its rate of change, and $\hat{d}_{\tau,i}$ is the estimate of the equivalent disturbance $d_{\tau,i}$. The observer gains satisfy $\ell_{\tau 1}, \ell_{\tau 2}, \ell_{\tau 3} > 0$.

To address the heading disturbance caused by changes in $\psi_{\mathrm{ref},i}$, an extended state observer is designed as follows:

$$\begin{cases} \dot{\hat{e}}_{\psi,i} = \hat{z}_{\psi 2,i} + \ell_{\psi 1}(e_{\psi,i} - \hat{e}_{\psi,i}) \\ \dot{\hat{z}}_{\psi 2,i} = \omega_i + \hat{d}_{\psi,i} + \ell_{\psi 2}(e_{\psi,i} - \hat{e}_{\psi,i}) \\ \dot{\hat{d}}_{\psi,i} = \ell_{\psi 3}(e_{\psi,i} - \hat{e}_{\psi,i}) \end{cases} \tag{35}$$

where $\hat{e}_{\psi,i}$ and $\hat{z}_{\psi 2,i}$ denote the estimates of the heading error and its rate of change, respectively, and $\hat{d}_{\psi,i}$ is the estimate of the equivalent disturbance in the heading channel. The observer gains satisfy $\ell_{\psi 1}, \ell_{\psi 2}, \ell_{\psi 3} > 0$.

By integrating the estimated disturbances as compensation into the Eq. (30), the control input can be expressed as:

$$\mathbf{u}_i = \boldsymbol{\Psi}_i^{-1}(\cdot)\Big(K_1|s_i|^{\alpha} \odot \mathrm{sgn}(s_i) K_2 s_i - \boldsymbol{\phi}_i(\cdot) - \Lambda\,\hat{\mathbf{d}}_i\Big) \tag{36}$$

where

$$\hat{\mathbf{d}}_i = \begin{bmatrix} \hat{d}_{\tau,i} \\ \hat{d}_{\psi,i} \end{bmatrix}, \qquad \Lambda = \mathrm{diag}(\lambda_\tau, \lambda_\psi) \succ \mathbf{0} \tag{37}$$

The control framework is illustrated in Fig. 4 and summarized in Algorithm 2. A*-based replanning triggers upon detecting a dynamic obstacle or USV–USV conflict. The sliding-mode controller then strictly follows the planned path for synchronization and tracking. Consequently, spatial safety is handled at the planning layer and not a hard controller guarantee; rather, a conservative threshold $d_{\mathrm{trig}} > 2R$ maintains a buffer short of the geometric limit.

3.3. Stability of the navigation controller

To analyze the stability and convergence of the proposed navigation controller, a Lyapunov-based analysis is conducted. For the ith USV, the sliding surface vector $\mathbf{s}_i(t)$ is defined as in Eq. (24). Consider the Lyapunov candidate function

$$V_i(t) = \frac{1}{2}\mathbf{s}_i^{\top}(t)\mathbf{s}_i(t) \tag{38}$$

which is positive definite with respect to $\mathbf{s}_i(t)$ and satisfies $V_i(t) = 0$ if and only if $\mathbf{s}_i(t) = \mathbf{0}$.

Taking the time derivative of $V_i(t)$ along the system trajectories yields

$$\dot{V}_i(t) = \mathbf{s}_i^{\top}(t)\dot{\mathbf{s}}_i(t) \tag{39}$$

With Eq. (27), the above equation can be written as:

$$\dot{V}_i = -\mathbf{s}_i^{\top}(t)K_1|\mathbf{s}_i(t)|^{\alpha} \odot \mathrm{sgn}(\mathbf{s}_i(t)) - \mathbf{s}_i^{\top}(t)K_2\mathbf{s}_i(t) \tag{40}$$

Algorithm 2: Pseudocode of Formation Navigation Control with Disturbance Observer

```
Input: Initial remaining path lengths $\{L^0_{\mathrm{rem},i}\}$, initial speeds $\{v^0_i\}$,
       sampling period $\Delta t$, small constant $\varepsilon > 0$, controller gains
       $(K_1, K_2, \Lambda)$, observer gains, maximum time $T_{\max}$
Output: Control inputs $\{\mathbf{u}_i\}$
Initialize $t \leftarrow 0$;
Initialize $\tau_i \leftarrow L^0_{\mathrm{rem},i}/(v^0_i + \varepsilon)$ for all $i$, compute the shared baseline $\bar{\tau}$,
  and construct the priority-adjusted reference $\tau^{\mathrm{ref}}_i$ according to
  Eq. (19)– (21);
Initialize tracking errors $\boldsymbol{e}_i$, sliding surface $\boldsymbol{s}_i$, and observer states for
  all USVs;
while $t \le T_{\max}$ do
    Update remaining path length $L_{\mathrm{rem},i}$ from A* / path tracker;
    Compute $\tau_i \leftarrow L_{\mathrm{rem},i}/(v_i + \varepsilon)$, update the shared baseline $\bar{\tau}$, and
      construct $\tau^{\mathrm{ref}}_i$ according to Eq. (19)– (21);
    Update time error $e_{\tau,i} \leftarrow \tau_i - \tau^{\mathrm{ref}}_i$ and sliding surface $s_i$ by Eq. (24);
    Update disturbance estimate $\hat{\mathbf{d}}_i$ using the observer by
      Eq. (34)– (35);
    Compute control input $\mathbf{u}_i = [u_i, \omega_i]^{\top}$ using SMC and compensation
      by Eq. (36)– (37);
    Apply $\mathbf{u}_i$ and propagate USV kinematics to obtain $\boldsymbol{\eta}_i(t + \Delta t)$;
    $t \leftarrow t + \Delta t$;
end while
```

Because

$$s|s|^{\alpha}\mathrm{sgn}(s) = |s|^{1+\alpha} \tag{41}$$

We know that

$$\mathbf{s}_i^{\top}(t)K_1|\mathbf{s}_i(t)|^{\alpha} \odot \mathrm{sgn}(\mathbf{s}_i(t)) = \sum_j k_{1j}|s_{ij}|^{1+\alpha_j} \tag{42}$$

and

$$\mathbf{s}_i^{\top}(t)K_2\mathbf{s}_i(t) \ge \lambda_{\min}(K_2)\|\mathbf{s}_i(t)\|^2 \tag{43}$$

Hence,

$$\dot{V}_i \le -\sum_j k_{1j}|s_{ij}|^{1+\alpha_j} - \lambda_{\min}(K_2)\|\mathbf{s}_i(t)\|^2, \qquad \forall \mathbf{s}_i(t) \ne \mathbf{0} \tag{44}$$

Let $\alpha_{\min} = \min_j \alpha_j$. Since $0 < \alpha_j < 1$, $1 + \alpha_j \in (1, 2)$. By norm equivalence in finite-dimensional spaces, there exists a positive constant $c_a > 0$ make

$$\sum_j k_{1j}|s_{ij}|^{1+\alpha_j} \ge c_a\|\mathbf{s}_i(t)\|^{1+\alpha_{\min}} \tag{45}$$

Moreover, since

$$V_i = \frac{1}{2}\|\mathbf{s}_i(t)\|^2 \tag{46}$$

there exist positive constants c_1 and c_2 such that

$$\dot{V}_i \le -c_1 V_i^{\frac{1+\alpha_{\min}}{2}} - c_2 V_i \tag{47}$$

Noting that

$$0 < \frac{1 + \alpha_{\min}}{2} < 1 \tag{48}$$

The above inequality satisfies the standard sufficient condition for finite-time stability. Therefore, the Lyapunov function $V_i(t)$ converges to zero in finite time, and the sliding vector $\mathbf{s}_i(t)$ reaches the origin within a finite settling time.

Once the sliding motion is established, i.e.,

$$\mathbf{s}_i(t) = \mathbf{0} \tag{49}$$

the following relation holds:

$$\mathbf{e}_i(t) + \Lambda\int_0^t \mathbf{e}_i(\zeta)\,d\zeta = \mathbf{0} \tag{50}$$

Differentiating the above equation yields the reduced-order error dynamics

$$\dot{\mathbf{e}}_i(t) + \Lambda \mathbf{e}_i(t) = \mathbf{0} \tag{51}$$

Since $\Lambda = \mathrm{diag}(\lambda_\tau, \lambda_\psi) \succ 0$, the reduced-order system is exponentially stable, and the tracking errors satisfy

$$\mathbf{e}_i(t) \to \mathbf{0}, \qquad t \to \infty. \tag{52}$$

Therefore, the proposed controller guarantees finite-time reachability of the sliding surface and asymptotic convergence of the tracking errors, providing explicit stability guarantees for the USV formation navigation system.

3.4. Stability of the disturbance observer

For the remaining-time channel, according to the disturbance-affected dynamics in Eqs. (32)–(33) and the observer structure in Eq. (34), the estimation errors are defined as:

$$\tilde{\tau}_i = \tau_i - \hat{\tau}_i, \quad \tilde{z}_{\tau 2,i} = \dot{\tau}_i - \hat{z}_{\tau 2,i}, \quad \tilde{d}_{\tau,i} = d_{\tau,i} - \hat{d}_{\tau,i} \tag{53}$$

The observer error dynamics can be written as:

$$\dot{\tilde{\mathbf{x}}}_{\tau,i} = A_\tau \tilde{\mathbf{x}}_{\tau,i} + \begin{bmatrix} 0 \\ 0 \\ \dot{d}_{\tau,i}(t) \end{bmatrix}, \qquad \tilde{\mathbf{x}}_{\tau,i} = \begin{bmatrix} \tilde{\tau}_i \\ \tilde{z}_{\tau 2,i} \\ \tilde{d}_{\tau,i} \end{bmatrix} \tag{54}$$

where

$$A_\tau = \begin{bmatrix} -l_{\tau 1} & 1 & 0 \\ -l_{\tau 2} & 0 & 1 \\ -l_{\tau 3} & 0 & 0 \end{bmatrix} \tag{55}$$

The characteristic polynomial of A_τ is

$$p_\tau(s) = s^3 + l_{\tau 1} s^2 + l_{\tau 2} s + l_{\tau 3} \tag{56}$$

According to the Routh–Hurwitz criterion, if the observer gains satisfy:

$$l_{\tau 1} > 0, \quad l_{\tau 2} > 0, \quad l_{\tau 3} > 0, \quad l_{\tau 1} l_{\tau 2} > l_{\tau 3} \tag{57}$$

A_τ is Hurwitz and the corresponding homogeneous error system is exponentially stable. Furthermore, if $\dot{d}_{\tau,i}(t)$ is bounded, the observer error system is a stable linear system driven by a bounded input. Therefore, the estimation error of the remaining-time channel is uniformly ultimately bounded, and it can converge to a sufficiently small neighborhood of the origin when the disturbance varies slowly.

For the heading channel, the disturbance observer in Eq. (35) has the same structure as that of the remaining-time channel. Define the estimation errors as:

$$\tilde{e}_{\psi,i} = e_{\psi,i} - \hat{e}_{\psi,i}, \quad \tilde{z}_{\psi 2,i} = \dot{e}_{\psi,i} - \hat{z}_{\psi 2,i}, \quad \tilde{d}_{\psi,i} = d_{\psi,i} - \hat{d}_{\psi,i} \tag{58}$$

As above, the observer error dynamics can be expressed as:

$$A_\psi = \begin{bmatrix} -l_{\psi 1} & 1 & 0 \\ -l_{\psi 2} & 0 & 1 \\ -l_{\psi 3} & 0 & 0 \end{bmatrix} \tag{59}$$

whose characteristic polynomial is

$$p_\psi(s) = s^3 + l_{\psi 1} s^2 + l_{\psi 2} s + l_{\psi 3} \tag{60}$$

Similarly, if

$$l_{\psi 1} > 0, \quad l_{\psi 2} > 0, \quad l_{\psi 3} > 0, \quad l_{\psi 1} l_{\psi 2} > l_{\psi 3} \tag{61}$$

A_ψ is Hurwitz and the corresponding observer error system is exponentially stable. Under the boundedness assumption of $\dot{d}_{\psi,i}(t)$, the estimation error of the heading channel is also uniformly ultimately bounded.

4. Experiments and results analysis

In this section, a simulation environment is established to evaluate the proposed algorithm. Based on the simulation results, the effectiveness, performance, generalization capability, and robustness of the formation-control framework are discussed.

The success rate is defined as the ratio of successful formation completions to the total number of tests. The path length refers to the average traveled distance of all USVs, and the number of steps denotes the average number of discrete motion steps taken by the USVs in the successful formation. The final velocity errors quantify the deviation between the actual terminal velocities and the velocities required by the formation specification; similarly, the final yaw errors measure the difference between the actual and the desired heading angles at arrival. These metrics are formally expressed as:

$$path\ length \triangleq \frac{\sum_{i=1}^{n} \int_0^{t^i_{end}} v_t \mathrm{d}t}{n} \tag{62}$$

$$step \triangleq \frac{\sum_{i=1}^{n} \frac{t^i_{end}}{\Delta t}}{n} \tag{63}$$

$$velocity\ error \triangleq \frac{\sum_{i=1}^{n} \left| v_i - v_i^{final} \right|}{n} \tag{64}$$

$$yaw\ error \triangleq \frac{\sum_{i=1}^{n} \left| \psi_i - \psi_i^{final} \right|}{n} \tag{65}$$

where t^i_{end} is the time the ith USV reaches its target, n is the number of USVs in the cluster.

4.1. Experimental scenes

To comprehensively evaluate the performance of the proposed methods under varying environmental complexity, two distinct test scenarios are designed.

Scenario 1: Low-Complexity Environment. This scenario comprises three static obstacles and one dynamic obstacle. The static obstacles are sparsely distributed, while the dynamic obstacle follows a predefined trajectory. This setup is designed to evaluate the fundamental obstacle-avoidance capability of the algorithms and to verify their effectiveness in global path planning and synchronized control.

Scenario 2: High-Complexity Environment. This scenario includes eleven static obstacles and three dynamic obstacles. The high number of static obstacles forms several bottleneck passages, substantially narrowing the feasible paths and increasing the difficulty of local collision avoidance. Meanwhile, multiple dynamic obstacles move in different directions, introducing significant environmental uncertainty. This scenario is therefore used primarily to assess the robustness and stability of the algorithms in dense obstacle fields and under strong dynamic disturbances.

4.2. Formation tests results

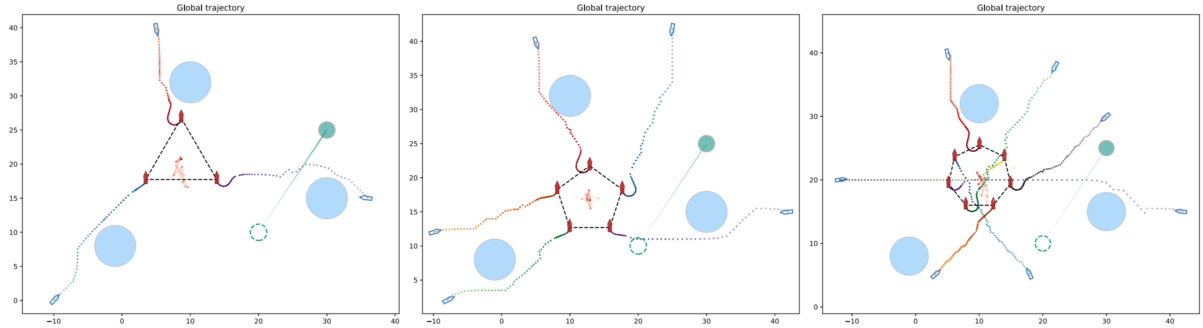
In the simulation experiments, all USVs share the same motion model and are subject to identical control constraints. The initial positions are randomly distributed within the operational area, the initial heading angles are arbitrarily chosen from $[0°, 360°)$, and the initial surge speeds are set to nonzero values.

In the current implementation, the surge speed of each USV is bounded by

$$0 < v_i(t) \le v_{\max} \tag{66}$$

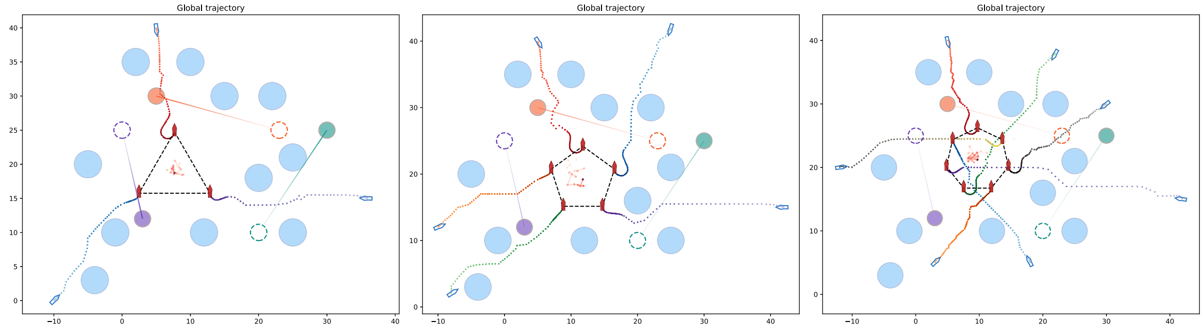
where $v_{\max} = 1.5$ unit/step. To maintain trajectory smoothness and control feasibility, the yaw rate is also limited as

$$|\omega_i(t)| \le \omega_{\max} \tag{67}$$



(a) (b) (c)

Fig. 5. Formations of 3, 5, and 7 USVs in Scene 1.



(a) (b) (c)

Fig. 6. Formations of 3, 5, and 7 USVs in Scene 2.

with $\omega_{\max} = 60^\circ/\text{step}$. During the simulations, control inputs are applied only to the surge thrust and yaw moment. We do not model actuator dynamics (e.g., propulsion lag, rudder servo dynamics) or high-fidelity hydrodynamic effects (e.g., added mass, nonlinear damping, parameter uncertainties). Therefore, the results should be interpreted under the assumption of ideal low-level actuation.

As shown in Fig. 5, in Scenario 1, after quickly determining the formation position based on requirements and environmental parameters, all USVs start from dispersed initial positions and smoothly converge to their assigned target locations while avoiding obstacles. For each formation size (3, 5, and 7 USVs), the desired formation is successfully achieved. The overall trajectories appear smooth, without noticeable path congestion or collisions.

Fig. 6 presents the corresponding results in Scenario 2. Owing to the substantially larger number of static obstacles and stronger dynamic disturbances, the USVs must execute frequent obstacle-avoidance maneuvers and path adjustments during navigation. Despite the increased environmental difficulty, all USVs still reach their designated target positions after necessary detours and corrections, eventually forming the prescribed formations.

Table 1 presents a quantitative comparison of formation quality across varying fleet sizes, formation geometries, and environmental complexities. Overall, the proposed A*-ETR framework successfully generates formations with higher success rates under all tested conditions, achieving acceptable terminal synchronization errors.

Performance declines with increasing environmental complexity and fleet size. In the high-complexity setting of Scenario 2, denser obstacles and narrower passages force longer detours, amplify path-planning disturbances, and consistently elevate velocity and yaw errors. As the number of USVs increases from 3 to 7, the growing coordination difficulty leads to progressively lower success rates and higher terminal errors. The influence of formation geometry is secondary, with the Polygon shape typically requiring slightly longer paths and yielding marginally larger errors than the simpler Line formation under identical conditions.

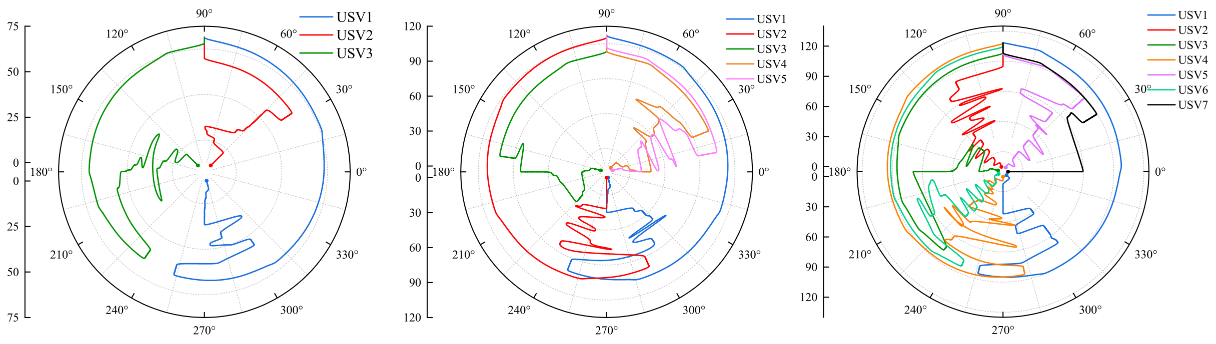
The primary cause of failure is target-point infeasibility. During ETR optimization, target points are generated from a fixed geometric template. In dense environments, some points may fall inside obstacles or in regions that violate safety constraints, causing the A* planner to fail and stalling the entire planning-control loop. This issue is most pronounced in Scenario 2 with larger formations, explaining the observed decline in success rates.

The heading angle of each USV is presented in Figs. 7 and 8, respectively. During the formation, the heading angles exhibit relatively frequent variations. These variations arise from the combined influence of several factors typical of complex environments, including path replanning, heading-rate constraints, and the closed-loop interaction between navigation and formation control.

The proposed method employs an ETR strategy to dynamically update the global formation center in response to USV states and obstacle distribution. Once the ETR-optimized center shifts, the target position for each USV is recomputed, which in turn triggers A*-based path replanning toward the updated setpoints. Consequently, each USV continuously adjusts its heading to follow the newly replanned path, giving rise to the rapid variations observed in the heading-angle profiles. This dynamic reflects the close coupling between formation-position optimization and reactive path planning.

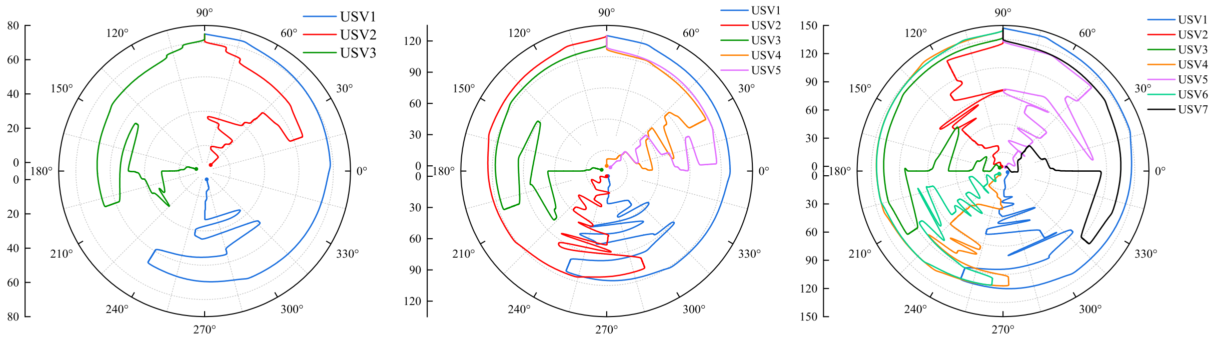
Figs. 9 and 10 present the speed profiles of formations consisting of 3, 5, and 7 USVs in Scenario 1 and Scenario 2, respectively.

Instead of enforcing identical speeds, the proposed controller adjusts each USV's velocity to synchronize the remaining arrival times. This coordination results in the purposeful, non-uniform speed variations observed in the early stage. As the formation advances, the differences in remaining arrival times diminish, and all speeds converge to the prescribed cruising value, confirming successful spatiotemporal synchronization.



(a) (b) (c)

Fig. 7. Yaw of 3, 5, and 7 USVs during formation in Scene 1.



(a) (b) (c)

Fig. 8. Yaw of 3, 5, and 7 USVs during formation in Scene 2.

Table 1
Formation performance reported separately for each fleet size, formation geometry, and scenes.

N	Formation	Success rate (%)		Avg steps		Avg path		Velocity error (%)		Yaw error (°)	
		Scene 1	Scene 2	Scene 1	Scene 2	Scene 1	Scene 2	Scene 1	Scene 2	Scene 1	Scene 2
3	Line	97.8	95.2	75.4	86.9	33.4	36.5	2.9	3.8	0.12	0.20
	V	97.1	94.7	78.1	89.2	34.7	38.3	3.4	4.6	0.15	0.28
	Polygon	96.4	94.1	71.5	83.6	36.1	39.6	3.7	4.9	0.12	0.31
5	Line	93.3	88.2	101.3	121.1	35.9	38.4	3.8	4.7	0.18	0.28
	V	91.8	87.5	105.3	123.4	36.4	43.2	4.3	5.4	0.24	0.34
	Polygon	89.8	86.9	114.5	127.3	37.8	45.1	4.6	5.2	0.30	0.40
7	Line	85.1	75.3	120.5	140.9	36.3	43.4	4.9	5.8	0.33	0.36
	V	84.4	74.0	121.6	142.6	37.9	44.5	5.4	6.3	0.32	0.42
	Polygon	83.6	73.8	127.8	149.7	40.5	46.8	5.8	6.1	0.38	0.37

Compared with Scenario 1, Scenario 2 exhibits larger speed fluctuations. This is primarily due to frequent local replanning in obstacle-dense environments, which introduces abrupt changes in L_{rem} and $\tau = L_{rem}/v$. Although the integrated disturbance observer compensates for these variations, residual speed oscillations remain unavoidable under high environmental complexity.

To further investigate the source of the velocity oscillations in Fig. 10, the corresponding surge control inputs are presented in Fig. 11. The control inputs remain bounded and relatively smooth throughout the process and do not exhibit dense high-frequency switching typically associated with severe SMC chattering.

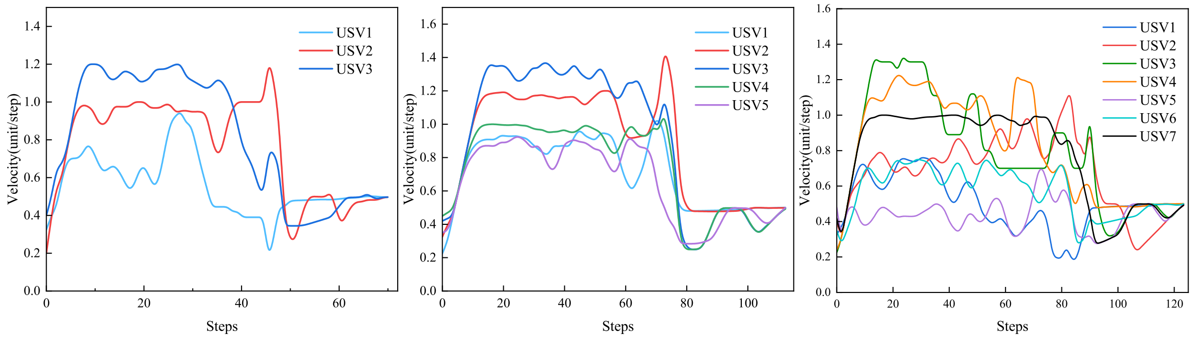
Each USV and obstacle is assigned a safety radius $R = 1$, and replanning is triggered when the pairwise distance falls below the threshold d_{trig}. The resulting trigger statistics and the observed minimum inter-vessel/obstacle distances are summarized in Table 2 for $d_{\text{trig}} = 4.0$. As shown therein, no triggers occur in smaller-scale cases where the minimum distance is larger than 4.0. Trigger counts increase with fleet size and environmental density, particularly for $N = 7$, and the trigger-step ranges indicate that events concentrate in the later stage of formation when the fleet is more compact near the target configuration.

Table 2
Safety-trigger statistics across different fleet sizes and environments.

N	Environment	Recorded triggers	Min distance (unit)	Trigger step range (steps)
3	Scene 1	0	9.82	–
	Scene 2	0	9.47	–
5	Scene 1	0	5.86	–
	Scene 2	61	3.91	106–115
7	Scene 1	154	3.86	112–123
	Scene 2	322	3.74	132–144

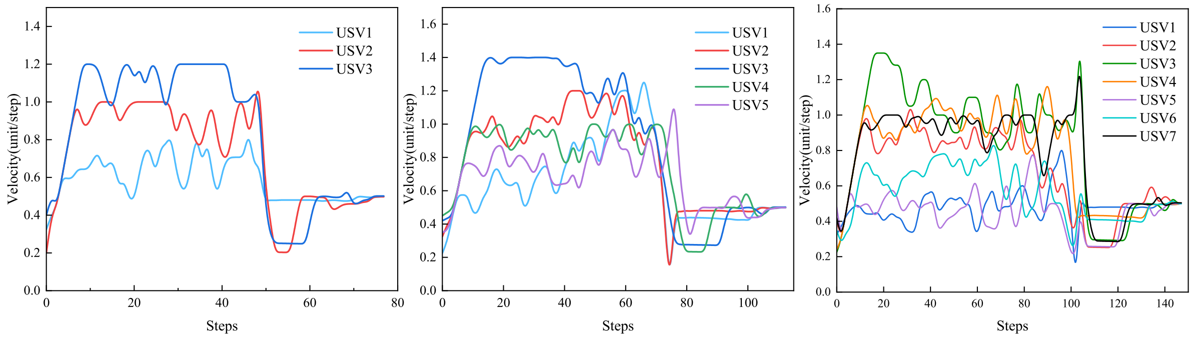
4.3. Analysis of comparative experiments

To systematically evaluate the performance of the proposed method, comparative experiments are conducted in Scenario 1 and Scenario 2. The proposed approach is benchmarked against four methods: the HG-PPO method introduced in Zhong et al. (2026), MAPPO, MADDPG, and the Artificial Potential Field (APF) method. All methods are tested



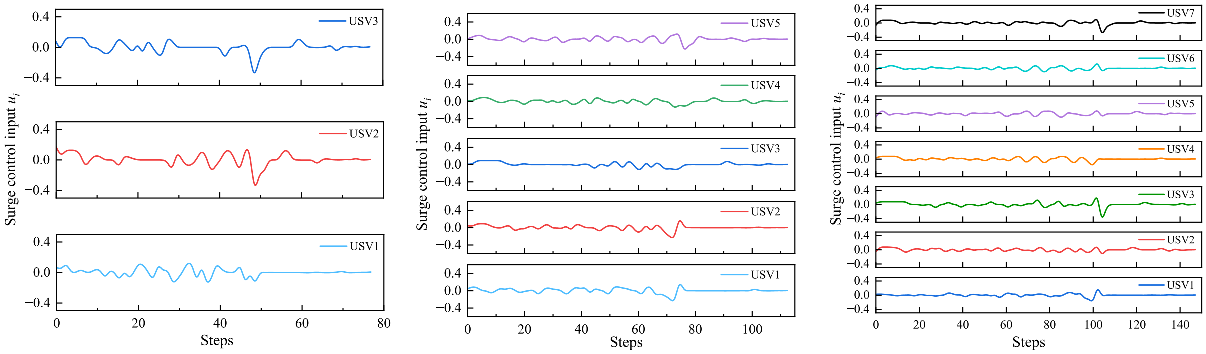
(a) (b) (c)

Fig. 9. Velocity of 3, 5, and 7 USVs during formation in Scene 1.



(a) (b) (c)

Fig. 10. Velocity of 3, 5, and 7 USVs during formation in Scene 2.



(a) (b) (c)

Fig. 11. Time-domain surge control inputs of 3, 5 and 7 USVs in Scene 2.

under identical environmental layouts, USV kinematic models, motion constraints, safety constraints, fleet sizes, and formation geometries. For MAPPO and MADDPG, a unified reward structure is adopted:

$$r_t = -\alpha_1 - \alpha_2 \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{p}_i - \mathbf{p}_{i,d}\| - \alpha_3 e_{\text{sync}} - \alpha_4 \Phi_{\text{safe}} + \alpha_5 \Phi_{\text{term}} \tag{68}$$

where $\alpha_1, \dots, \alpha_5$ are weighting coefficients, $\mathbf{p}_i$ and $\mathbf{p}_{i,d}$ denote the current and target positions of the ith USV, e_{sync} denotes the synchronization error, Φ_{safe} and Φ_{term} denote the safety penalty and terminal reward, respectively.

Key hyperparameters, the training budget, and the number of random seeds are listed in Table 3. Baseline parameters were adopted from commonly reported values in the original studies and then moderately adjusted under the same training budget. Performance is evaluated via success rate and mean steps to formation completion, with results summarized in Table 4.

Table 3
Key hyperparameters, training budgets, and random seeds of MAPPO and MADDPG.

Item	MAPPO	MADDPG
Learning rate	5×10^{-4}	1×10^{-3}
Mini-batch size	256	256
Discount factor γ	0.99	0.95
Clip ratio	0.2	–
Training budget	20 000 eps.	20 000 eps.
Number of random seeds	5	5

As shown in Table 4, under identical fleet sizes and formation geometries, the proposed method exhibits a moderate performance decline when transitioning from the low-complexity Scenario 1 to the high-complexity Scenario 2. Specifically, its success rate drops by only

Table 4
Experimental results of different algorithms under different environments and formation shapes.

N	Environment	Method	Success rate (%)			Average formation steps		
			Line	V	Polygon	Line	V	Polygon
3	Scene 1	**Our method**	**97.8**	**97.1**	**96.4**	**75.4**	**78.1**	**71.5**
		HG-PPO	95.6	95.0	94.8	83.2	84.7	79.6
		MAPPO	95.1	94.5	94.0	86.7	96.2	81.3
		APF	94.1	93.2	92.8	78.4	88.5	74.5
		MADDPG	93.4	92.6	92.1	92.2	101.1	87.5
	Scene 2	**Our method**	**95.2**	**94.7**	**94.1**	**86.9**	**89.2**	**83.6**
		HG-PPO	92.3	92.0	91.7	95.1	103.6	87.2
		MAPPO	91.6	91.1	90.4	99.4	108.6	91.7
		APF	89.3	88.5	87.7	88.4	96.5	81.5
		MADDPG	87.8	86.9	86.1	112.4	122.6	108.4
5	Scene 1	**Our method**	**93.3**	**91.8**	**89.8**	**101.3**	**105.3**	**114.5**
		HG-PPO	86.8	85.7	84.1	162.4	171.6	183.2
		MAPPO	85.9	85.1	83.7	156.3	165.8	177.6
		APF	82.8	81.2	79.1	145.6	154.4	165.3
		MADDPG	80.6	79.4	77.8	162.5	171.4	182.6
	Scene 2	**Our method**	**88.2**	**87.5**	**86.9**	**121.1**	**123.4**	**127.3**
		HG-PPO	79.5	78.3	76.1	185.7	193.5	205.8
		MAPPO	78.6	77.4	75.8	179.8	188.9	201.0
		APF	70.0	71.8	69.1	168.1	176.3	188.4
		MADDPG	67.8	66.5	65.1	200.1	210.5	222.3
7	Scene 1	**Our method**	**85.1**	**84.4**	**83.6**	**120.5**	**121.6**	**127.8**
		HG-PPO	78.4	77.1	75.6	228.3	241.6	254.8
		MAPPO	77.5	76.3	74.8	224.1	237.5	250.4
		APF	71.4	69.6	67.6	227.3	227.3	232.3
		MADDPG	68.6	67.3	65.9	224.7	238.9	252.3
	Scene 2	**Our method**	**75.3**	**74.0**	**73.8**	**140.9**	**142.6**	**149.7**
		HG-PPO	62.3	60.8	58.9	269.4	274.6	282.6
		MAPPO	60.9	59.6	57.7	261.7	269.8	279.3
		APF	54.3	52.3	53.0	245.2	252.6	262.6
		MADDPG	51.5	50.2	49.8	275.4	286.7	298.6

3%–12%, and the average formation steps increase by about 10–40 steps. It should be clarified that the Average Formation Steps are the discrete control simulation steps of the navigation layer. In contrast, the benchmark methods suffer more severe degradation under dense obstacles and narrow passages, with larger reductions in success rates and substantially longer formation steps.

As the formation scales from 3 to 5 and then 7 USVs under the same environmental and geometric conditions, the proposed method shows a gradual decrease in success rate (about 12%–15% in Scenario 1 and 18%–22% in Scenario 2) and a controlled rise in average steps (from roughly 70–90 to 120–150 steps). By comparison, the other methods degrade more sharply, often exceeding 250 steps in the 7-USV case.

Across different formation shapes (Line, V-shape, Polygon), the proposed method maintains stable performance. For a given environment and fleet size, success-rate differences remain within 1%–3%, and step counts vary by no more than 10–20 steps, even in Scenario 2 or with larger fleets. The benchmark methods, however, display significantly larger performance variations across shapes, especially in complex or large-scale settings.

In all, our method maintains higher success rates and requires significantly fewer formation steps across all tested conditions. Notably, the performance degradation of our method under increased environmental complexity and larger fleet scales is more gradual and controlled compared to other methods, which suffer from sharper declines in success rates and substantially longer step counts, especially in complex, large-scale scenarios. Furthermore, our method shows stable performance across different formation shapes, with minimal variation, whereas competing methods exhibit larger performance fluctuations depending on the geometry.

4.4. Analysis of ETR and A* disturbance observer

4.4.1. Analysis of A*-ETR

To validate the efficiency of the elliptical trust-region optimization strategy, we compared its convergence behavior with that of the conventional circular trust-region method. Both TR and ETR use the same objective function (Eq. (6)), update rules, stopping criteria, initial formation center, environmental setup, and formation configuration. The trust-region radius is adjusted adaptively based on the reduction of the objective value. Optimization stops when the change in the objective falls below a preset threshold or the maximum iteration count is reached. Experiments are conducted in Scenario 1 with 3 USVs, and the results are shown in Fig. 12.

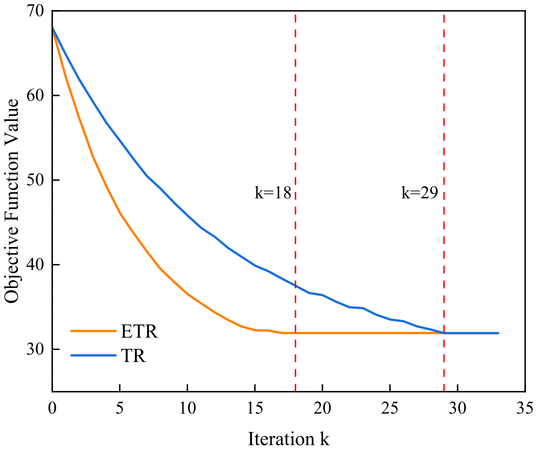


Fig. 12. Convergence behavior of TR and ETR.

Fig. 12 shows that both methods exhibit monotonic convergence, reducing the objective function steadily until stable convergence is achieved. However, ETR descends faster in early iterations and stabilizes after about 18 steps, whereas TR requires roughly 29 iterations to reach a comparable level. This demonstrates that the elliptical scaling in ETR better exploits directional search information, reduces

Table 5
A*-ETR stability statistics with different initialization and obstacle configurations.

Case	Samples	Succ. rate (%)	Final Obj. ($\mu \pm \sigma$)	Loc. Fluct. ($\mu \pm \sigma$)
ETR cases	100	96.0	36.38 ± 1.82	2.01 ± 0.31

unproductive trial steps, and thus accelerates convergence. Once near the optimum, the two methods behave similarly, indicating that ETR improves convergence speed without compromising solution quality. It should be clarified that the “Iteration k” refers to the optimization-layer iterations of the ETR algorithm (Algorithm 1), rather than physical motion steps.

To further assess the stability of the A*-ETR, a supplementary experiment was conducted for the 3-USV case, covering only the optimization stage from the initialization center to the first optimized formation center. Obstacle complexity was at the same level as in Scenario 2. A total of 100 ETR test cases were constructed, with results summarized in Table 5. The table reports the success rate, final objective value, and local fluctuation indicator J_{local}. The final objective value reflects the objective-function level after optimizing the first formation center; J_{local} characterizes local variations of the convergence curve around its overall descending trend, defined as

$$J_{\text{local}} = \sqrt{\frac{1}{T}\sum_{k=1}^{T}\left(f_k - \bar{f}(k)\right)^2} \tag{69}$$

where f_k is the objective value at the kth iteration, $\bar{f}(k)$ is the smoothed trend of f_k, and T is the total number of iterations.

As shown in Table 5, A*-ETR successfully optimized the first formation center with a success rate of 96%. The final objective value exhibits modest variation across samples, as changes in obstacle distribution directly reshape the A*-induced objective landscape. Meanwhile, the local fluctuation indicator remains relatively low, indicating only mild local oscillations in the convergence curves.

Table 6 shows that the trigger frequencies of all backup methods increase with environmental complexity and fleet size. B1 remains the dominant method, indicating that most infeasible instances are resolved by shrinking the trust region and re-solving the local subproblem, while B2 and B3 offer additional safeguards in more challenging scenarios. Compared with the no-backup baseline, the proposed mechanism increases runtime by only 4.1%–9.1%. The residual failures are mainly associated with persistent target-point infeasibility, limited feasibility recovery within the allowed backup budget, and repeated infeasibility in dense-obstacle configurations.

Collectively, these results suggest that A*-ETR maintains stable numerical behavior under variations in both initialization center and obstacle arrangement.

4.4.2. Analysis of A disturbance observer*

To validate the effectiveness of the proposed A*-based disturbance observer in improving dynamic performance under complex conditions, this study selects the high-complexity Scenario 2 and compares the speed and heading responses of 3 USVs with and without the observer. All experiments use identical initial conditions, planning parameters, control settings, and obstacle layouts.

Table 7 compares the averaged formation performance in Scene 2 with and without the disturbance observer, where the results are aggregated over the three formation geometries for each fleet size. The results show that, without the observer, the success rate drops slightly but consistently across all fleet sizes—by about 1–2%—with the reduction becoming more pronounced as the formation scale increases. This trend suggests that in obstacle-dense environments with frequent replanning, uncompensated equivalent disturbances accumulate during the terminal synchronization phase, gradually degrading the system’s ability to meet strict terminal constraints.

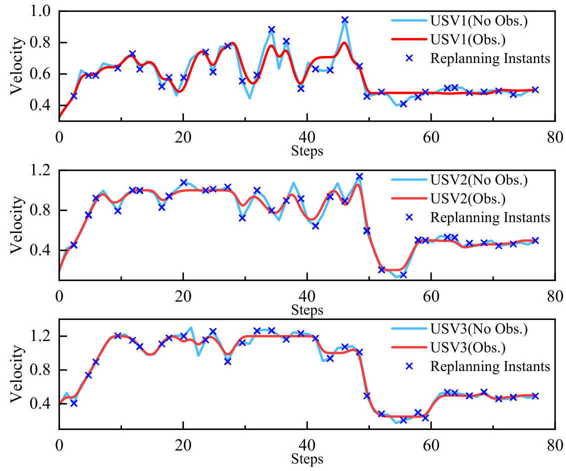


Fig. 13. Velocity comparison of 3 USVs in Scene 2.

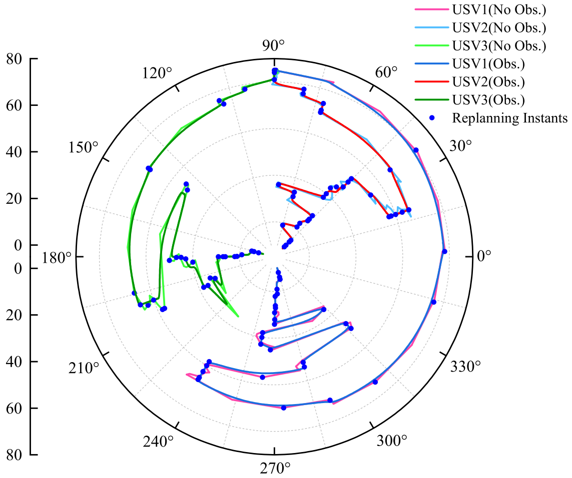


Fig. 14. Yaw angle comparison of 3 USVs in Scene 2.

The average number of formation steps is more noticeably affected by the absence of the observer. Without compensation, step counts increase by roughly 5%–12% for all fleet sizes. This indicates that, when replanning-induced variations in remaining path length and reference heading are not actively mitigated, the sliding-mode controller must rely on more frequent speed and heading adjustments to maintain arrival-time consistency, thereby prolonging the convergence process.

As shown in Figs. 13 and 14, in obstacle-dense areas and during frequent path replanning, velocities exhibit abrupt accelerations and decelerations, while heading angles display high-frequency oscillations interspersed with large corrective turns.

Figs. 15 and 16 show the estimation errors of the remaining-time disturbance and the yaw disturbance for the three USVs in Scene 2. It can be observed that the estimation errors remain small and fluctuate around zero during the process. Although slight fluctuations appear at several instants due to path replanning and maneuver adjustments, the errors quickly return to a small range.

With the A*-based disturbance observer, both speed and heading profiles become markedly smoother. By estimating the disturbances induced by path replanning and injecting appropriate compensation into the control law, the observer mitigates the direct propagation of sudden path-level changes to the control layer. Consequently, the controller no longer triggers frequent aggressive accelerations, decelerations, or

Table 6
Hierarchical backup statistics under different fleet sizes and environments.

N	Environment	B1 trigger (%)	B2 trigger (%)	B3 trigger (%)	Residual failure (%)	Runtime increase (%)
3	Scene 1	8.6	1.8	0.4	2.2	4.1
	Scene 2	15.8	4.6	1.5	4.1	5.6
5	Scene 1	12.9	3.5	0.9	7.1	5.0
	Scene 2	23.4	8.7	3.2	10.8	7.2
7	Scene 1	18.7	5.9	1.8	13.9	6.3
	Scene 2	31.6	13.4	5.8	23.2	9.1

Table 7
Comparison of formation quality with/without the observer in Scene 2, averaged over the 3 formation geometries.

N	Success rate (%)		Avg steps		Avg path		Velocity error (%)		Yaw error (°)	
	No Obs.	Obs.	No Obs.	Obs.	No Obs.	Obs.	No Obs.	Obs.	No Obs.	Obs.
3	93.92	94.67	94.96	86.57	40.23	38.13	4.57	4.43	0.276	0.264
5	86.63	87.53	137.81	123.93	44.63	42.23	5.36	5.10	0.365	0.342
7	73.16	74.37	161.87	144.40	48.01	44.90	6.49	6.07	0.398	0.381

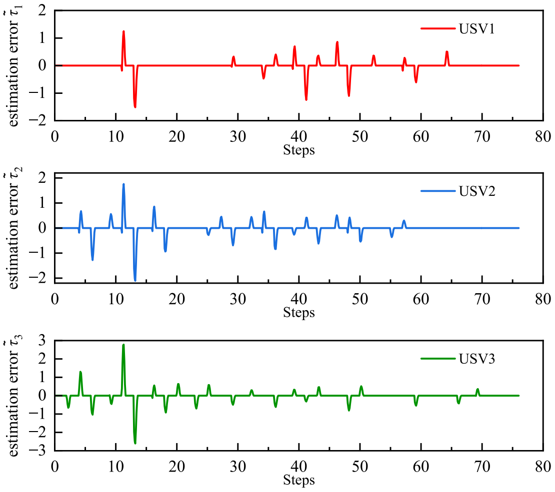


Fig. 15. Estimation errors of the remaining-time of 3 USVs in Scene 2.

abrupt heading corrections during synchronization, leading to a more continuous and stable dynamic response of the overall system.

To quantitatively evaluate fluctuations in speed and heading during formation, two metrics are adopted: the root-mean-square (RMS) of the incremental changes and the standard deviation of the residual signal.

The RMS of the incremental variation captures the overall magnitude of step-to-step state changes:

$$\mathrm{RMS}_{\Delta y} = \sqrt{\frac{1}{T-1}\sum_{k=1}^{T-1}\left(y(k+1)-y(k)\right)^2} \tag{70}$$

where $y(k)$ denotes the speed $v(k)$ or heading angle $\psi(k)$ at time step k, and T is the total number of steps.

To further characterize high-frequency oscillations, the residual signal is obtained by removing the low-frequency trend (using a moving-average filter). Its standard deviation is defined as:

$$\sigma(\mathrm{residual}) = \sqrt{\frac{1}{T}\sum_{k=1}^{T}\left(y(k)-\bar{y}(k)\right)^2} \tag{71}$$

where $\bar{y}(k)$ is the smoothed trend of the original signal $y(k)$.

As shown in Table 8, after introducing the disturbance observer, the RMS of speed increments ($\mathrm{RMS}_{\Delta v}$) and the residual standard deviation $\sigma(\mathrm{residual}_v)$ decrease by about 25% and 28%, respectively, indicating effective suppression of high-frequency speed oscillations. Similarly, the RMS of heading-angle increments ($\mathrm{RMS}_{\Delta\psi}$) and the residual standard deviation $\sigma(\mathrm{residual}_\psi)$ are reduced by approximately 16% and 17%, confirming that the observer also alleviates heading fluctuations.

Table 8
Comparison of speed and heading fluctuation metrics with and without the disturbance observer in Scene 2 (3 USVs).

Metric type	Indicator	Value		Improvement
		With observer	Without observer	(%)
Speed	$\mathrm{RMS}_{\Delta v}$	0.10295	0.13776	25.26
	$\sigma(\mathrm{residual}_v)$	0.08102	0.11277	28.16
Heading	$\mathrm{RMS}_{\Delta\psi}$	11.25149	13.23705	16.31
	$\sigma(\mathrm{residual}_\psi)$	8.19108	9.98912	17.42

Although the improvement in heading-related metrics is somewhat smaller than that for speed, the overall trend is consistent. These results verify that the proposed observer contributes significantly to enhancing the dynamic smoothness of the system in complex environments.

4.4.3. Control-input analysis

To examine local control variations around replanning instants more directly, a branch-comparison analysis was conducted, as shown in Fig. 17. At each replanning instant, two post-event branches were initialized from the same system state: With replanning, which retains the online replanning mechanism, and No replanning, which discards the current replanning update while the sliding-mode controller continues along the previous reference. Control inputs were then compared over the interval until the next original replanning instant. Three quantitative metrics were further computed: the peak command increment $\Delta u_{\max}^{(r)}$, the input-rate peaks $\dot{u}_{\max}^{(r)}$, and the total variation TV_u.

$$\Delta u_{\max}^{(r)} = \max_{n,n+1\in I_r} |u[n+1]-u[n]| \tag{72}$$

$$\dot{u}_{\max}^{(r)} = \max_{n,n+2\in I_r} |u[n+2]-2u[n+1]+u[n]| \tag{73}$$

$$TV_u = \sum_{r\in\mathcal{R}}\sum_{n\in I_r} |u[n+1]-u[n]| \tag{74}$$

Here, I_r is the rth replanning interval, $\mathcal{R}$ is the set of all replanning intervals, $u[n]$ is the surge control input at step n.

As shown in Table 9, the With replanning branch yields consistently larger values across all metrics; notably, TV_u rises from 1.0867 to 4.2826. Discarding the replanning update still allows closed-loop SMC regulation, yet local variations become markedly smaller. Thus, the observed oscillations stem chiefly from replanning-triggered reference updates, not residual SMC switching.

Table 9

Branch-comparison for the surge control input of USV1.

Case	Intervals	Max ($\Delta u_{\max}$)	Max ($\dot{u}_{\max}$)	TV_u
Replanning	33	0.2962	0.2506	4.2826
No replanning	33	0.0965	0.0834	1.0867

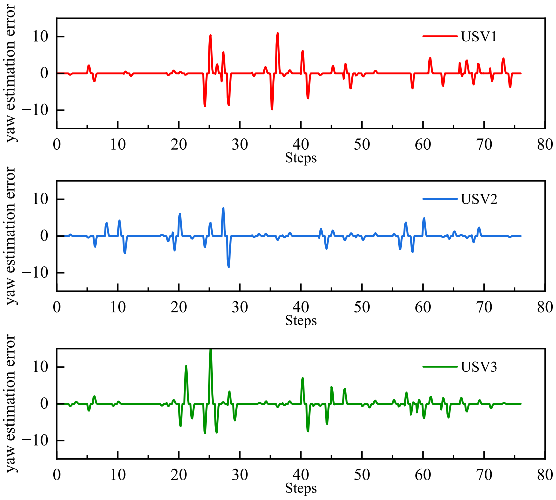


Fig. 16. Estimation errors of the yaw of 3 USVs in Scene 2.

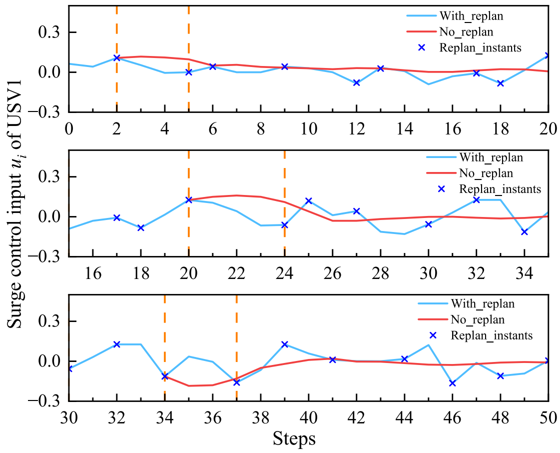


Fig. 17. Local branch-comparison of surge control inputs around replanning instants for USV1 in Scene 2.

Since actuator saturation, rate limits, and detailed actuator dynamics are highly specific to a given USV platform and are not explicitly incorporated in the present model, the reported results characterize local input variations at the control-layer level only and do not constitute a strict hardware-safety guarantee.

4.5. Effect of the realistic factors

To further evaluate the robustness of the proposed method under more realistic ocean conditions, a time-varying ocean disturbance is introduced in Scenario 2. The ocean disturbance is incorporated as an additional velocity term in the planar motion equations, expressed as

$$\dot{x} = v\cos\psi + A\sin(\omega t), \qquad \dot{y} = v\sin\psi + A\cos(\omega t) \tag{75}$$

where A and ω denote the amplitude and variation frequency of the ocean disturbance, respectively.

To evaluate its performance under ocean disturbances, we repeated the simulations and summarized the results in Table 10. The data show only a marginal decline in success rates (0.95%–2.79%), with moderate increases in average steps, path length, and velocity/yaw errors. These findings indicate that the method remains effective under the tested ocean-disturbance setting, with only moderate performance degradation. Although the impact scales with fleet size — larger formations exhibit greater performance degradation — the framework continues to complete the formation task across the tested simulation cases.

Table 10

Formation performance of the proposed method with/without ocean currents in Scene 2.

N	Metric	No current	Time-varying current	Change
3	Success rate	94.67%	93.72%	−0.95%
	Avg steps	86.57	92.61	+6.04
	Avg path	38.13	39.46	+1.33
	Velocity error	4.43%	4.68%	+0.25%
	Yaw error	0.264°	0.286°	+0.022°
5	Success rate	87.53%	85.96%	−1.57%
	Avg steps	123.93	132.84	+8.91
	Avg path	42.23	43.95	+1.72
	Velocity error	5.10%	5.46%	+0.36%
	Yaw error	0.342°	0.371°	+0.029°
7	Success rate	74.37%	71.58%	−2.79%
	Avg steps	144.40	156.27	+11.87
	Avg path	44.90	46.88	+1.98
	Velocity error	6.07%	6.42%	+0.35%
	Yaw error	0.381°	0.428°	+0.047°

Table 11

Formation performance of the proposed method with different communication constraints in Scene 2.

N	Metric	Ideal communication	Communication constraints	Degradation
3	Success rate	94.7%	92.4%	−2.3%
	Avg steps	86.6	96.8	+10.2
	Avg path	38.13	40.14	+2.01
	Velocity error	4.43%	5.12%	+0.69%
	Yaw error	0.26°	0.32°	+0.06°
5	Success rate	87.5%	83.9%	−3.6%
	Avg steps	123.9	139.5	+15.6
	Avg path	42.23	45.47	+3.24
	Velocity error	5.10%	6.08%	+0.98%
	Yaw error	0.34°	0.43°	+0.09°
7	Success rate	74.4%	69.2%	−5.2%
	Avg steps	144.4	166.7	+22.3
	Avg path	44.90	49.36	+4.46
	Velocity error	6.07%	7.26%	+1.19%
	Yaw error	0.38°	0.49°	+0.11°

Beyond the idealized communication assumption, an empirical stress test under communication constraints is included in Scenario 2, injecting random delays of 1–3 control steps and occasional information unavailability into inter-USV data sharing. Simulation results under communication constraints are summarized in Table 11.

Compared to the ideal communication case, the success rate decreases slightly, while average steps, path length, and velocity/yaw errors increase to some extent. Notably, this performance degradation becomes more pronounced with larger fleets: success rates drop by 2.3% for N = 3 and by 5.2% for N = 7, accompanied by substantially greater increases in steps, path length, and control errors. Despite these challenges, the overall degradation remains bounded, and the system stably accomplishes the formation task across all tested scenarios. These results suggest that the proposed method retains acceptable performance under the tested communication-constrained conditions, though the impact scales with fleet size—highlighting an important consideration for large-scale deployments.

5. Conclusion

The proposed framework enables efficient and synchronized formation assembly for multi-USV systems. The sliding-mode controller and disturbance observer are theoretically proven to be convergent and stable under the stated assumptions, and simulations across various scenarios confirm the effectiveness and strong numerical performance of the overall framework.

Compared to benchmark methods under identical conditions, the proposed approach achieves significantly higher success rates (ranging from 75.3% to 97.8%) while requiring remarkably fewer formation steps (71.5–149.7). It maintains terminal velocity errors within 6% and heading errors below 0.5°. Notably, the elliptical trust region accelerates optimization-layer convergence, and the integrated disturbance observer reduces speed and heading fluctuations by 16%–28%, thereby improving dynamic smoothness in the tested simulation scenarios.

The proposed framework assumes ideal communication and sufficient environmental knowledge — both often impractical in marine settings — and lacks actuator dynamics or detailed hydrodynamic modeling (only simple amplitude/rate limits were enforced). Future work should pursue decentralized coordination, environmental model integration, and realistic propulsion/rudder models with parameter uncertainties to validate the framework under real-world conditions.

CRediT authorship contribution statement

Weibo Zhong: Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Shaohan Cai:** Writing – original draft, Software, Methodology, Formal analysis. **Yang Gu:** Writing – original draft, Methodology. **Xin Su:** Writing – original draft, Methodology. **Yazhou Zhu:** Writing – review & editing, Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank reviewers for their detailed and constructive comments, which helped us improve the quality of this work. This work was supported by National Natural Science Foundation of China (12172308) and National Key Research and Development Program of China (2020YFA0710902).

Data availability

The authors do not have permission to share data.

References

Cui, Z., Guan, W., Zhang, X., 2024. USV formation navigation decision-making through hybrid deep reinforcement learning using self-attention mechanism. Expert Syst. Appl. 256, 124906. http://dx.doi.org/10.1016/j.eswa.2024.124906.

Dong, Z., Tan, F., Yu, M., Xiong, Y., Li, Z., 2024. A bio-inspired sliding mode method for autonomous cooperative formation control of underactuated USVs with ocean environment disturbances. J. Mar. Sci. Eng. 12, http://dx.doi.org/10.3390/jmse12091607.

Er, M.J., Ma, C., Liu, T., Gong, H., 2023. Intelligent motion control of unmanned surface vehicles: A critical review. Ocean Eng. 280, 114562. http://dx.doi.org/10.1016/j.oceaneng.2023.114562.

Fang, Y., Lavaei, J., Na, S., 2025. High probability complexity bounds of trust-region stochastic sequential quadratic programming with heavy-tailed noise. URL: https://arxiv.org/abs/2503.19091.

Gu, Y., Wang, P., Rong, Z., Wei, H., Yang, S., Zhang, K., Tang, Z., Han, T., Si, Y., 2024. Vessel intrusion interception utilising unmanned surface vehicles for offshore wind farm asset protection. Ocean Eng. 299, 117395. http://dx.doi.org/10.1016/j.oceaneng.2024.117395.

Jin, K., Wang, J., Wang, H., Liang, X., Guo, Y., Wang, M., Yi, H., 2022. Soft formation control for unmanned surface vehicles under environmental disturbance using multi-task reinforcement learning. Ocean Eng. 260, 112035. http://dx.doi.org/10.1016/j.oceaneng.2022.112035.

Li, C., Yao, L., Mi, C., 2025. Fusion algorithm based on improved A* and DWA for USV path planning. J. Mar. Sci. Appl. 24, 224–237. http://dx.doi.org/10.1007/s11804-024-00434-1.

Liu, S., Chen, H., Shen, M., Chen, Y., 2026. Distributed H_∞ time-varying formation tracking of 3-DOF USvs: A singular control approach. Nonlinear Dynam. 114 (1), 11. http://dx.doi.org/10.1007/s11071-025-11887-6.

Matarić, M.J., 1995. Issues and approaches in the design of collective autonomous agents. Robot. Auton. Syst. 16, 321–331. http://dx.doi.org/10.1016/0921-8890(95)00053-4.

Meng, J., Tan, H., Jiang, L., Qian, C., Xiao, H., Hu, Z., Li, G., 2024. Robust finite-time sliding mode control of unmanned surface vehicle with active compensation of pose estimation uncertainty. Ocean Eng. 304, 117831. http://dx.doi.org/10.1016/j.oceaneng.2024.117831.

Mohammadi, A., Hajinezhad, D., Garcia, A., 2025. A trust-region approach for computing Pareto fronts in multiobjective derivative-free optimization. Optim. Lett. 19, 233–266. http://dx.doi.org/10.1007/s11590-024-02173-2.

Mou, Z., Zhang, Y., Gao, F., Wang, H., Zhang, T., Han, Z., 2021. Deep reinforcement learning based three-dimensional area coverage with UAV swarm. IEEE J. Sel. Areas Commun. 39, 3160–3176. http://dx.doi.org/10.1109/JSAC.2021.3088718.

Sang, T., Xiao, J., Xiong, J., Xia, H., Wang, Z., 2023. Path planning method of unmanned surface vehicles formation based on improved A* algorithm. J. Mar. Sci. Eng. 11, 176. http://dx.doi.org/10.3390/jmse11010176.

Waltz, R.A., Morales, J.L., Nocedal, J., Orban, D., 2006. An interior algorithm for nonlinear optimization that combines line search and trust region steps. Math. Program. 107, 391–408. http://dx.doi.org/10.1007/s10107-004-0560-5.

Wang, N., Gao, Y., Zhang, X., 2021. Data-driven performance-prescribed reinforcement learning control of an unmanned surface vehicle. IEEE Trans. Neural Netw. Learn. Syst. 32 (12), 5456–5467. http://dx.doi.org/10.1109/TNNLS.2021.3056444.

Wang, P., Liu, R., Tian, X., Zhang, X., Qiao, L., Wang, Y., 2023. Obstacle avoidance for environmentally-driven USVs based on deep reinforcement learning in large-scale uncertain environments. Ocean Eng. 270, 113670. http://dx.doi.org/10.1016/j.oceaneng.2023.113670.

Xiaofei, Y., Yilun, S., Wei, L., Hui, Y., Weibo, Z., Zhengrong, X., 2022. Global path planning algorithm based on double DQN for multi-tasks amphibious unmanned surface vehicle. Ocean Eng. 266, 112809. http://dx.doi.org/10.1016/j.oceaneng.2022.112809.

Xing, B., Yu, M., Liu, Z., Tan, Y., Sun, Y., Li, B., 2023. A review of path planning for unmanned surface vehicles. J. Mar. Sci. Eng. 11, 1556. http://dx.doi.org/10.3390/jmse11081556.

Xu, R., Chen, H., Zheng, J., Zhang, J., Xue, M.-A., Long, X., Tang, Y., 2026. Reinforcement learning-based iterative learning consensus control for multiple unmanned surface vessels with time-varying delay. J. Franklin Inst. http://dx.doi.org/10.1016/j.jfranklin.2025.108284.

Yu, Z., Yuan, Y., Tian, P., 2024. An efficient trust region algorithm with bounded iteration sequence for unconstrained optimization and its application in support vector machine. J. Comput. Appl. Math. 449, 115956. http://dx.doi.org/10.1016/j.cam.2024.115956.

Yu, C., Zhang, Z., Liu, X., Yang, C., Wang, Y., Liang, X., 2025. Verification method for formation control of USVs based on virtual-real integration. Ocean Eng. 335, 121677. http://dx.doi.org/10.1016/j.oceaneng.2025.121677.

Yuan, Y.-x., 2015. Recent advances in trust region algorithms. Math. Program. 151, http://dx.doi.org/10.1007/s10107-015-0893-2.

Zhang, J., Yu, S., Yan, Y., Zhao, Y., 2024. Fixed-time sliding mode trajectory tracking control for marine surface vessels with input saturation and prescribed performance constraints. Nonlinear Dynam. 112, http://dx.doi.org/10.1007/s11071-024-09918-9.

Zhong, W., Li, H., Meng, Y., Yang, X., Feng, Y., Ye, H., Liu, W., 2022. USV path following controller based on DDPG with composite state-space and dynamic reward function. Ocean Eng. 266, 112449. http://dx.doi.org/10.1016/j.oceaneng.2022.112449.

Zhong, W., Tan, Q., Yan, S., Meng, Y., Ye, N., 2025. USVs cooperative hunting method based on advantage decomposition and sequential decision. Ocean Eng. 318, 120123. http://dx.doi.org/10.1016/j.oceaneng.2024.120123.

Zhong, W., Yan, S., Tan, Q., Ye, N., 2026. USVs formation generation method using HG-PPO with dynamic virtual structures. Ocean Eng. 349, 124091. http://dx.doi.org/10.1016/j.oceaneng.2025.124091.