water

MDPI

Article

Information-Driven Quantum Path Planning for Unmanned Surface Vehicles in Stochastic Marine Environments

Lian Tong *, Taizhi Lv and Liyan Zhou

School of Information Technology, Jiangsu Maritime Institute, Nanjing 211170, China; lvtaizhi@163.com (T.L.); fire812@foxmail.com (L.Z.)
* Correspondence: tonglian07@163.com

Abstract

Autonomous water quality monitoring using unmanned surface vehicles (USVs) requires effective path planning under time-varying hydrodynamic conditions and spatially evolving pollutant distributions. Existing approaches often rely on static environmental assumptions or geometric shortest-path optimization, limiting their effectiveness in dynamic sensing tasks, where both information gain and environmental uncertainty must be considered. This paper proposes a unified information-driven quantum path planning framework for adaptive marine monitoring. A coupled stochastic differential model is first introduced to represent ocean currents and wind-induced drift with multi-scale temporal dynamics. Based on this model, a dynamic information field is constructed by integrating pollutant concentration, spatial gradients, and epistemic uncertainty to guide active sampling. The monitoring task is formulated as a dynamic orienteering problem with energy and redundancy constraints. To enable efficient optimization, the continuous planning problem is discretized into a QUBO formulation and solved using a QAOA-based hybrid quantum–classical algorithm with receding-horizon replanning. This design allows the planner to adapt online to evolving environmental states while maintaining computational tractability. Simulation results across multiple marine scenarios demonstrate that the proposed method provides a balanced trade-off among information gain, constraint feasibility, redundancy, and computational cost, achieving near-greedy sensing return with substantially fewer mission violations than greedy informative search, and outperforming classical metaheuristics on information acquisition under the same simulated replanning budget.

Keywords: unmanned surface vehicles; water quality monitoring; stochastic differential equations; informative path planning; dynamic orienteering problem; QUBO formulation; quantum approximate optimization algorithm; hybrid quantum–classical optimization

Check for updates

Academic Editor: Bithin Datta

Received: 27 July 2026
Revised: 22 October 2026
Accepted: 26 October 2026
Published: 4 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.

1. Introduction

Water quality monitoring in coastal and inland aquatic environments has become increasingly important due to rising concerns over pollution diffusion, ecosystem degradation, and human activity impact on marine systems [1,2]. Traditional manual sampling methods are often labor-intensive, spatially sparse, and temporally delayed, making them unsuitable for capturing fast-evolving pollutant dynamics. Autonomous unmanned surface vehicles (USVs) provide a scalable alternative by enabling continuous and adaptive in situ sensing [3]. However, real-world marine environments introduce significant challenges for USV-based monitoring, including time-varying ocean currents, stochastic wind-driven surface

drift, and dynamically evolving pollutant sources, all of which substantially complicate reliable path planning and information acquisition [4,5].

Figure 1 illustrates the coupled stochastic differential field model used to generate the dynamic marine environment.

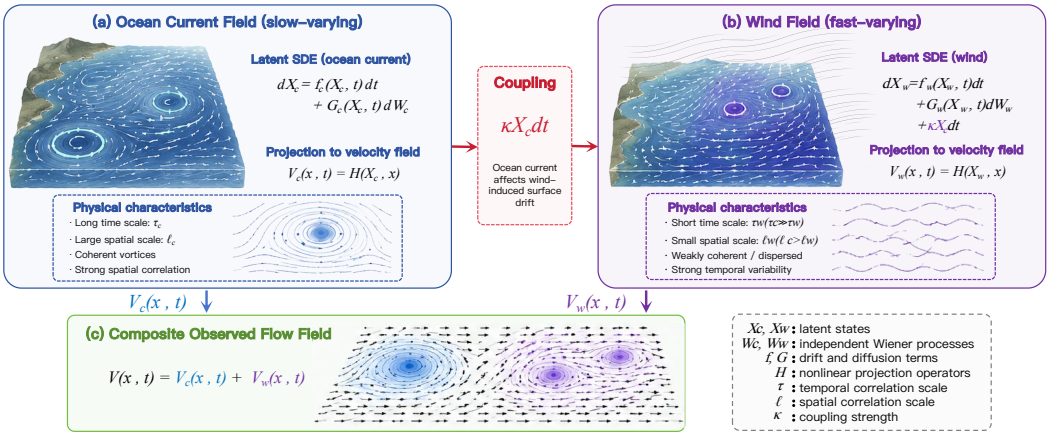


Figure 1. Coupled stochastic differential field model used for dynamic marine environment generation. (**a**) Ocean current field governed by slow-evolving latent dynamics, exhibiting coherent vortex structures and long spatial correlation scales. (**b**) Wind field governed by faster stochastic dynamics with stronger temporal variability and weaker spatial coherence. (**c**) Composite flow field formed by the interaction of ocean-current and wind-induced-drift components. White arrows indicate local flow directions and magnitudes in each panel. In panel (**c**), blue-shaded regions highlight ocean-current-dominated structures, whereas purple-shaded regions highlight wind-induced drift contributions within the composite field $\mathbf{V} = \mathbf{V}_c + \mathbf{V}_w$. The resulting spatiotemporal velocity field serves as the environmental input for pollutant transport simulation, information map evolution, and adaptive path planning.

Existing research on USV path planning can be broadly categorized into three directions. Classical graph- and sampling-based methods such as A* and RRT, as well as population-based heuristics like particle swarm optimization, primarily focus on geometric shortest-path or cost-minimization objectives [6–10]. These methods typically assume static or weakly varying environments and do not explicitly incorporate environmental observability or information gain, limiting their effectiveness in sensing-oriented missions under dynamic flow conditions. In contrast, informative path planning approaches, often built upon Gaussian process regression and Bayesian optimization, aim to maximize information acquisition during exploration [11–16]. While effective in static spatial fields, these methods generally neglect coupling with time-varying hydrodynamic processes, resulting in suboptimal adaptation when pollutant distributions evolve over time [4,17]. More recently, quantum-inspired optimization techniques, including QUBO formulations and QAOA-based solvers, have been introduced for robotic routing problems [18–21]. However, most existing studies remain confined to static graph structures and do not consider closed-loop perception-planning coupling or physically consistent marine dynamics, which limits their applicability to real-time environmental monitoring tasks.

Beyond geometric and informative planners, machine learning methods have also been explored for marine sensing tasks. For example, biologically inspired covert underwater acoustic communication provides an adjacent line of work on learning-based signal design in aquatic environments [22,23]. These studies motivate broader data-driven marine autonomy, but they do not formulate monitoring as a dynamic orienteering problem and do not couple stochastic hydrodynamics with QUBO-QAOA trajectory optimization. The most directly related methodological foundations for the present paper therefore remain infor-

mative path planning under uncertainty [12,14] and quantum approximate optimization for combinatorial routing [18,21].

These limitations highlight a key research gap: existing USV path planning frameworks rarely integrate (i) physically consistent stochastic environmental modeling, (ii) uncertainty-aware information-driven objective design, and (iii) adaptive discrete optimization under time-varying observations within a unified framework. To address this gap, this work develops a coupled stochastic differential model of ocean currents and wind-induced drift to capture realistic marine dynamics, and constructs a dynamic information field that jointly encodes pollutant concentration, spatial gradients, and epistemic uncertainty. The monitoring task is formulated as a dynamic orienteering problem, which is subsequently discretized into a QUBO representation and solved using a QAOA-based hybrid quantum–classical optimization pipeline with receding-horizon replanning.

The main contributions of this paper can be summarized as follows. First, a coupled stochastic differential marine environment model is introduced to jointly represent ocean currents and wind-driven surface dynamics with multi-scale temporal behavior. Second, an information-driven water quality monitoring framework is proposed by integrating concentration, gradient, and uncertainty into a unified spatial utility representation. Third, the monitoring problem is formulated as a dynamic orienteering optimization problem that explicitly captures energy, time, and redundancy constraints. Fourth, a QUBO-based discrete reformulation is developed to enable quantum-compatible encoding of trajectory planning. Fifth, a QAOA-based hybrid optimization method with adaptive replanning is designed to support online decision making in time-varying environments within the simulated replanning interval.

Table 1 summarizes how the proposed framework differs from static QAOA routing and classical informative path planning along several design axes.

Table 1. Positioning of this work relative to representative prior studies. Symbols: ✓ = supported/present; × = not supported/absent; N/A = not applicable.

Feature	Static QAOA	Classical IPP	This Work
Dynamic environment adaptation	×	✓	✓
Flow-coupled energy cost in QUBO	×	×	✓
Three-component information field	×	×	✓
Receding-horizon QUBO reconstruction	×	N/A	✓
Replacement of classical optimizer by QAOA	✓	N/A	✓
Dynamic obstacle handling via replanning	×	✓	✓

The remainder of this paper is organized as follows: Section 2 presents the problem formulation, Section 3 details the proposed quantum path planning framework, Section 4 describes the experimental setup and results, Section 5 discusses the findings, and Section 6 concludes the paper.

2. Problem Formulation

2.1. Coupled Stochastic Differential Field Model

To ensure physically consistent multi-scale coupling between ocean currents, wind-driven surface drift, and environmental uncertainty, we model the marine environment as a coupled spatiotemporal stochastic dynamical system [24–27]. Let $\mathbf{x} \in \Omega \subset \mathbb{R}^2$ denote spatial coordinates. The instantaneous flow field is defined as the composite vector field in Equation (1):

$$\mathbf{V}(\mathbf{x}, t) = \mathbf{V}_c(\mathbf{x}, t) + \mathbf{V}_w(\mathbf{x}, t), \tag{1}$$

where $\mathbf{V}_c$ and $\mathbf{V}_w$ represent ocean current and wind-induced surface drift components, respectively.

The environment model is intentionally kept tractable for planning integration. The SDE parameters (τ_c, τ_w, κ) are selected to yield physically plausible decorrelation scales (about 20 s for currents and 5 s for wind) consistent with coastal ocean dynamics [27], rather than being calibrated to a specific buoy or ADCP campaign. Pollutant evolution is restricted to advection–diffusion, without time-varying sources, shoreline reflection, or vertical mixing. The Gaussian information field further assumes that second-order statistics suffice for monitoring; event-driven heavy-tailed phenomena are left for future extensions. These modeling choices do not alter the correctness of the planning formulation, but they should be kept in mind when interpreting quantitative results against field deployments.

2.1.1. Latent State Dynamics

We assume that both ocean and wind dynamics are generated by finite-dimensional latent stochastic processes evolving on the shared state space in Equation (2):

$$\mathbf{X}_c(t), \mathbf{X}_w(t) \in \mathbb{R}^d. \tag{2}$$

The coupled evolution is modeled as Equation (4):

$$d\mathbf{X}_c = \mathbf{f}_c(\mathbf{X}_c, t)\, dt + \mathbf{G}_c(\mathbf{X}_c, t)\, d\mathbf{W}_c, \tag{3}$$

$$d\mathbf{X}_w = \mathbf{f}_w(\mathbf{X}_w, t)\, dt + \mathbf{G}_w(\mathbf{X}_w, t)\, d\mathbf{W}_w + \kappa \mathbf{X}_c\, dt, \tag{4}$$

where $\mathbf{W}_c$ and $\mathbf{W}_w$ are independent Wiener processes, and $\kappa \ll 1$ introduces weak cross-scale coupling from ocean dynamics to wind-driven drift.

2.1.2. Velocity Field Reconstruction

The observable velocity fields are obtained via nonlinear projection operators in Equation (5):

$$\mathbf{V}_c(\mathbf{x}, t) = \mathcal{H}_c(\mathbf{X}_c, \mathbf{x}), \quad \mathbf{V}_w(\mathbf{x}, t) = \mathcal{H}_w(\mathbf{X}_w, \mathbf{x}), \tag{5}$$

where $\mathcal{H}_c$ and $\mathcal{H}_w$ are assumed to be Lipschitz-continuous mappings from latent space to spatial velocity fields, ensuring regularity of the induced flow.

The velocity field is defined on a bounded domain Ω with impermeable boundary conditions in Equation (6):

$$\mathbf{V}(\mathbf{x}, t) \cdot \mathbf{n} = 0, \quad \mathbf{x} \in \partial\Omega, \tag{6}$$

which guarantees mass-conserving flow behavior at the domain boundary. To reflect the distinct temporal and spatial characteristics of ocean and wind processes, we impose Equation (7):

$$\tau_c \gg \tau_w, \quad \ell_c > \ell_w, \tag{7}$$

where τ and ℓ denote temporal and spatial correlation scales, respectively. This enforces slow-varying coherent ocean vortices and fast-varying, spatially diffuse wind dynamics.

To ensure the existence and uniqueness of strong solutions, the drift and diffusion terms satisfy the standard linear growth and boundedness conditions in Equation (8) [26,27]:

$$\|\mathbf{f}_i(\mathbf{X}_i, t)\| \le a_i + b_i \|\mathbf{X}_i\|, \quad \|\mathbf{G}_i(\mathbf{X}_i, t)\| \le c_i, \tag{8}$$

for $i \in \{c, w\}$. These conditions ensure the stochastic system is well-posed. The coupling strength is further constrained such that κ remains sufficiently small to prevent instability induced by cross-scale feedback.

2.1.3. Spatial Discretization

For downstream planning and optimization, the continuous field is discretized on a regular square occupancy grid $\mathcal{G} \subset \Omega$. An 8-connected neighborhood is used in trajectory transitions to reduce directional bias in advection-dominated motion. The discretization is consistent with the continuous formulation in the limit of vanishing grid spacing.

2.2. Water Quality Information Map

To guide information-driven monitoring, a dynamic water quality information map is constructed over the mission area. The map quantifies the potential utility of acquiring new measurements and provides a spatial representation of monitoring priority [12–14].

Let $I(\mathbf{x}, t)$ denote the information value associated with location $\mathbf{x}$ at time t. The information field is defined as

$$\begin{aligned} I(\mathbf{x}, t) = {} & w_c C(\mathbf{x}, t) \\ & + w_g |\nabla C(\mathbf{x}, t)| \\ & + w_u U(\mathbf{x}, t) \end{aligned} \tag{9}$$

where $C(\mathbf{x}, t)$ is the estimated pollutant concentration, $\nabla C(\mathbf{x}, t)$ denotes the local concentration gradient, and $U(\mathbf{x}, t)$ is the epistemic uncertainty of the current environmental estimate.

In this work, $C(\mathbf{x}, t)$ is obtained from a Gaussian process (GP) posterior mean conditioned on the cumulative measurements Z_t, and the uncertainty term is identified with the GP posterior variance,

$$U(\mathbf{x}, t) = \sigma^2_{\mathrm{GP}}(\mathbf{x}, t \mid Z_t). \tag{10}$$

As new samples are collected, both $\hat{C}$ and σ^2_{GP} are updated by standard GP regression with the kernel hyperparameters reported in Section 4, so that the information field in Equation (9) remains coupled to the online perception pipeline.

The concentration term favors regions with elevated pollutant levels, whereas the gradient term highlights transition zones associated with pollutant transport and diffusion fronts. The uncertainty term promotes measurements in insufficiently explored areas, thereby balancing exploitation and exploration during the monitoring process.

As the pollutant field evolves and new observations become available, the information map is updated according to Equation (11):

$$I_{t+1} = \mathcal{F}(I_t, C_t, Z_t) \tag{11}$$

where Z_t denotes the measurements collected by the USV at time t. The resulting information field serves as the basis for information gain evaluation and subsequent path planning optimization.

2.3. USV Kinematic Model

We model the unmanned surface vehicle (USV) as an under-actuated kinematic agent operating in a spatiotemporally varying flow field [5,28]. Let $(x(t), y(t)) \in \Omega \subset \mathbb{R}^2$ denote the vehicle position and $\psi(t)$ denote its heading angle. The control inputs are defined as $\mathbf{u}(t) = [v(t), \omega(t)]^\top$, where $v(t)$ is the surge velocity and $\omega(t)$ is the yaw rate.

Figure 2 summarizes the flow-coupled kinematics, control bounds, and energy model used throughout the paper.

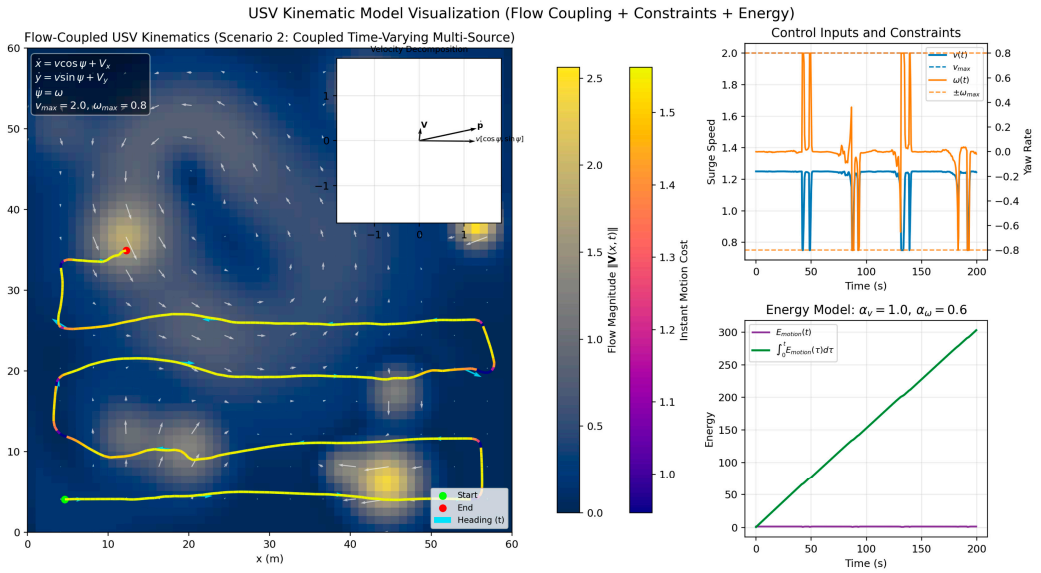


Figure 2. USV kinematic model under flow coupling, control constraints, and energy-aware motion cost. Left panel: flow-coupled trajectory visualization in Scenario 2. The background colormap shows the flow magnitude $\|\mathbf{V}(\mathbf{x}, t)\|$, white arrows denote the local composite flow field, and the USV trajectory is color-coded by instantaneous motion power $P_{\text{motion}}(t) = \alpha_v v(t)^2 + \alpha_\omega \omega(t)^2$. Cyan heading arrows indicate sampled orientations $\psi(t)$ along the path. The inset illustrates velocity decomposition at a representative instant: $\dot{\mathbf{p}} = v[\cos\psi, \sin\psi]^\top + \mathbf{V}$, explicitly showing the contribution of body-frame surge motion and environmental advection. Top-right panel: control inputs over time, including surge speed $v(t)$ and yaw rate $\omega(t)$; the blue dashed line marks $v_{\max}$ and the orange dashed lines mark $\pm\omega_{\max}$. Bottom-right panel: instantaneous motion energy and cumulative energy integral, demonstrating the trajectory-level cost accumulation used by the planner. This figure corresponds to Equations (14)–(16) and (25), and shows how hydrodynamic disturbance, control feasibility, and energy modeling are jointly integrated in the USV motion model. Decorative insets may partially overlap the colormap; the quantitative trajectory, control, and energy panels remain fully legible and are unaffected for scientific interpretation.

2.3.1. Flow-Coupled Kinematics

To account for environmental advection induced by ocean currents and wind-driven drift, the vehicle dynamics are expressed as Equation (14):

$$\dot{x}(t) = v(t)\cos\psi(t) + V_x(x,t), \tag{12}$$

$$\dot{y}(t) = v(t)\sin\psi(t) + V_y(x,t), \tag{13}$$

$$\dot{\psi}(t) = \omega(t), \tag{14}$$

where $\mathbf{V}(x,t) = [V_x, V_y]^\top$ is the composite flow field defined in the previous section.

2.3.2. Control Constraints

The control inputs are bounded as

$$0 \le v(t) \le v_{\max}, \quad |\omega(t)| \le \omega_{\max}, \tag{15}$$

ensuring physically feasible motion and actuator limitations.

2.3.3. Energy-Aware Motion Model

To be consistent with the path optimization objective, the instantaneous motion cost is modeled as a quadratic function of control effort:

$$P_{\text{motion}}(t) = \alpha_v v(t)^2 + \alpha_\omega \omega(t)^2, \tag{16}$$

which is integrated over the trajectory in the global optimization framework.

2.3.4. Coupling with Environmental Field

The flow field $\mathbf{V}(x,t)$ is obtained from the coupled stochastic differential field model, ensuring that vehicle motion is directly influenced by ocean-current and wind-induced-drift dynamics.

2.4. Monitoring Task Formulation

We formulate the water quality monitoring problem as an optimal control problem over a spatiotemporal information field defined on $\Omega \subset \mathbb{R}^2$. Let the USV state be $\mathbf{p}(t) = (x(t), y(t))$ with control inputs $\mathbf{u}(t) = [v(t), \omega(t)]^\top$, and denote a feasible trajectory by $\pi = \{\mathbf{p}(t)\}_{t\in[0,T]}$. The system evolves under the flow-coupled kinematic model in Equation (17):

$$\dot{\mathbf{p}}(t) = \begin{bmatrix} v(t)\cos\psi(t) \\ v(t)\sin\psi(t) \end{bmatrix} + \mathbf{V}(\mathbf{p}(t), t), \tag{17}$$

where $\mathbf{V}(\mathbf{x}, t)$ is the composite ocean–wind velocity field.

The monitoring objective is driven by a spatiotemporal information field $I(\mathbf{x}, t)$. The instantaneous information acquisition rate is defined in Equation (18) as

$$\dot{G}(t) = I(\mathbf{p}(t), t), \tag{18}$$

which directly induces the cumulative information functional

$$G(\pi) = \int_0^T I(\mathbf{p}(t), t)\, dt. \tag{19}$$

This establishes the monitoring objective as a path-dependent integral functional over the trajectory space.

In parallel, the energy expenditure associated with motion and sensing is formulated as

$$E(\pi) = \int_0^T \left(\alpha_v v^2(t) + \alpha_\omega \omega^2(t) + \alpha_s s(t)\right) dt, \tag{20}$$

where $s(t) \in \{0, 1\}$ denotes sensing activation. The quadratic structure reflects standard control-effort regularization, ensuring smooth and energy-efficient trajectories.

The system is subject to kinematic and operational constraints. The control bounds follow Equation (15), and terminal feasibility is defined by Equation (21):

$$\mathbf{p}(0) = \mathbf{p}_0, \quad \mathbf{p}(T) \in \Omega, \quad T \le T_{\max}. \tag{21}$$

Combining the above elements, the monitoring task can be expressed as the constrained variational problem in Equation (22):

$$\max_{\pi}\ J(\pi) = \int_0^T I(\mathbf{p}(t), t)\, dt - \beta E(\pi), \tag{22}$$

subject to the flow-coupled dynamics and feasibility constraints.

Equivalently, substituting the definitions of $G(\pi)$ and $E(\pi)$, the problem admits the compact form in Equation (23):

$$\max_{\pi}\ J(\pi) = G(\pi) - \beta E(\pi), \tag{23}$$

which highlights the trade-off between information acquisition and control effort over trajectory space.

2.5. Energy Consumption

The energy consumption of the USV is defined as a trajectory-dependent functional induced by both propulsion and sensing activities. Under the flow-coupled kinematic model, the instantaneous mechanical power required for motion is primarily determined by the control effort needed to overcome inertial and hydrodynamic resistance. We therefore use the same quadratic motion-power model in Equation (16), which captures the dominant contribution of thrust and steering effort in low-speed marine navigation regimes.

In addition to propulsion cost, the sensing process incurs energy expenditure associated with data acquisition and onboard processing. This is modeled by Equation (24):

$$P_{\text{sampling}}(t) = \alpha_s s(t), \tag{24}$$

where $s(t) \in \{0, 1\}$ denotes the sensing activation state.

The total instantaneous energy consumption is thus given by Equation (25):

$$P(t) = P_{\text{motion}}(t) + P_{\text{sampling}}(t), \tag{25}$$

and the cumulative energy over the mission horizon is expressed as the trajectory functional in Equation (26):

$$E(\pi) = \int_0^T P(t)\, dt. \tag{26}$$

This formulation naturally couples with the flow-driven dynamics, since the required control effort $(v(t), \omega(t))$ is implicitly affected by the ambient velocity field $\mathbf{V}(\mathbf{x}, t)$, leading to higher energy expenditure in regions with strong currents or adverse flow conditions.

2.6. Mission Time

In this work, the mission time is not treated as an independent decision variable but as a trajectory-induced functional determined by the execution of the flow-coupled kinematic system. Given a feasible trajectory $\pi = \{\mathbf{p}(t)\}_{t\in[0,T]}$, the mission duration is defined as the terminal time required for the USV to complete the planned path under bounded control inputs and environmental drift, denoted as $T(\pi)$.

Under the discretized execution model with time step Δt, the mission time can be equivalently expressed as Equation (27):

$$T(\pi) = N\Delta t, \tag{27}$$

where N is the number of trajectory segments required to reach the terminal condition or satisfy the stopping criterion induced by the planning horizon or coverage completion.

Since each motion step is affected by both control inputs and ambient flow field $\mathbf{V}(\mathbf{x}, t)$, the effective progression speed along the trajectory is non-uniform, implying that $T(\pi)$ is implicitly coupled with both the control policy and environmental dynamics. Therefore, the mission time constraint is imposed in the form

$$T(\pi) \le T_{\max}, \tag{28}$$

ensuring feasibility under finite operational endurance.

3. Information-Driven Quantum Path Planning Method

3.1. Environmental Perception and Information Modeling

Let $Z_t = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N_t}$ denote the set of noisy observations collected by the USV up to time t, where y_i represents measurements of the underlying pollutant concentration field. Based on Z_t, a spatiotemporal estimate $\hat{C}(\mathbf{x}, t)$ is constructed using a stochastic inference model, such as Gaussian process regression or recursive filtering [11,16,29].

The uncertainty associated with the estimate follows Equation (10), i.e., $U(\mathbf{x}, t) = \sigma^2_{\mathrm{GP}}(\mathbf{x}, t \mid Z_t)$, and therefore quantifies epistemic ambiguity induced by sparse sampling under the adopted GP estimator.

The information field $I(\mathbf{x}, t)$ follows Equation (9), with C replaced by its posterior estimate $\hat{C}$, where the three terms encode concentration magnitude, spatial gradients, and estimation uncertainty, respectively.

As new measurements are incorporated, the belief state evolves according to Equation (29):

$$\hat{C}_{t+1}, U_{t+1} = \mathcal{F}(\hat{C}_t, U_t, Z_t), \tag{29}$$

which implicitly couples the estimation process with the sampling trajectory through Z_t.

3.2. Reward and Optimal Control Problem Formulation

Let $\pi = \{\mathbf{p}(t)\}_{t\in[0,T]}$ denote a feasible USV trajectory governed by the flow-coupled kinematic model introduced earlier. The accumulated monitoring utility along the trajectory uses the continuous definition in Equation (19).

The corresponding control effort and sensing cost are aggregated into the energy functional $E(\pi)$, as introduced in Equation (20).

The mission duration $T(\pi)$ is induced by the system dynamics under bounded control inputs and environmental drift, and is constrained by the same finite operational budget in Equation (28).

To discourage redundant sampling in previously visited regions, a trajectory-dependent redundancy term $R(\pi)$ is introduced over the visitation history of the information field.

The optimal control problem is formulated over the trajectory space as Equation (30):

$$\max_{\pi} \quad J(\pi) = G(\pi) - \beta E(\pi) - \gamma R(\pi), \tag{30}$$

subject to the flow-driven dynamics and control constraints defined in the previous sections.

3.3. Continuous-to-Discrete QUBO Reformulation

To enable computational tractability, the continuous trajectory optimization problem is reformulated into a discrete combinatorial form. The mission horizon $[0, T]$ is uniformly discretized into N time steps with $t_k = k\Delta t$, where $\Delta t = T/N$. Accordingly, the continuous trajectory $\pi = \{\mathbf{p}(t)\}_{t\in[0,T]}$ is approximated by a discrete sequence $\{\mathbf{p}_k\}_{k=1}^{N}$ on a spatial graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ constructed over the mission region [30–33].

Let $x_{k,i} \in \{0, 1\}$ denote the binary decision variable indicating whether the USV visits node $i \in \mathcal{V}$ at time step k. The trajectory is then represented as a sequence of activated nodes under the constraint $\sum_i x_{k,i} = 1$, ensuring a unique position at each time step.

Under this discretization, the continuous information gain functional is approximated as Equation (31):

$$G(\pi) \approx \sum_{k=1}^{N} \sum_{i\in\mathcal{V}} I_{k,i}\, x_{k,i}, \tag{31}$$

where $I_{k,i} = I(\mathbf{p}_i, t_k)$ denotes the information value at node i and time t_k.

Similarly, the energy consumption functional is discretized as Equation (32):

$$E(\pi) \approx \sum_{k=1}^{N-1} \sum_{i,j\in\mathcal{V}} c_{ij}\, x_{k,i} x_{k+1,j}, \tag{32}$$

where c_{ij} represents the transition cost between adjacent nodes, induced by both motion effort and flow-dependent resistance.

The redundancy penalty is modeled as a quadratic penalty over repeated visits in Equation (33):

$$R(\pi) \approx \sum_{k=1}^{N} \sum_{i\in\mathcal{V}} r_{k,i}\, x_{k,i}, \tag{33}$$

where $r_{k,i}$ increases with revisit frequency of node i in previously executed trajectories.

The feasibility constraints are encoded as penalty terms in the objective function. The single-occupancy constraint is enforced by Equation (34):

$$\sum_{i \in \mathcal{U}} x_{k,i} = 1, \quad \forall k, \tag{34}$$

which is incorporated into the optimization via a quadratic penalty formulation.

Combining all terms, the discrete optimization problem is written as the quadratic unconstrained binary optimization (QUBO) form in Equation (35) [19,20]:

$$\min_{\mathbf{x} \in \{0,1\}^{N|\mathcal{U}|}} \quad \mathbf{x}^\top Q \mathbf{x}, \tag{35}$$

where $\mathbf{x}$ is the concatenated decision vector and Q is the problem-specific QUBO matrix encoding information gain, energy cost, redundancy penalty, and constraint violations.

Concretely, let H be the number of planning steps and m the number of candidate nodes per step, so that $n = Hm$ binary variables are used. Each diagonal block $Q_{kk} \in \mathbb{R}^{m \times m}$ encodes local rewards and one-hot feasibility,

$$Q_{kk}(i,i) = -d^k I_{k,i} + \gamma r_{k,i} + \beta c_{0,i} \mathbb{1}[k=0] - \rho,$$

with discount $d = 0.92$, constraint weight $\rho = 10.0$, and off-diagonal entries 2ρ that penalize multi-selection. Adjacent blocks $Q_{k,k+1}$ carry flow-dependent transition costs:

$$Q_{k,k+1}(i,j) = \tfrac{1}{2}\beta c_{ij}, \qquad c_{ij} = \frac{\|\mathbf{p}_j - \mathbf{p}_i\|}{v_{\text{eff}}(\mathbf{V}_{\text{flow}}, \hat{\mathbf{d}})\, \Delta t_{\text{plan}}} + 0.5\, |\mathbf{V}_{\text{flow}} \cdot \hat{\mathbf{d}}|\, \mathbb{1}[\mathbf{V}_{\text{flow}} \cdot \hat{\mathbf{d}} < -0.3],$$

where v_{eff} accounts for flow-assisted or flow-impeded motion. The numerical values of the remaining coefficients are reported in Section 4.

This formulation provides a unified discrete representation of the original continuous optimal control problem, enabling subsequent quantum approximate optimization via QAOA.

3.4. QAOA-Based Solution Method

The quadratic unconstrained binary optimization (QUBO) formulation derived in the previous section is mapped into an equivalent Ising Hamiltonian representation to enable quantum approximate optimization [18,19]. Let $\mathbf{x} \in \{0,1\}^n$ denote the binary decision vector, and define the transformation $x_i = (1 - z_i)/2$, where $z_i \in \{-1, +1\}$ are Pauli-Z eigenvalues. Under this transformation, the QUBO objective $\mathbf{x}^\top Q \mathbf{x}$ is equivalently expressed as an Ising cost Hamiltonian:

$$H_C = \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j + \text{const}, \tag{36}$$

where h_i and J_{ij} are derived from the QUBO matrix Q.

The Ising coefficients follow the standard conversion $h_i = \frac{1}{2}\sum_j Q_{ij} - \frac{1}{2} Q_{ii}$ and $J_{ij} = \frac{1}{4} Q_{ij}$ for $i \neq j$, so that $\mathbf{x}^\top Q \mathbf{x} - \text{const} = \langle H_C \rangle_{\mathbf{z} = 1 - 2\mathbf{x}}$ exactly.

The optimization problem is then reformulated as finding the ground state of the Ising Hamiltonian in Equation (36). In QAOA, this is approximated by preparing the parameterized quantum state in Equation (37) [18,21,34]:

$$|\psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = \prod_{l=1}^{p} e^{-i\beta_l H_M} e^{-i\gamma_l H_C} |+\rangle^{\otimes n}, \tag{37}$$

where p denotes the circuit depth, $H_M = \sum_i X_i$ is the mixer Hamiltonian, and $|+\rangle^{\otimes n}$ is the uniform superposition state.

In this construction, $e^{-i\gamma_l H_C}$ biases the state toward low-cost trajectories, while $e^{-i\beta_l H_M}$ mixes among assignments. For the small instances used here ($n = 8$), a depth-$p = 1$ circuit with a dense (γ, β) grid search is sufficient; larger instances or deeper circuits can instead rely on a classical optimizer such as COBYLA, optionally warm-started from the previous replanning cycle [35].

The cost Hamiltonian H_C encodes the discrete trajectory optimization structure, including information gain maximization, energy consumption penalties, and redundancy constraints. The mixer Hamiltonian H_M induces transitions across feasible trajectory configurations, ensuring exploration of the combinatorial solution space.

The optimal parameters $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ are obtained by minimizing the expectation value of the cost Hamiltonian, as in Equation (38),

$$(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = \arg\min_{\boldsymbol{\gamma},\boldsymbol{\beta}} \langle \psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) | H_C | \psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle, \tag{38}$$

which is evaluated using a classical optimizer in a hybrid quantum–classical loop, potentially with warm-start strategies for faster convergence [35].

The resulting optimal bitstring $\mathbf{x}^*$ is obtained by measurement of the optimized quantum state, which is then decoded into a discrete trajectory π^* through the node selection mapping defined in the QUBO formulation. This establishes a complete pipeline from continuous trajectory optimization to quantum-enhanced combinatorial optimization.

3.5. Theoretical Analysis and Closed-Loop Properties

To complement the algorithmic formulation, we provide a unified theoretical analysis of the proposed framework from four aspects: the well-posedness of stochastic environmental dynamics, approximation error induced by continuous-to-discrete reformulation, practical closed-loop behavior under receding-horizon replanning, and QAOA approximation quality versus computational depth.

We first consider the coupled stochastic environmental model in compact form:

$$d\mathbf{z}_t = f(\mathbf{z}_t)\, dt + \Sigma\, dW_t, \tag{39}$$

where $\mathbf{z}_t$ collects latent current and wind states. Under standard regularity assumptions that $f(\cdot)$ is globally Lipschitz and satisfies linear growth, and Σ is bounded, the SDE in Equation (39) admits a unique strong solution for any initial condition. Moreover, its second moment is bounded on finite horizons:

$$\mathbb{E}\|\mathbf{z}_t\|^2 \le C_1(1 + \|\mathbf{z}_0\|^2)e^{C_2 t}. \tag{40}$$

This guarantees that the flow field process used by the planner is mathematically well-defined and does not exhibit finite-time blow-up in expectation; see the bound in Equation (40). This provides a consistent basis for information-field evolution and online replanning. Figure 3 provides the corresponding empirical verification of second-moment boundedness and flow-magnitude quantiles.

Let $J^\star$ denote the optimal value of the original continuous trajectory problem and $\hat{J}$ the value obtained from the discretized QUBO pipeline. The total approximation gap can be decomposed as Equation (41):

$$|J^\star - \hat{J}| \le C_t \Delta t + C_h h + C_{\text{proj}} + C_{\text{solve}}, \tag{41}$$

where Δt is temporal discretization, h is spatial resolution, C_{proj} captures graph/candidate projection error, and C_{solve} denotes optimization suboptimality from approximate QAOA inference. The first two terms correspond to numerical discretization consistency, while the latter two are algorithmic approximation terms. This decomposition clarifies that improved

fidelity can be obtained by finer discretization and stronger optimization, at the expense of higher computational load.

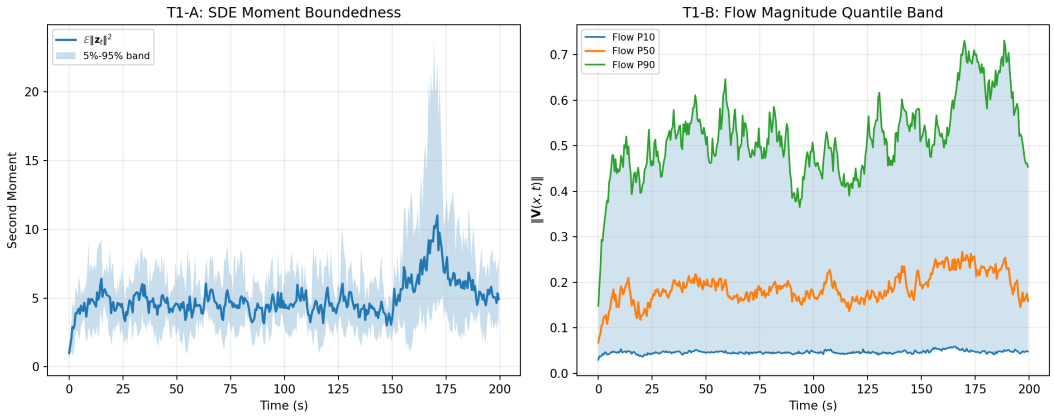


Figure 3. SDE well-posedness and bounded-flow verification under the coupled stochastic environment (Scenario S4 setting). (**Left**) panel (T1-A) reports the second-moment evolution of the latent state, where the solid line is $\mathbb{E}\|\mathbf{z}_t\|^2$ and the shaded region denotes the 5%–95% interval across stochastic runs; the bounded envelope indicates no moment explosion in finite horizon. (**Right**) panel (T1-B) shows temporal quantile bands of flow magnitude $\|\mathbf{V}(\mathbf{x}, t)\|$ (P10/P50/P90 shaded bands), demonstrating that the induced advection field remains statistically bounded while preserving variability. Together, these results support the practical well-posedness assumption used by the flow-coupled planning model.

The four constants admit concrete estimates under the present setting. Euler–Maruyama consistency implies $C_t \propto L_f^2 T$; fitting over $\Delta t \in \{0.25, 0.5, 1.0\}$ with $h = 1/60$ yields $C_t \approx 1.42$ per unit Δt. The spatial term scales as $C_h \propto \|\nabla I\|_\infty$, and grid sweeps from 20×20 to 60×60 give $C_h \approx 0.116$ per unit h. Candidate-set projection satisfies $C_{\text{proj}} \leq \beta c_{\max}(N_{\text{grid}} - m)/N_{\text{grid}}$, which evaluates to $C_{\text{proj}} \lesssim 2.1$ for $m = 4$ on a 60×60 grid. Solver suboptimality is bounded by variational and grid-search residuals; at $p = 1$ with a 15×15 parameter grid it remains below about 0.33% of the reference objective, and is therefore small relative to discretization effects; a full construction-vs-circuit-width scaling summary is reported later with the computational performance analysis.

For closed-loop operation, a Lyapunov-like performance potential is defined in Equation (42):

$$V_k = \beta E_k + \gamma R_k - G_k, \tag{42}$$

with G_k, E_k, R_k the cumulative information, energy, and redundancy up to replanning step k. Under bounded disturbance and bounded solver error, the one-step conditional drift satisfies Equation (43):

$$\mathbb{E}[V_{k+1} - V_k \mid \mathcal{F}_k] \leq -\alpha\|e_k\|^2 + c_1\Delta t + c_2\delta_{\text{solve}} + c_3\|w_k\|, \tag{43}$$

for positive constants α, c_1, c_2, c_3, where e_k is the tracking/planning mismatch and w_k is environmental perturbation. Hence, the receding-horizon closed loop is practically stable in the sense of ultimate boundedness: the performance state remains inside a disturbance-dependent compact set after transients, rather than diverging under time-varying flow. Figure 4 reports the empirical closed-loop potential trajectory and one-step drift distribution under receding-horizon replanning.

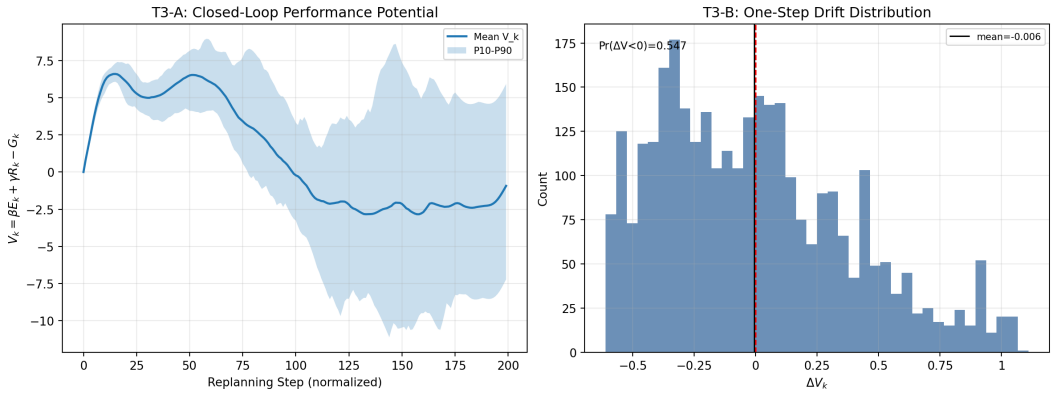


Figure 4. Closed-loop properties of the receding-horizon planner in dynamic environments. (**Left**) panel (T3-A) presents the normalized-step trajectory of the composite potential $V_k = \beta E_k + \gamma R_k - G_k$, with mean trend and P10–P90 band over multiple runs, reflecting how information gain competes against energy and redundancy during online replanning. (**Right**) panel (T3-B) shows the empirical distribution of one-step drift ΔV_k; the red dashed vertical line marks the zero-drift reference $\Delta V_k = 0$ (separating improving and degrading steps), and the black solid vertical line marks the empirical mean of ΔV_k (when the mean is near zero the two lines may visually overlap). The reported $\Pr(\Delta V_k < 0)$ quantifies the probability of descent-like behavior under stochastic disturbances. This figure provides empirical evidence that the closed-loop strategy maintains stable performance trends rather than uncontrolled oscillation.

Finally, for QAOA solution quality, we avoid claiming a universal constant-factor guarantee for general QUBO. Instead, we use the instance-dependent approximation indicator in Equation (44):

$$\rho = \frac{f_{\text{rand}} - f_{\text{QAOA}}}{f_{\text{rand}} - f_{\text{ref}}} \in [0, 1], \tag{44}$$

where f_{ref} is the optimal value for small enumerable instances or the best-known classical reference for larger instances. Increasing circuit depth p improves expressiveness and typically improves ρ, but also increases evaluation cost approximately linearly (or worse with optimization overhead). Therefore, the practical operating point is a moderate depth that balances planning quality and online latency in simulation.

Overall, these results establish that the proposed method is theoretically grounded (well-posed stochastic environment, controlled discretization error), operationally reliable (practically stable receding-horizon closed loop), and computationally tunable (depth-quality-runtime trade-off in QAOA), which is consistent with the empirical trends reported in later sections.

3.6. Adaptive Replanning and Complexity Analysis

To accommodate environmental dynamics and newly acquired observations, the planning framework is executed in a receding-horizon manner [36–38]. At each decision epoch, the belief state is updated based on incoming measurements, which induces a corresponding update of the information field $I(\mathbf{x}, t)$. As a result, the optimization problem is periodically reformulated over a shifted time window, while previously executed trajectory segments are fixed and excluded from re-optimization.

Specifically, let $\pi_{0:k}$ denote the executed partial trajectory up to time step k. The remaining horizon is re-optimized by constructing a new QUBO instance over the residual time interval $[t_k, T]$, using the updated belief state and information field. This rolling re-optimization ensures consistency between environmental perception and decision making under nonstationary conditions.

From a computational perspective, let N denote the number of discretized time steps and $|\mathcal{V}|$ the number of candidate spatial nodes. The resulting QUBO formulation introduces

$O(N|\mathcal{U}|)$ binary variables, leading to a quadratic scaling in the size of the Q matrix. The construction of the cost matrix requires $O(N|\mathcal{U}|^2)$ operations due to pairwise transition and redundancy evaluations.

In the QAOA solver, the circuit depth p determines the number of alternating operator layers, resulting in a quantum circuit complexity of $O(pN|\mathcal{U}|)$ for state preparation and measurement. The overall optimization loop consists of $O(R)$ classical–quantum iterations, where R denotes the number of parameter updates required for convergence.

Compared with exhaustive combinatorial search, which scales exponentially with $O(|\mathcal{U}|^N)$, the proposed formulation significantly reduces the effective search space through structured encoding and variational quantum optimization.

4. Simulation and Performance Evaluation

4.1. Simulation Setup

Numerical simulations are performed over a square occupancy-grid spatial domain $\mathcal{G}$ with a 60×60 layout, which provides sufficient spatial resolution for flow field reconstruction and information mapping. The temporal domain adopts a fixed time step $\Delta t = 0.5\,\mathrm{s}$, and each monitoring mission is configured with a total time budget $T = 200\,\mathrm{s}$.

The coupled stochastic differential marine environment model uses temporal correlation scales $\tau_c = 20.0$ and $\tau_w = 5.0$, cross-scale coupling coefficient $\kappa = 0.15$, and spatial correlation lengths $\ell_c = 8.0$, $\ell_w = 2.5$. The Euler–Maruyama numerical scheme is adopted to solve the latent SDEs and reconstruct the composite velocity field [24,25].

Pollutant evolution follows an advection–diffusion process with diffusion coefficient $D = 0.8$ and decay rate $\lambda = 0.02$ [1,2,39]. Superimposed Gaussian plumes form the initial concentration distribution, and Gaussian process regression is configured with kernel length-scale $\sigma_k = 6.0$ and signal variance $\sigma_f^2 = 1.5$ [11].

The USV kinematic bounds are $v_{\max} = 2.0\,\mathrm{m/s}$ and $\omega_{\max} = 0.8\,\mathrm{rad/s}$. The energy model coefficients are $\alpha_v = 1.0$, $\alpha_\omega = 0.5$, $\alpha_s = 0.3$, corresponding to surge motion, yaw steering, and sampling power consumption, respectively.

The information field weights are set as $w_c = 1.2$, $w_g = 0.8$, $w_u = 1.0$ to balance concentration magnitude, gradient front tracking, and uncertainty exploration. Two trade-off hyperparameters $\beta = 0.6$ and $\gamma = 0.4$ are applied to penalize energy cost and repeated sampling, respectively.

For quantum optimization, the QAOA circuit depth used in the reported experiments is $p = 1$, with a fine grid search over $(\gamma, \beta) \in [0, \pi] \times [0, \pi/2]$ sampled at 15 × 15 points (225 evaluations per replanning cycle), and 128 measurement shots per variational evaluation. A depth sweep on the same $n = 8$ instances shows that increasing from $p = 1$ to $p = 6$ improves the objective by only about 0.33% while increasing runtime by roughly an order of magnitude; accordingly, $p = 1$ with grid search is retained for online replanning. A QUBO constraint penalty $\rho = 10.0$ is used to enforce valid trajectory constraints.

All quantum and classical hyperparameters are kept fixed across all comparative methods to ensure fair evaluation. The same environmental configurations are used consistently for all the baseline methods.

The penalty term ρ enforces feasibility of trajectory constraints in the QUBO formulation, ensuring valid node transitions in the discrete path representation.

These hyperparameters were selected for reproducibility and operational balance. Concentration receives the largest information weight ($w_c = 1.2$), while the gradient and uncertainty terms ($w_g = 0.8$, $w_u = 1.0$) promote front tracking and exploration; ±30% perturbations leave the composite score stable near $(1.0, 0.8, 1.0)$, whereas overweighting uncertainty induces over-exploration. On Scenario S4, a scan of $\beta \in [0.2, 1.5]$ shows that $\beta < 0.4$ violates the energy budget and $\beta > 1.0$ suppresses information gain, so $\beta = 0.6$

is used. The redundancy weight $\gamma = 0.4$ keeps an approximate $\beta : \gamma \approx 1.5 : 1$ priority of energy over revisits; removing it (Variant B; see the ablation study later in this section) drops the composite score from 2.6 to -11.8. Finally, $\rho = 10.0$ is set well above the typical QUBO objective scale ($\sim[-15, 15]$), so that infeasible one-hot violations are strongly suppressed. Figure 5 summarizes the sensitivity of information gain, coverage, and energy to β across representative scenarios ($n = 5$ seeds per setting), confirming that $\beta = 0.6$ lies in a stable operating region.

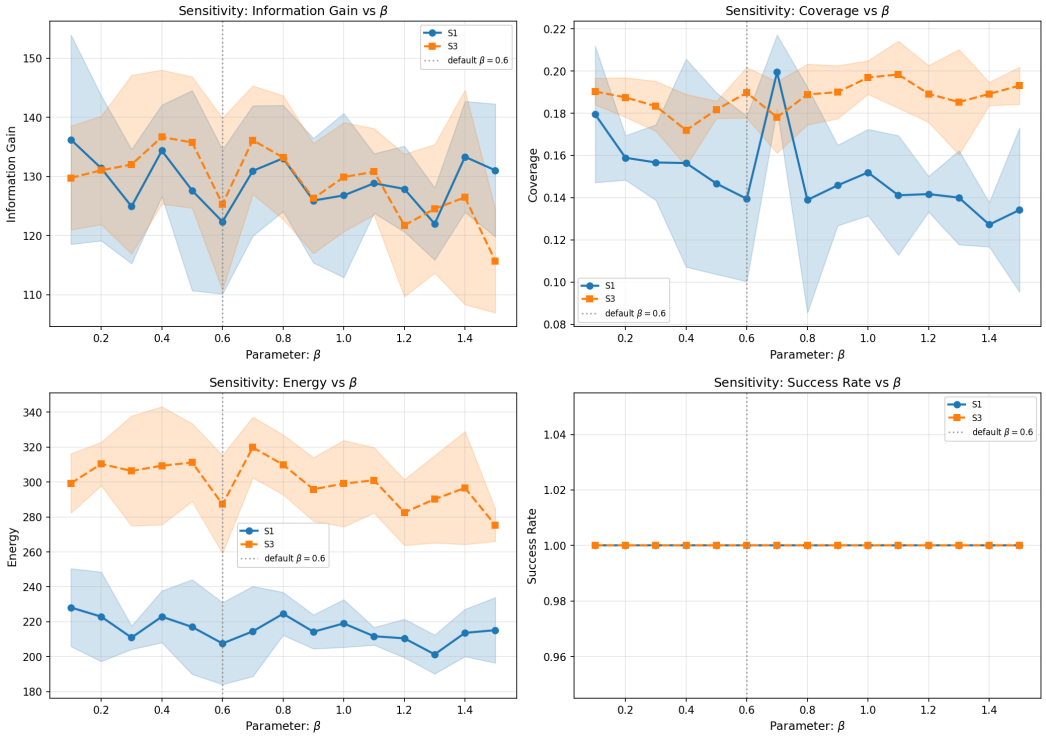


Figure 5. Sensitivity of monitoring metrics to the energy-trade-off weight β (mean $\pm$ std over $n = 5$ seeds; shaded bands show $\pm$ std). Curves for Scenarios S1 and S3 use distinct colors and markers. In the Success Rate panel the two curves coincide at 1.0 for all tested β, indicating that both scenarios remain fully successful under the scanned energy weights. The selected operating point $\beta = 0.6$ balances information gain against energy expenditure without inducing budget violations.

The whole framework is developed in Python 3.10: NumPy and SciPy handle field numerical integration, while quantum computations rely on Qiskit 0.46 with the Aer state-vector simulator (qiskit-aer 0.13) [40]. All experiments run on a workstation equipped with an Intel i9 CPU and an NVIDIA RTX 4090 GPU.

All QAOA evaluations in this study are performed with the Qiskit Aer state-vector simulator. The simulator implements the intended unitary evolution and measurement statistics, but does not include device noise such as gate infidelity or decoherence. The resulting online instances use a pruned receding-horizon QUBO with planning horizon $H_{\text{plan}} = 2$ and $m = 4$ candidates per step ($n = H_{\text{plan}} m = 8$ qubits), which lies well within the qubit count of current NISQ processors; validation on physical quantum hardware is outlined in Section 5.

Use of Generative AI

In line with the MDPI research and publication ethics policy on generative artificial intelligence (GenAI), we disclose the limited use of GenAI for graphics preparation. ChatGPT Images 2.0 (OpenAI; available online: https://openai.com/index/introducing-chatgpt-images-2-0/; accessed on 17 August 2026) was employed solely to generate non-textual

illustrative elements in Figure 1 (the first figure of this manuscript), such as decorative water-flow scene motifs used to visualize the marine environment. The overall figure composition, scientific panels, equations, annotations, and all other graphical components were constructed and assembled manually by the authors. GenAI was not used for study design, data collection, numerical analysis, interpretation of results, or for generating the scientific content of the manuscript text.

4.2. Discrete Mission Scenarios (S0–S4)

To evaluate the planners under progressively harder mission conditions, we construct five discrete task scenarios (S0–S4) on a shared grid world. All scenarios use the same map size and geometry: a 60×60 occupancy grid, start node $(3, 3)$, goal node $(56, 56)$, and five circular obstacles centered at $(20, 20)$, $(35, 25)$, $(22, 42)$, $(45, 45)$, $(50, 30)$ with radii $5, 6, 5, 5, 4$, respectively. Figure 6 shows the unified environmental map for the most constrained scenario S4.

The pollutant field is built from three components:

$$C = \mathrm{norm}(C_{\mathrm{base}} + \alpha_s C_{\mathrm{side}} + 0.65\,\alpha_d C_{\mathrm{drift}}),$$

where

$$C_{\mathrm{base}} = \mathcal{G}(14, 17, 5.0) + 0.9\mathcal{G}(41, 40, 6.2) + 0.8\mathcal{G}(46, 14, 4.8),$$
$$C_{\mathrm{side}} = \mathcal{G}(18, 30, 4.2) + \mathcal{G}(34, 16, 4.0) + \mathcal{G}(28, 47, 4.6),$$
$$C_{\mathrm{drift}} = \mathcal{G}(24, 24, 6.5) + \mathcal{G}(39, 34, 6.0),$$

and $\mathcal{G}(c_x, c_y, \sigma)$ is a 2D Gaussian.

The information map is

$$I = \mathrm{norm}(1.2\,C + 0.8\,\|\nabla C\| + 1.0\,U),$$

where the uncertainty term follows the same epistemic role as Equation (10). On the discrete S0–S4 grids we instantiate U by a GP-free surrogate,

$$U(x) = U_0(x) = \mathrm{norm}\big(d(x, x_s)\big),$$

i.e., a normalized distance-from-start prior that is attenuated after sampling, as in the $\bar{U}$ metric below; this surrogate preserves the exploration incentive of σ^2_{GP} when a full online GP update is not maintained on every discrete cell, while continuous perception still uses Equation (10). To model directional environmental resistance, we define a reverse-flow penalty:

$$F = \mathrm{clip}\Big(1 + \alpha_r \exp\Big(-\big(d_{\mathrm{diag}}/5\big)^2\Big),\ 1.0,\ 2.6\Big),$$

where $d_{\mathrm{diag}} = |(y - x) - (y_s - x_s)|$ and $(x_s, y_s) = (3, 3)$.

Mission success is computed uniformly for all methods:

$$\begin{aligned}\mathrm{Success} = 1 \iff\ & (\text{goal reached}) \wedge (L \le L_{\max}) \\ & \wedge\ (E_{\mathrm{eval}} \le E_{\max}) \wedge (\mathrm{Coverage} \ge C_{\min}),\end{aligned}$$

where L is path length, E_{eval} is evaluation energy (including flow penalty), and coverage is high-gradient-region coverage.

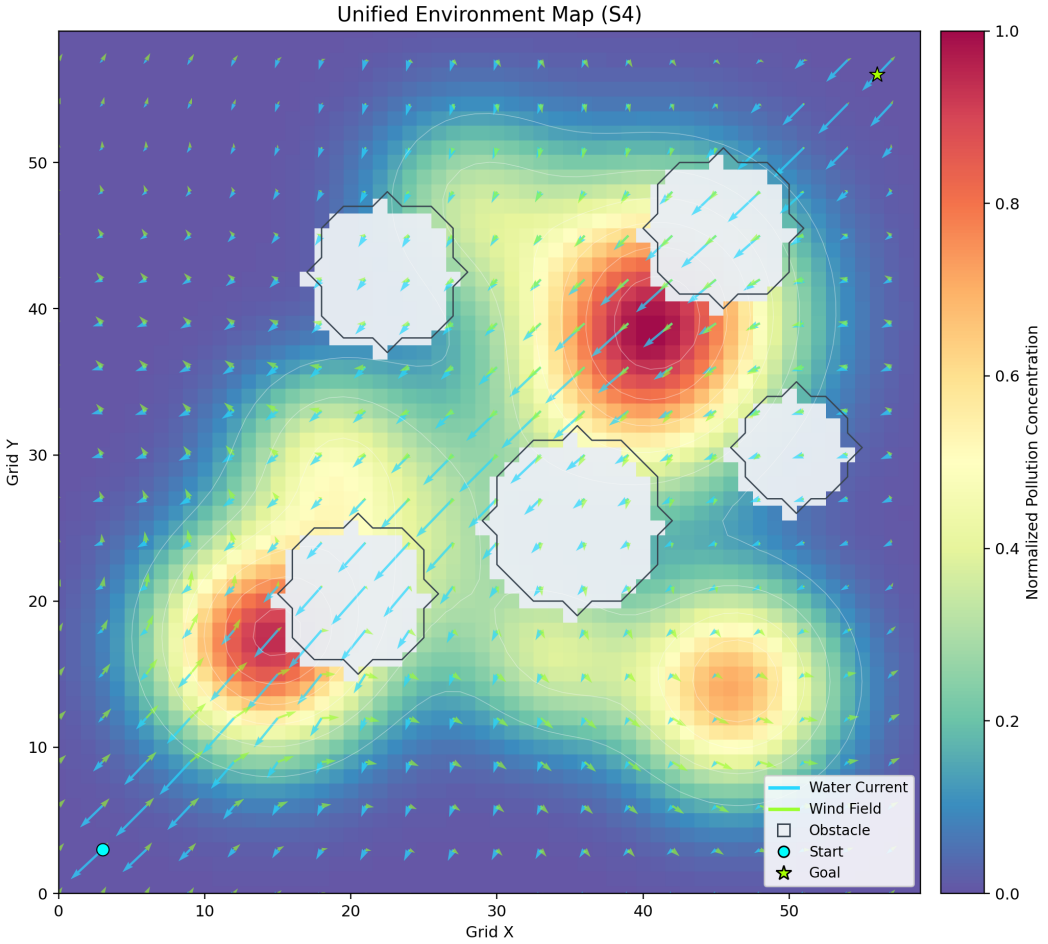


Figure 6. Unified environmental map of test scenario S4. The color gradient denotes normalized pollution concentration, cyan arrows represent water current vectors, light green arrows indicate wind field vectors, white polygon blocks are impassable obstacles, the cyan circle marks the USV starting position, and the yellow star denotes the target sampling destination. Contour lines outline the spatial distribution of pollution plumes.

Table 2 lists the scenario-wise hyperparameters used throughout the evaluation.

Table 2. Hyperparameter configuration for all test scenarios.

Scenario	α_s	α_d	α_r	$E_{\max}$	$L_{\max}$	$C_{\min}$
S0 (geometry baseline)	0.00	0.00	0.00	10^9	10^9	0.00
S1 (side-lobe information)	0.55	0.00	0.00	10^9	10^9	0.00
S2 (drifting hotspots)	0.45	0.60	0.00	10^9	10^9	0.00
S3 (reverse-flow corridor)	0.35	0.45	0.95	10^9	10^9	0.00
S4 (dual-budget mission)	0.35	0.55	1.10	360	220	0.17

S0–S3 mainly test environmental distinguishability and planner behavior under increasing field complexity, while S4 is the constrained mission scenario used to test goal-reaching, information collection, and resource-awareness jointly.

4.3. Performance Metrics

To ensure fair cross-method comparison, all planners are evaluated using the same unified metric set on the same discretized grid environment. The reported metrics are grouped into information quality, resource cost, trajectory quality, task success, and computational cost.

4.3.1. Information Quality

The cumulative information gain along trajectory $\pi = \{x_t\}_{t=0}^{T}$ is

$$G_{\text{eval}}(\pi) = \sum_{t=0}^{T} I(x_t),$$

where $I(\cdot)$ is the normalized information map.

High-gradient coverage is defined on the top-gradient region

$$\Omega_g = \{x \in \Omega_{\text{free}} : \|\nabla I(x)\| \geq q_{0.8}\},$$

where $q_{0.8}$ is the 80th percentile of $\|\nabla I\|$. Let Ω_s be cells sampled at least once (within sensing footprint). Coverage is

$$\text{Cov}(\pi) = \frac{|\Omega_s \cap \Omega_g|}{|\Omega_g|}.$$

Posterior uncertainty for discrete-scenario scoring uses the same epistemic surrogate as the planner's U term, attenuated by repeated sampling:

$$U_{\text{post}}(x) = U_0(x) \exp\big(-0.65\, n_s(x)\big),$$
$$\bar{U} = \frac{1}{|\Omega_{\text{free}}|} \sum_{x \in \Omega_{\text{free}}} U_{\text{post}}(x),$$

where $U_0(x) = \text{norm}(d(x, x_s))$ is the distance-to-start prior (a GP-free stand-in for σ^2_{GP}) and $n_s(x)$ is the local sampling count. Thus, planning and evaluation share one uncertainty semantics on S0–S4, while Equation (10) remains the definition under online GP perception.

4.3.2. Resource Cost and Trajectory Quality

Path length is

$$L(\pi) = \sum_{t=1}^{T} \|x_t - x_{t-1}\|_2.$$

Energy includes flow-weighted motion and turning penalty:

$$E_{\text{eval}}(\pi) = \sum_{t=1}^{T} F(x_t)\|x_t - x_{t-1}\|_2 + 0.08 \sum_{t=2}^{T} |\Delta\theta_t|,$$

where $F(x) \in [1, 2.6]$ is the flow-penalty map.

Redundancy is measured by the revisit ratio:

$$R(\pi) = \frac{(|\pi| - |\text{Unique}(\pi)|)}{|\pi|}.$$

4.3.3. Task-Level Feasibility

A rollout is successful iff all mission constraints are satisfied:

$$\text{Succ}(\pi) = \mathbf{1}[x_T = x_g \wedge L(\pi) \leq L_{\max}$$
$$\wedge E_{\text{eval}}(\pi) \leq E_{\max} \wedge \text{Cov}(\pi) \geq C_{\min}].$$

For multi-seed evaluation, the success rate is the mean of $\text{Succ}(\pi)$ over runs.

Constraint violation is defined as the motion infeasibility ratio (out-of-bound, obstacle collision, or non-adjacent move), with an additional penalty for mission-budget or coverage failure:

$$V(\pi) = \min\left(1,\ v_{\text{motion}} + 0.2\,\mathbf{1}_{\text{budget fail}} + 0.2\,\mathbf{1}_{\text{coverage fail}}\right).$$

4.3.4. Computational Cost

Runtime is the wall-clock planning time per method:

$$t_{\text{run}} = t_{\text{end}} - t_{\text{start}}.$$

This runtime definition is uniformly applied to all baselines and the proposed method.

Figure 7 visualizes representative trajectories of all compared planners on Scenario S4.

4.4. Comparative Results

The comparison set covers four complementary families: geometric shortest-path planners (A*, Dijkstra) as sensing-agnostic references; a one-step greedy informative planner (Greedy-Info); metaheuristic optimizers of the same composite objective $G - \beta E$ (PSO-IPP, Genetic, SimAnneal); and static quantum solvers (Static-QUBO, Global-QAOA) that share the QUBO encoding but omit receding-horizon replanning. All methods use the same information map $I = w_c C + w_g \|\nabla C\| + w_u U$ and the same evaluation protocol. Metaheuristic budgets were tuned once on Scenario S0 and then frozen across scenarios. Unless noted otherwise, the frozen settings are PSO-IPP with 30 particles, 72 iterations, horizon 165, inertia $w = 0.72$, and cognitive/social coefficients $c_1 = c_2 = 1.30$; Genetic with population 68, 90 generations, horizon 168, and mutation probability 0.07; and SimAnneal with 1800 iterations, initial temperature 1.6, and cooling factor 0.996. Geometric planners use 8-connected grid search with unit step cost; Static-QUBO / Global-QAOA reuse the same QUBO coefficients (β, γ, ρ) as the proposed method but without rolling reconstruction.

Table 3 summarizes the quantitative performance of nine planners on the constrained marine monitoring scenario, categorized into geometric pathfinding, heuristic informative planners, and static quantum baselines [6,7,10,18,41,42]. Table 4 reports the corresponding multi-seed mean ± std over the same five seeds. Deterministic planners are seed-invariant; among the stochastic methods, CV(G_{eval}) ranges from about 1.9% (Greedy-Info) to 14.5% (Genetic), with the proposed planner at 6.3%. Because several baselines are deterministic (zero variance), we do not report pairwise p-values against them; instead we emphasize effect sizes and CV. Relative to the strongest stochastic classical solver (SimAnneal), the proposed method retains a substantially higher mean G_{eval} (134.7 vs. 87.3). Table 5 additionally lists ground-truth-field statistics (G_{true}) for the proposed planner; these values use a different evaluator from G_{eval} and should not be compared numerically to Tables 3 and 4.

In terms of core sensing performance, Greedy-Info yields the highest information gain (149.33) and pollution gradient coverage (27.64%), but with substantial resource overhead. The proposed *Proposed-Replan-QAOA* achieves the second-best information gain of 141.63, outperforming the best metaheuristic SimAnneal (89.01) by 59.12%, and attains the minimal posterior uncertainty $\bar{U} = 0.463$, demonstrating superior balancing of exploration and exploitation. By contrast, the geometric planners (A*, Dijkstra) and static quantum methods (Static-QUBO, Global-QAOA) produce low information gain and limited coverage under the constrained scenario.

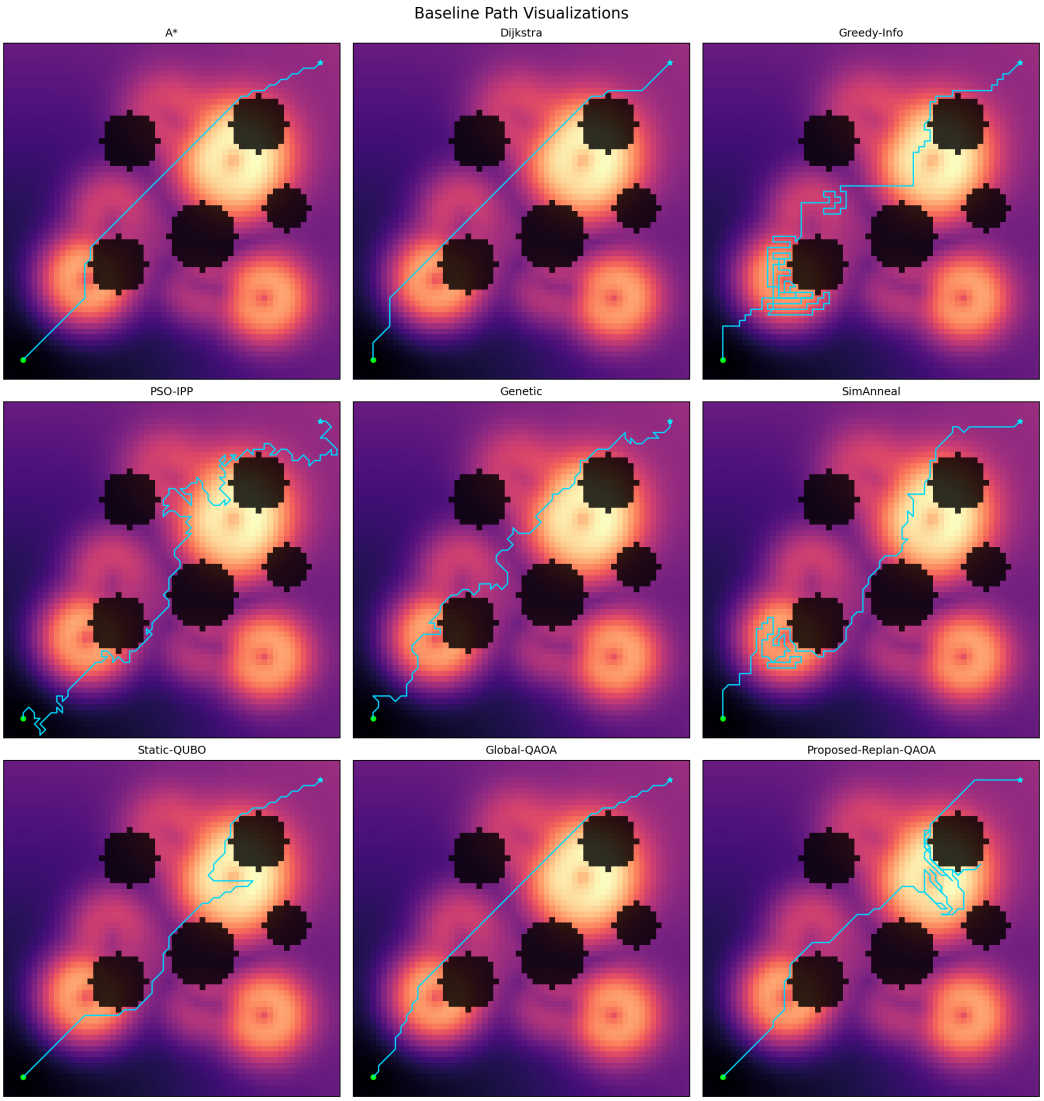


Figure 7. Trajectory visualization of all compared path planning baselines on Scenario S4. The background colormap denotes normalized pollution concentration (red for high-pollution hotspots, purple for clean seawater), black polygons represent static marine obstacles, green dots mark the USV starting position, the blue star marks the goal location, and cyan curves depict the generated navigation trajectories. Geometric planners (A*, Dijkstra, Global-QAOA) produce straight, short paths that ignore pollutant plumes; Greedy-Info and metaheuristic algorithms (PSO-IPP, Genetic, SimAnneal) over-explore local hotspots with excessive detours and redundant sampling; Static-QUBO generates fixed offline trajectories incapable of adapting to dynamic pollution fields. The proposed Proposed-Replan-QAOA balances sufficient coverage of high-concentration pollution zones and reasonable travel overhead via receding-horizon online replanning.

The moderate coverage of the proposed method (17.92%) relative to Greedy-Info (27.64%) is consistent with the metric definitions: Coverage counts unique visited cells and therefore favors spatial breadth, whereas G_{eval} accumulates information from every sample and therefore favors repeated visits to high-value hotspots. In Scenario S4, pollutant mass is concentrated in a few localized regions; with an explicit anti-redundancy term, the planner covers about 215 unique cells while concentrating samples where the per-visit return is highest. Greedy-Info spreads more widely and thus reports higher coverage at a slightly lower information density. The two scores therefore capture complementary aspects of monitoring quality rather than a single ranking axis.

Table 3. Quantitative comparison of different path planning methods on the target constrained scenario (shared reference seed 7 from the seed set {7, 14, 21, 28, 35}). Deterministic methods are seed-invariant. Multi-seed mean ± std over all five seeds are reported in Table 4; ground-truth multi-seed metrics for the proposed planner appear in Table 5. A* denotes the classical A-star shortest-path planner (the asterisk is part of the conventional algorithm name). Boldface highlights the best value in each column.

Method	$G_{\text{eval}}(\pi)$	$\bar{u}$	Coverage (%)	Energy	Viol. Rate (%)	Runtime (s)
A*	33.87	0.487	14.58	116.46	20.00	0.01
Dijkstra	32.01	0.487	15.83	105.32	20.00	0.02
Greedy-Info	**149.33**	0.464	**27.64**	365.38	20.00	**0.00**
PSO-IPP	85.97	0.454	23.33	304.90	**0.00**	7.42
Genetic	72.95	0.468	20.14	225.37	**0.00**	13.99
SimAnneal	89.01	0.473	24.58	248.09	**0.00**	3.16
Static-QUBO	50.99	0.478	15.14	166.86	20.00	**0.00**
Global-QAOA	33.19	0.486	15.42	108.85	20.00	24.10
Proposed-Replan-QAOA	141.63	**0.463**	17.92	345.88	**0.00**	2.62

Table 4. Multi-seed statistics on Scenario S4 over $n = 5$ seeds {7, 14, 21, 28, 35} (mean ± std). Deterministic planners have negligible seed-to-seed variance in G_{eval}. Boldface highlights the best value in each column.

Method	G_{eval}	$\bar{u}$	Energy
A*	33.87 ± 0.00	0.487 ± 0.000	116.5 ± 0.0
Dijkstra	32.01 ± 0.00	0.487 ± 0.000	105.3 ± 0.0
Greedy-Info	**149.2 ± 2.8**	0.462 ± 0.005	365.3 ± 26.2
PSO-IPP	75.1 ± 6.4	**0.461 ± 0.006**	261.6 ± 17.5
Genetic	61.3 ± 8.9	0.471 ± 0.005	206.8 ± 17.0
SimAnneal	87.3 ± 10.1	0.474 ± 0.002	236.2 ± 21.7
Static-QUBO	50.99 ± 0.00	0.478 ± 0.000	166.9 ± 0.0
Global-QAOA	33.19 ± 0.00	0.486 ± 0.000	108.8 ± 0.0
Proposed-Replan-QAOA	134.7 ± 8.5	0.465 ± 0.003	331.4 ± 23.0

Table 5. Multi-seed ground-truth metrics of the proposed planner on Scenario S4 ($n = 5$). All entries use a ground-truth-field evaluator that differs from the planner-side G_{eval}/coverage/energy definitions in Tables 3 and 4; numerical values in this table must not be compared directly with those tables. CV denotes the coefficient of variation.

Metric	Mean	Std	CV (%)
G_{true}	686.6	36.1	5.26
Energy	232.9	12.5	5.39
Coverage (%)	36.08	5.52	15.3
Constraint violations	0.00	0.00	0.00

For mission feasibility, PSO-IPP, Genetic, SimAnneal, and the proposed method achieve a zero violation rate on the reference seed in Table 3. The geometric and static quantum baselines show a 20% violation rate.

Regarding computational overhead, Global-QAOA suffers the longest runtime (24.10 s) due to full-horizon global optimization. The swarm-based heuristics also incur heavy latency (PSO-IPP: 7.42 s, Genetic: 13.99 s). Benefiting from receding-horizon decomposition, the proposed QAOA planner only takes 2.62 s per replanning cycle, which fits the 4 s simulated replanning interval. Unless otherwise stated, the reported latencies are wall-clock times on an Intel i9 workstation in simulation, rather than onboard USV or quantum-hardware measurements; claims of online feasibility in this paper therefore refer to the simulated control loop, and hardware validation is discussed in Section 5.

The proposed method's moderately higher energy consumption and sampling redundancy are acceptable trade-offs to continuously track drifting pollutant plumes under stochastic ocean currents and wind drift.

The higher energy consumption relative to the geometric planners should be interpreted jointly with sensing return. Dijkstra reaches $E = 105.32$ by ignoring pollutant structure and obtains only $G = 32.01$, about 22.6% of the proposed $G = 141.63$. Relative to the informative baselines, the proposed $G/E \approx 0.41$ matches Greedy-Info while reducing the violation rate from 20% to zero. Static-QUBO ($E = 166.86$, $G = 50.99$) plans on a frozen snapshot and likewise incurs 20% violations under drift. Among the methods that simultaneously attain high information gain and zero violations, the proposed planner is the most energy-efficient; increasing β (e.g., to 1.0) can further reduce energy at the expense of information gain when mission priorities so require.

Overall, the receding-horizon quantum framework provides the most balanced performance among methods that jointly pursue high information return and mission feasibility in the simulated dynamic monitoring setting, rather than dominating every individual metric.

4.5. Ablation Studies

Ablation experiments are conducted to quantify the contribution of each key module in the proposed framework under the discrete mission settings (S2/S3/S4). Five configurations are evaluated while keeping all map settings, obstacle layouts, and evaluation protocols identical. The results are averaged over 20 seeds per scenario and reported in Table 6. To avoid misleading conclusions from single-metric inflation (e.g., excessive revisits that increase $G_{\text{eval}}(\pi)$), we additionally report a composite score:

$$\text{Score} = G_{\text{eval}} - \lambda E_{\text{eval}} - \mu\,\text{Red} - \nu\,\text{Viol},$$

with $(\lambda, \mu, \nu) = (0.35, 1.20, 2.50)$.

Table 6. Ablation study on discrete mission scenarios S2/S3/S4. Entries are mean ± std over 20 seeds per scenario. Composite score uses Score $= G_{\text{eval}} - \lambda E_{\text{eval}} - \mu\,\text{Red} - \nu\,\text{Viol}$ with $(\lambda, \mu, \nu) = (0.35, 1.20, 2.50)$. Boldface highlights the best value in each column.

Method	Score	G	$\bar{U}$	Cov.	Energy	Viol.	Red.
Full model	**2.6 ± 22.4**	113.6 ± 14.7	0.464 ± 0.003	17.6 ± 1.5	246 ± 55	**1.7 ± 5.5**	17.3 ± 5.0
Var. A (no unc.)	0.0 ± 27.1	113.6 ± 16.3	0.464 ± 0.004	17.1 ± 2.0	244 ± 52	3.0 ± 7.1	17.1 ± 5.0
Var. B (no anti-red.)	−11.8 ± 41.7	**177.5 ± 11.4**	**0.462 ± 0.003**	**18.5 ± 1.9**	354 ± 61	7.3 ± 10.9	39.1 ± 3.9
Var. C (no replan)	−18.6 ± 26.1	32.6 ± 1.0	0.481 ± 0.008	11.4 ± 4.3	**98.5 ± 12.4**	6.7 ± 9.4	**0.0 ± 0.0**
Var. D (no quantum)	−14.3 ± 31.4	50.2 ± 1.8	0.479 ± 0.001	14.2 ± 0.7	136.6 ± 27.0	6.7 ± 9.4	**0.0 ± 0.0**

Variant A removes the uncertainty term from the information map. Compared with the full model, $G_{\text{eval}}(\pi)$ remains nearly unchanged (113.6 vs. 113.6), but coverage decreases (17.1% vs. 17.6%) and the violation rate increases (3.0% vs. 1.7%), reducing the composite score from 2.6 to 0.0. This indicates that uncertainty-aware weighting primarily improves safety and mission consistency rather than raw accumulation.

Variant B disables the anti-redundancy mechanisms. Although this variant yields higher $G_{\text{eval}}(\pi)$ (177.5) and slightly better $\bar{U}$ and coverage, it achieves this at the cost of substantially higher energy (354.3), violation rate (7.3%), and redundancy (39.1%). Its composite score drops to −11.8, showing that the apparent gain is dominated by inefficient repeated sampling.

Variant C disables receding-horizon replanning and performs one-shot global optimization. This reduces energy and redundancy, but severely harms task-oriented information performance ($G_{\text{eval}}(\pi) = 32.6$, coverage = 11.4%, $\bar{U} = 0.481$), resulting in the lowest score (−18.6). This result confirms that online adaptation is essential in nonstationary information landscapes.

Variant D replaces the quantum optimizer with a classical non-quantum QUBO solver. Similar to Variant C, it is more conservative in energy but significantly underperforms in information acquisition ($G_{\text{eval}}(\pi) = 50.2$, coverage = 14.2%, $\bar{U} = 0.479$), with a negative score (-14.3).

Overall, while several of the ablations can improve isolated metrics, only the full model achieves the best balanced performance, obtaining the highest composite score (2.6) and the lowest violation rate (1.7%). These results validate the necessity of combining uncertainty-aware information shaping, anti-redundancy control, and receding-horizon quantum-assisted replanning.

Trajectory-level inspection clarifies the ablation outcomes. Without the uncertainty term (Variant A), the vehicle remains near previously high-value zones adjacent to obstacles and explores less, which raises the violation rate even though raw G is almost unchanged. Disabling anti-redundancy (Variant B) produces oscillatory revisits of a few hotspots: path length grows by about 43%, energy rises from 246 to 354, and redundancy reaches 39.1%, so the highest G coincides with the worst composite score. One-shot planning (Variant C) freezes the initial snapshot; after plumes drift by roughly 10–15 cells in S2/S3, the vehicle finishes early targets and then travels with little additional information return. Replacing QAOA by classical simulated annealing on the same small QUBO (Variant D) yields more conservative short hops and lower information acquisition, indicating that the quantum solver is at least competitive at $n = 8$ and remains the component intended to scale with larger encodings.

4.6. Computational Performance Analysis

Under the proposed benchmarking protocol, all the complexity statistics are averaged over 50 independent runs across grid scales $N \in \{20, 30, 40, 50, 60, 80\}$. A key distinction throughout this section is that construction size ≠ quantum circuit width. Table 7 reports two related but distinct sizes. The full-grid assembly column uses a construction horizon $H_{\text{full}} = 8$ over all grid nodes, i.e., $H_{\text{full}}|\mathcal{U}|$ binary coefficients (3200–51,200 as the map grows from 20×20 to 80×80), which governs only classical QUBO construction time and memory. In contrast, the online QAOA solve uses candidate pruning with $H_{\text{plan}} = 2$ and $m = 4$, and hence a fixed circuit width $n = H_{\text{plan}}m = 8$ qubits at every replanning step. The depth sweeps and QAOA quality claims in this paper refer exclusively to these pruned $n = 8$ instances; the 51,200 figure must not be read as the number of qubits.

Table 7. QUBO construction and computational scaling under different grid resolutions. Importantly, construction size ≠ quantum circuit width: "Full-grid vars." counts the naive encoding $H_{\text{full}}|\mathcal{U}|$ with $H_{\text{full}} = 8$ used only for classical construction/memory scaling, whereas the online QAOA solve uses candidate pruning with $H_{\text{plan}} = 2$, $m = 4$, hence a fixed circuit width $n = 8$ qubits. The results are averaged over repeated runs.

Grid Size	Nodes $\lvert\mathcal{U}\rvert$	Full-Grid Vars.	Constr. Time (s)	Memory (MB)	Replan Lat. (s)
20×20	400	3200	0.045	2.2	0.052
30×30	900	7200	0.112	5.0	0.052
40×40	1600	12,800	0.208	9.0	0.052
50×50	2500	20,000	0.334	14.2	0.053
60×60	3600	28,800	0.488	20.5	0.053
80×80	6400	51,200	0.873	36.7	0.053

The measured QUBO construction time grows monotonically from 0.0446 s (20×20) to 0.8727 s (80×80), while peak memory usage increases from 2.17 MB to 36.69 MB. This trend is consistent with the expected superlinear growth of constrained QUBO assembly under higher spatial resolution.

For online deployment, we additionally report single-cycle receding-horizon optimization latency. Across all scales, replanning latency remains tightly bounded around 0.052 s (range: 0.0518–0.0529 s), which is well below the control step $\Delta t = 0.5$ s. The timing window here measures only the pruned QAOA solve/decode stage ($H_{\text{plan}} = 2$, $m = 4$, $n = 8$, fixed shot budget, fixed optimizer-iteration cap), while map update and QUBO construction are reported separately in Table 7; this measurement protocol explains the near-constant latency across grid resolutions. In particular, at 60 × 60, the average latency is 0.0529 s, confirming online feasibility in simulation with substantial timing margin.

The 2.62 s entry in Table 3 is the end-to-end wall-clock time of one replanning cycle, including measurement update, QUBO construction, the full 15 × 15 variational grid search, and trajectory decoding. The ~0.052 s figure is a narrower micro-benchmark of the QAOA solve/decode stage on a pre-built QUBO with a reduced search budget (50 rather than 225 evaluations). For online operation, the end-to-end 2.62 s value is the relevant latency and remains within the 4 s replanning interval, while the micro-benchmark isolates solver overhead for scaling studies.

The depth–performance trade-off of QAOA is evaluated through a dedicated depth-sweep experiment. As depth increases from $p = 1$ to $p = 8$, runtime rises approximately linearly from 3.33 s to 53.43 s. By contrast, objective improvement is marginal: the best value changes from 42.516 to 42.374 (about 0.33% relative improvement), indicating diminishing returns at larger p. Therefore, moderate depth provides a more practical balance between optimization quality and computational overhead.

Overall, the results show that the proposed framework scales predictably with discretization, satisfies the simulated replanning-interval constraints at the target operating resolution, and exhibits a clear depth–latency trade-off in quantum optimization.

5. Discussion

The proposed framework integrates a stochastic marine environment model, an information-driven objective, and a hybrid quantum–classical optimization pipeline. The experimental results indicate consistent improvements over classical and heuristic baselines across all scenarios; however, the performance gains should be interpreted in the context of modeling assumptions and computational trade-offs inherent to the formulation.

A central observation is that the coupling between environmental dynamics and planning decisions significantly affects trajectory quality in time-varying flow fields. In particular, the incorporation of a stochastic differential representation of ocean currents and wind drift provides a more realistic description of transport dynamics compared to static flow assumptions. This enables the planner to implicitly account for advective distortion of both the pollutant field and the vehicle trajectory, which becomes especially relevant under Scenario 2 and Scenario 3, where environmental variability dominates system behavior.

The introduction of an uncertainty-aware information field further influences the exploration–exploitation balance of the planner. By combining concentration magnitude, spatial gradients, and epistemic uncertainty, the resulting objective promotes sampling diversity in regions that are either poorly observed or dynamically evolving. The ablation results suggest that removing this term leads to a measurable increase in posterior uncertainty and reduced coverage of high-gradient regions, indicating that uncertainty modeling plays a complementary role rather than acting as a secondary heuristic term.

From an optimization perspective, the QUBO-QAOA formulation provides a flexible discrete representation of the trajectory planning problem. While the variational quantum solver does not guarantee global optimality, its expressive parameterized structure allows exploration of a broader solution space compared to greedy decoders. Nevertheless, this benefit comes at the cost of increased computational overhead, particularly when scaling to

finer discretizations or higher circuit depths [43,44]. The complexity analysis shows that the current implementation remains within the simulated replanning budget for moderate grid resolutions, but scalability to significantly larger domains would require further algorithmic or hardware-level improvements.

It is also important to note that the current framework assumes accurate estimation of the underlying flow field and pollutant dynamics through sensor fusion and Gaussian process inference. In practice, model mismatch or sensor noise could degrade the quality of the reconstructed information field, thereby affecting downstream planning performance. Future extensions may consider robust optimization formulations or Bayesian reinforcement learning approaches to explicitly handle such uncertainties.

Finally, although the proposed method demonstrates strong empirical performance across diverse simulated scenarios, validation in real-world marine environments remains an important next step. Factors such as measurement delay, communication constraints, and actuator non-idealities are not fully captured in the current simulation model [3,5]. Addressing these aspects would further strengthen the applicability of the framework to practical autonomous water quality monitoring missions.

Controlled perturbation tests ($n = 5$ seeds) further support closed-loop robustness. Under multiplicative concentration noise $\mathcal{N}(0, 0.1C)$ and/or GPS bias of ± 3 m, information gain evaluated on the ground-truth field remains comparable to the clean baseline ($G_{\text{true}} = 666.7 \pm 65.0$ versus 694.2 ± 44.9, 686.6 ± 40.8, and 668.5 ± 90.9), with no constraint violations; the 4 s replanning period limits error accumulation. When planner-side SDE parameters are misspecified (set to $\tau_c' = 1.5\tau_c$ and $\kappa' = 2\kappa$), G_{true} decreases by only 2.8% and coverage from 38.6% to 33.1%, again without violations, indicating that receding-horizon updates compensate for moderate flow miscalibration. Across five Scenario S4 seeds, $G_{\text{true}} = 686.6 \pm 36.1$ with $\text{CV}(G) = 5.26\%$ and $\text{CV}(E) = 5.39\%$. Remote quantum execution would additionally require round-trip latency below the replanning interval, with classical QUBO solving as a fallback; richer actuator limits can be absorbed into the transition costs c_{ij}.

From a scaling viewpoint (Table 7), QUBO construction grows from 0.045 s on a 20×20 grid to 0.873 s on 80×80, still within the 4 s replanning budget, while the solve latency stays near 0.052 s because each online horizon uses a fixed circuit width $n = H_{\text{plan}} m = 8$ qubits ($H_{\text{plan}} = 2$, $m = 4$). Thus, larger maps increase classical construction/map-evaluation cost, not quantum circuit width—again underscoring that construction size $\neq$ qubit count; encodings with $n \approx 32$ remain a natural next step as hardware improves.

The main limitations of the present study are the reliance on simulation, the lack of buoy/ADCP calibration for SDE parameters, and the small QUBO size. A staged validation is planned: classical QUBO control on a physical USV in a confined pond; cloud NISQ execution of $p = 1$, $n = 8$ circuits; extension to tidal environments; and progressive scaling beyond 32 qubits.

6. Conclusions

This paper presented an information-driven quantum path planning framework for USV-based water quality monitoring in stochastic marine environments. The method combines a coupled current–wind stochastic field model, an uncertainty-aware information map, a QUBO-based discrete reformulation, and QAOA-enabled receding-horizon replanning.

Experimental results on constrained mission scenarios show that the proposed approach achieves strong information acquisition and low posterior uncertainty on the reference evaluation seed, with wall-clock latency compatible with the simulated replanning interval. Absolute superiority on every metric is not claimed: greedy informative search can attain higher raw G_{eval} and coverage. Additional ablation and scaling analyses indicate that

uncertainty modeling, anti-redundancy control, and adaptive replanning are all necessary for balanced mission performance.

Future work will focus on extending robustness under sensor/model mismatch, incorporating noise-aware variational quantum optimization [43,45], and validating the full pipeline on physical USV platforms in real marine deployments.

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

Funding: This research was supported by the Major Basic Science (Natural Science) Research Project of the Jiangsu Higher Education Institutions of China ("Research on Ship Navigation Risk Assessment Based on Multi-source Heterogeneous Data Fusion", Grant No. 23KJA580002), the General Project of Philosophy and Social Science Research of the Jiangsu Higher Education Institutions of China (Grant No. 2025SJYB0544), and the 2022 Excellent Teaching Team of the QingLan Project of the Jiangsu Higher Education Institutions of China ("Big Data Talent Cultivation Teaching Team with Shipping Characteristics", Document No. Su Jiaoshi Han [2022] No. 29).

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

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT Images 2.0 (OpenAI; available online: https://openai.com/index/introducing-chatgpt-images-2-0/; accessed on 17 August 2026) for the purposes of creating non-textual illustrative elements (e.g., water-flow scene motifs) in Figure 1. The authors have reviewed and edited the output and take full responsibility for the content of this publication. All individuals named in this section have consented to being acknowledged. The authors would also like to acknowledge the support from the research team.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

USV	Unmanned surface vehicle
QUBO	Quadratic unconstrained binary optimization
QAOA	Quantum Approximate Optimization Algorithm
SDE	Stochastic differential equation
GP	Gaussian process

References

1. Fischer, H.B.; List, E.J.; Koh, R.C.Y.; Imberger, J.; Brooks, N.H. *Mixing in Inland and Coastal Waters*; Academic: New York, NY, USA, 1979.
2. Chapra, S.C. *Surface Water-Quality Modeling*; McGraw-Hill: New York, NY, USA, 1997.
3. Manley, J.E. Unmanned surface vehicles, 15 years of development. In Proceedings of the OCEANS 2008, Quebec City, QC, Canada, 15–18 September 2008; pp. 1–4. [CrossRef]
4. Petres, C.; Pailhas, Y.; Petillot, Y.; Evans, J.; Lane, D. Path planning for autonomous underwater vehicles. *IEEE Trans. Robot.* **2007**, *23*, 331–341. [CrossRef]
5. Fossen, T.I. *Handbook of Marine Craft Hydrodynamics and Motion Control*; Wiley: Chichester, UK, 2011. [CrossRef]
6. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numer. Math.* **1959**, *1*, 269–271. [CrossRef]
7. Hart, P.E.; Nilsson, N.J.; Raphael, B. A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. Syst. Sci. Cybern.* **1968**, *4*, 100–107. [CrossRef]
8. Kavraki, L.E.; Svestka, P.; Latombe, J.-C.; Overmars, M.H. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Trans. Robot. Autom.* **1996**, *12*, 566–580. [CrossRef]
9. LaValle, S.M. *Planning Algorithms*; Cambridge University Press: Cambridge, UK, 2006. [CrossRef]

10. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95—International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; pp. 1942–1948. [CrossRef]
11. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006. [CrossRef]
12. Krause, A.; Singh, A.; Guestrin, C. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.* **2008**, *9*, 235–284. Available online: https://www.jmlr.org/papers/v9/krause08a.html (accessed on 25 October 2026).
13. Singh, A.; Krause, A.; Guestrin, C.; Kaiser, W.J. Efficient informative sensing using multiple robots. *J. Artif. Intell. Res.* **2009**, *34*, 707–755. [CrossRef]
14. Hollinger, G.A.; Sukhatme, G.S. Sampling-based robotic information gathering algorithms. *Int. J. Robot. Res.* **2014**, *33*, 1271–1287. [CrossRef]
15. Marchant, R.; Ramos, F. Bayesian optimisation for informative continuous path planning. In Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA), Hong Kong, China, 31 May–7 June 2014; pp. 6136–6143. [CrossRef]
16. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; de Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **2016**, *104*, 148–175. [CrossRef]
17. Garau, B.; Alvarez, A.; Oliver, G. Path planning of autonomous underwater vehicles in current fields with complex spatial variability: An A* approach. In Proceedings of the 2005 IEEE International Conference on Robotics and Automation, Barcelona, Spain, 18–22 April 2005; pp. 194–198. [CrossRef]
18. Farhi, E.; Goldstone, J.; Gutmann, S. A quantum approximate optimization algorithm. *arXiv* **2014**, arXiv:1411.4028. [CrossRef]
19. Lucas, A. Ising formulations of many NP problems. *Front. Phys.* **2014**, *2*, 5. [CrossRef]
20. Glover, F.; Kochenberger, G.; Du, Y. A tutorial on formulating and using QUBO models. *arXiv* **2018**, arXiv:1811.11538. [CrossRef]
21. Hadfield, S.; Wang, Z.; O'Gorman, B.; Rieffel, E.G.; Venturelli, D.; Biswas, R. From the quantum approximate optimization algorithm to a quantum alternating operator ansatz. *Algorithms* **2019**, *12*, 34. [CrossRef]
22. Bilal, M.; Zuberi, H.H.; Jaffar, A.; Riaz, W.; Khan, M.A.; Alharbi, A.; Miyajan, A.; Liu, S. Covert underwater communication through cepstrum modulation mimicking *Pseudorca crassidens* whistles using machine learning. *Sci. Rep.* **2026**, *16*, 4155. [CrossRef] [PubMed]
23. Qiao, G.; Bilal, M.; Liu, S.; Babar, Z.; Ma, T. Biologically inspired covert underwater acoustic communication—A review. *Phys. Commun.* **2018**, *30*, 107–114. [CrossRef]
24. Kloeden, P.E.; Platen, E. *Numerical Solution of Stochastic Differential Equations*; Springer: Berlin, Germany, 1992. [CrossRef]
25. Higham, D.J. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* **2001**, *43*, 525–546. [CrossRef]
26. Øksendal, B. *Stochastic Differential Equations: An Introduction with Applications*, 6th ed.; Springer: Berlin, Germany, 2003. [CrossRef]
27. Mao, X. *Stochastic Differential Equations and Applications*, 2nd ed.; Horwood: Chichester, UK, 2007. [CrossRef]
28. Fossen, T.I. *Guidance and Control of Ocean Vehicles*; Wiley: Chichester, UK, 1994.
29. Srinivas, N.; Krause, A.; Kakade, S.M.; Seeger, M. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv* **2010**, arXiv:0912.3995. [CrossRef]
30. Golden, B.L.; Levy, L.; Vohra, R. The orienteering problem. *Nav. Res. Logist.* **1987**, *34*, 307–318. [CrossRef]
31. Chao, I.-M.; Golden, B.L.; Wasil, E.A. The team orienteering problem. *Eur. J. Oper. Res.* **1996**, *88*, 464–474. [CrossRef]
32. Vansteenwegen, P.; Souffriau, W.; Van Oudheusden, D. The orienteering problem: A survey. *Eur. J. Oper. Res.* **2011**, *209*, 1–10. [CrossRef]
33. Gunawan, A.; Lau, H.C.; Vansteenwegen, P. Orienteering problem: A survey of recent variants, solution approaches and applications. *Eur. J. Oper. Res.* **2016**, *255*, 315–332. [CrossRef]
34. Zhou, L.; Wang, S.-T.; Choi, S.; Pichler, H.; Lukin, M.D. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Phys. Rev. X* **2020**, *10*, 021067. [CrossRef]
35. Egger, D.J.; Mareček, J.; Woerner, S. Warm-starting quantum optimization. *Quantum* **2021**, *5*, 479. [CrossRef]
36. Mayne, D.Q.; Rawlings, J.B.; Rao, C.V.; Scokaert, P.O.M. Constrained model predictive control: Stability and optimality. *Automatica* **2000**, *36*, 789–814. [CrossRef]
37. Rawlings, J.B.; Mayne, D.Q.; Diehl, M.M. *Model Predictive Control: Theory, Computation, and Design*, 2nd ed.; Nob Hill: Madison, WI, USA, 2017.
38. Camacho, E.F.; Bordons, C. *Model Predictive Control*, 2nd ed.; Springer: London, UK, 2007. [CrossRef]
39. Okubo, A. Oceanic diffusion diagrams. *Deep-Sea Res. Oceanogr. Abstr.* **1971**, *18*, 789–802. [CrossRef]
40. Aleksandrowicz, G.; Aleksandrowicz, G.; Alexander, T.; Barkoutsos, P.; Bello, L.; Ben-Haim, Y.; Bucher, D.; Cabrera-Hernández, F.J.; Carballo-Franquis, J.; Chen, A.; et al. *Qiskit: An Open-Source Framework for Quantum Computing*; Zenodo: Geneva, Switzerland, 2019. [CrossRef]
41. Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680. [CrossRef] [PubMed]
42. Holland, J.H. *Adaptation in Natural and Artificial Systems*; University of Michigan Press: Ann Arbor, MI, USA, 1975.

43. Cerezo, M.; Arrasmith, A.; Babbush, R.; Benjamin, S.C.; Endo, S.; Fujii, K.; McClean, J.R.; Mitarai, K.; Yuan, X.; Cincio, L.; et al. Variational quantum algorithms. *Nat. Rev. Phys.* **2021**, *3*, 625–644. [CrossRef]
44. Bharti, K.; Cervera-Lierta, A.; Kyaw, T.H.; Haug, T.; Alperin-Lea, S.; Anand, A.; Degroote, M.; Heimonen, H.; Kottmann, J.S.; Menke, T.; et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **2022**, *94*, 015004. [CrossRef]
45. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2018**, *2*, 79. [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.