Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Minimizing Information Age and Sensor Energy in Ocean Monitoring Networks through Hybrid Acoustic and Optical Data Harvesting by Cooperative Underwater Vehicles

S. Marisargunam
Saveetha Engineering College
G. Arul Dalton
aruldalton@saveetha.ac.in

Saveetha Engineering College

Research Article

Keywords: Ocean monitoring · Underwater sensor networks, Age of information, Hybrid acoustic and optical communication, Autonomous underwater vehicles, Evolution strategies.

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11426077/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Minimizing Information Age and Sensor Energy in Ocean Monitoring Networks through Hybrid Acoustic and Optical Data Harvesting by Cooperative Underwater Vehicles

Marisargunam S[1], Arul Dalton G*[2]

[1] Saveetha Engineering College, Department of Electrical and Electronics Engineering, Chennai, India.
ORCID (Marisargunam S): 0009-0008-5157-2048, Email: marisargunams@saveetha.ac.in

[2] Saveetha Engineering College, Department of Computer Science and Engineering, Chennai, India
ORCID (Arul Dalton G): 0000-0002-5981-7176

Correspondence: Arul Dalton G (aruldalton@saveetha.ac.in)

Abstract

Ocean monitoring networks must deliver fresh measurements from battery-powered seabed sensors, yet underwater links impose a hard trade-off. Acoustic links reach kilometres but are slow and energy-hungry, whereas optical links are fast and energy-cheap but reach only metres. Autonomous underwater vehicles (AUVs) can harvest data by approaching sensors, but ocean currents perturb their motion. This paper proposes AquaFresh, a cooperative policy that decides, for every AUV and time slot, which sensor to serve and whether to use an acoustic or an optical link, minimising priority-weighted age of information (AoI) and sensor energy. AquaFresh combines Lyapunov drift-plus-penalty scores with a graph attention residual shared by all vehicles and is trained by antithetic evolution strategies on the team objective. The simulation model includes Thorp absorption, Wenz ambient noise, Rayleigh fading, co-channel acoustic interference, residual Doppler, water-type-dependent optical attenuation and a time-varying double-gyre current field. On held-out ocean realisations, AquaFresh reduced weighted AoI by 12.5% and sensor energy by 28.8% relative to a Lyapunov max-weight scheduler, and AoI by 28.0% relative to an AoI-greedy hybrid policy at equal energy. It retained the best AoI in zero-shot tests across water types, current strengths, 40 to 100 sensors and 2 to 5 vehicles, and an energy-weight sweep traced an AoI and energy front that dominates all baselines. Ablations indicate that end-to-end optimisation of the energy-aware objective drives most of the gain.

Keywords

Ocean monitoring · Underwater sensor networks, Age of information, Hybrid acoustic and optical communication, Autonomous underwater vehicles, Evolution strategies.

1 Introduction

Ocean monitoring—tracking temperature, salinity, dissolved oxygen, pollutants and harmful algal blooms—relies increasingly on seabed sensor networks that form the Internet of Underwater Things (IoUT) (Akyildiz et al. 2005; Jahanbakht et al. 2021). The value of such data decays quickly: an early-warning system for an oil plume or a hypoxic event needs fresh readings, not merely many readings. Age of Information (AoI) (Kaul et al. 2012; Yates et al. 2021), the time elapsed since the newest delivered sample was generated, is therefore a more appropriate figure of merit than throughput.

Underwater links make freshness hard to achieve. Acoustic links reach kilometres, but they offer only a few kbit/s, suffer frequency-dependent absorption, ambient noise, multipath fading and Doppler (Stojanovic 2007; Stojanovic and Preisig 2009), and consume substantial transmit energy from battery-limited sensors. Underwater optical wireless communication (UOWC) offers Mb/s rates at a tiny energy cost per bit, but only over metres to tens of metres, depending strongly on water turbidity (Zeng et al. 2017; Kaushal and Kaddoum 2016). Autonomous underwater vehicles (AUVs) can bridge the gap by travelling to sensors (Hollinger et al. 2012), but travel is slow and is pushed around by ocean currents. Each harvesting decision therefore couples three questions: which sensor to serve, by which link, and by which AUV.

Existing work typically fixes one of these dimensions: AUV tours with optical or acoustic links only (Farr et al. 2010; Hollinger et al. 2012), AoI scheduling on static single-hop links (Kadota et al. 2018), or multi-agent

reinforcement learning (MARL) for trajectory planning that ignores link-modality choice and currents (Yu et al. 2022; Lowe et al. 2017). In our experiments, standard cooperative MARL (MAPPO) failed to improve on a Lyapunov max-weight scheduler. The cause is poor credit assignment: several AUVs share a team reward whose response to any one decision is delayed and noisy.

This paper proposes **AquaFresh**, a freshness-aware, energy-aware policy for multi-AUV hybrid acoustic–optical data harvesting. The contributions are:

(1) A joint formulation of multi-AUV target selection and per-link modality choice (acoustic now, or optical after approach), minimising priority-weighted AoI and sensor energy. The model includes a time-varying double-gyre current field, Thorp absorption, Wenz noise, Rayleigh fading, co-channel acoustic interference, residual Doppler inter-carrier interference and Jerlov-type-dependent optical attenuation.

(2) A Lyapunov-prior-guided graph-attention policy. Each AUV scores a candidate set of sensor–modality pairs; its logits are the drift-plus-penalty scores of a Lyapunov scheduler (Neely 2010) plus a learned residual. The residual is produced by graph attention (Veličković et al. 2018) over the AUV's candidates and the other AUVs' states and intentions, which lets the AUVs learn to divide the sensor field among themselves.

(3) Training by antithetic evolution strategies (ES) (Salimans et al. 2017) on the episode-level objective with common random numbers. This avoids the multi-agent credit-assignment problem that stalled policy-gradient MARL in this setting.

(4) An experimental study over held-out ocean realisations with five training seeds, four ablations, an energy-weight sweep yielding an AoI–energy Pareto front, and zero-shot generalisation to unseen water types, current strengths, network sizes and AUV fleet sizes.

2 Related Work

Underwater channels. Acoustic propagation loss, absorption and noise are classically modelled by the spreading law with Thorp absorption (Urick 1983) and the Wenz ambient-noise spectra (Wenz 1962). Stojanovic related capacity to distance and showed an optimal distance-dependent operating frequency (Stojanovic 2007; Stojanovic and Preisig 2009). UOWC surveys (Zeng et al. 2017; Kaushal and Kaddoum 2016) document the Beer–Lambert attenuation across Jerlov water types (Jerlov 1976) and the resulting metre-scale ranges in coastal and turbid waters. Hybrid acoustic–optical modems have been demonstrated experimentally (Farr et al. 2010).

AUV-assisted data collection. Hollinger et al. (2012) formulated AUV tours that trade travel against acoustic transmission. Subsequent work optimised tours for throughput or energy but mostly assumed a single modality, static water and a single AUV. Routing in underwater sensor networks is surveyed in (Han et al. 2015). AUV-assisted harvesting presupposes accurate knowledge of node positions, which can be obtained by mobile localisation; for example, Hema et al. (2025) estimated the depth of underwater sensor nodes using mobile localisation techniques.

Age of Information and learning. AoI was introduced in (Kaul et al. 2012) and surveyed in (Yates et al. 2021). Max-weight and Whittle-index scheduling policies are near-optimal for single-hop broadcast networks (Kadota et al. 2018). MARL methods such as MADDPG (Lowe et al. 2017), MAPPO (Yu et al. 2022) and counterfactual baselines (Foerster et al. 2018) address multi-robot coordination but suffer from high variance with shared rewards. ES (Salimans et al. 2017) optimises episode return directly and is robust to delayed, shared rewards. Attention-based temporal graph convolutional networks have also proved effective for spatiotemporal ocean prediction, such as sea-level-rise forecasting (Swathy et al. 2025), which supports using graph attention to capture spatial interactions in ocean systems. Table 1 positions AquaFresh.

Table 1 Positioning of AquaFresh (✓ addressed, – not addressed; ***** only through its score design)

Work	AoI objective	Hybrid ac./opt.	Multi-AUV	Ocean currents	Energy-aware	Learned coordination
Hollinger et al. (2012)	–	–	–	–	✓	–
Farr et al. (2010)	–	✓	–	–	–	–
Kadota et al. (2018)	✓	–	–	–	–	–
MAPPO / MADDPG (Yu et al. 2022; Lowe et al. 2017)	–	–	✓	–	–	✓

Lyapunov max-weight (Neely 2010)	✓	✓*	–	–	✓	–
AquaFresh (this work)	✓	✓	✓	✓	✓	✓

3 System Model

Fig. 1 summarises the system in three layers. (a) The ocean-monitoring field comprises priority-weighted seabed sensors, a time-varying current field and an AUV fleet that relays data to a surface buoy. (b) The hybrid underwater channel offers each sensor–AUV pair either a long-range, low-rate, energy-costly acoustic link or a short-range, high-rate, energy-cheap optical link; each delivery resets the sensor's AoI and consumes its energy. (c) The AquaFresh decision engine, run by every AUV each slot, turns observations into Lyapunov-scored candidates, refines them with a shared graph-attention residual and executes an acoustic transfer, an optical approach or a hold. The resulting AoI and energy define the objective, which offline evolution-strategies training optimises. The following subsections formalise each block.

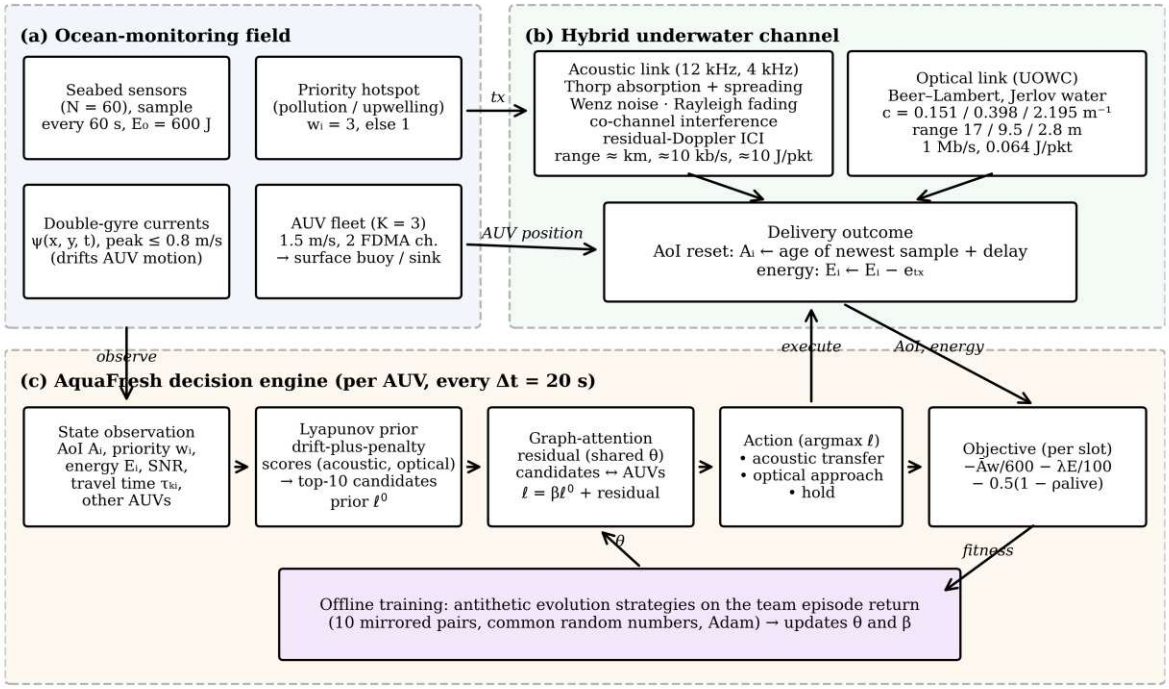


Fig. 1 Block diagram of the AquaFresh system: (a) ocean-monitoring field, (b) hybrid acoustic–optical underwater channel and (c) the per-AUV decision engine with its offline evolution-strategies training loop

A 1.5 km × 1.5 km ocean-monitoring field at 100 m depth contains N seabed sensors (N = 60 by default) and K AUVs (K = 3) cruising at 1.5 m/s (Fig. 2). Sensors inside a 400 m-radius hotspot (e.g., an upwelling or pollution plume) carry priority weight w_i = 3, others w_i = 1. Each sensor samples every T_s = 60 s and keeps only its newest sample. Time is slotted with Δt = 20 s, and an episode lasts 1 h (180 slots).

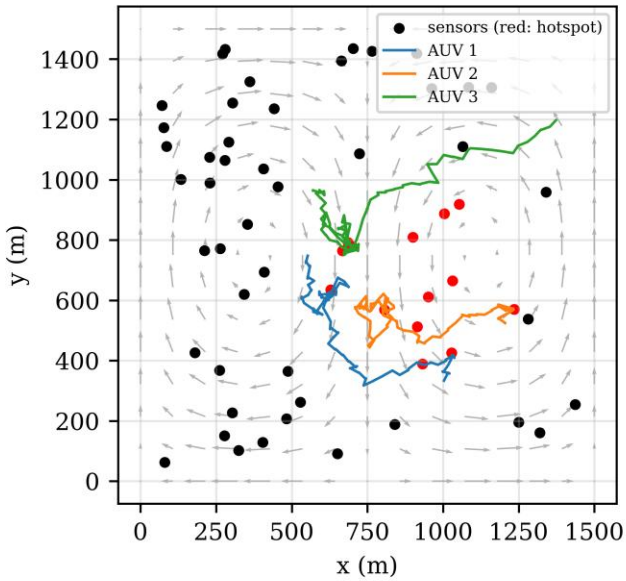


Fig. 2 Example ocean-monitoring field: seabed sensors (red: high-priority hotspot), the double-gyre current field (grey arrows) and the AUV trajectories produced by AquaFresh over one hour

3.1 Ocean current field

Currents follow the time-periodic double-gyre flow (Shadden et al. 2005), widely used as a benchmark for oceanic transport, scaled to the field with peak speed U_c (0.4 m/s by default, i.e. 27% of AUV speed):

$$\psi(x,y,t) = (U_c/\pi)\,\sin(\pi f(x,t))\,\sin(\pi y), \quad f = \varepsilon \sin(\omega t)\,x^2 + (1 - 2\varepsilon \sin(\omega t))\,x \tag{1}$$

with $u = -\partial\psi/\partial y$, $v = \partial\psi/\partial x$, $\varepsilon = 0.25$ and a period of 1 h. AUV motion is the commanded velocity plus the local current, so reaching a sensor takes an effective time $\tau_ki = d_ki / \max(v_AUV + c\cdot\hat{e}_ki, 0.3)$, which is shorter downstream and longer upstream.

3.2 Acoustic link

At carrier f_c = 12 kHz with bandwidth B = 4 kHz, the transmission loss over range d (m) combines spherical spreading and Thorp absorption α(f) (Urick 1983):

$$TL(d) = 20\log_{10} d + \alpha(f)\,d/1000, \quad \alpha(f) = 0.11f^2/(1+f^2) + 44f^2/(4100+f^2) + 2.75\cdot10^{-4}f^2 + 0.003 \tag{2}$$

The noise power spectral density is the Wenz model (Wenz 1962) (turbulence, shipping activity 0.7, wind 10 m/s, thermal), integrated over B. The source level is SL = 170.8 + 10 $\log_{10}$(η_t P_ac) dB re µPa at 1 m, with P_ac = 1 W and transducer efficiency η_t = 0.2. The instantaneous SINR at AUV k for sensor i is

$$\gamma_ki = SNR_ki \cdot h / (1 + \Sigma_{j\in\mathcal{J}_k} SNR_kj \cdot h_j), \quad h, h_j \sim Exp(1) \tag{3}$$

where $\mathcal{J}_k$ is the set of other simultaneous acoustic transmissions on the same FDMA channel (two channels; AUV k uses channel k mod 2). Residual Doppler after resampling (10% of v·f_c/c) causes inter-carrier interference with power $(\pi\, f_d/\Delta f)^2/3$ relative to the signal for subcarrier spacing Δf = 10 Hz, which caps γ. The rate is R = η B $\log_2(1 + \gamma)$ with η = 0.5. A 64 kbit packet succeeds if it fits into the slot and γ ≥ 8 dB. The sensor spends (P_ac + P_circ)·min(L/R, Δt) J, with P_circ = 0.5 W, whether or not the packet succeeds. An AUV approaching for an acoustic transfer stops once the mean SNR reaches 20 dB and receives while hovering, which reduces Doppler.

3.3 Optical link

The optical channel gain follows Beer–Lambert attenuation with beam-spreading loss, G(d) = e^(−c d) $(a_r/d)^2$ G_0, where the beam attenuation coefficient c is 0.151, 0.398 or 2.195 m^{-1} for clear-ocean, coastal and turbid-harbour Jerlov-type waters (Zeng et al. 2017; Jerlov 1976). With a −45 dB link-budget threshold, the usable range is 17.2, 9.5 and 2.8 m respectively. An optical transfer at 1 Mb/s and 1 W costs only 0.064 J per packet, but the AUV must first travel to within range.

3.4 Age of Information and energy

The AoI of sensor i increases by Δt every slot. On successful delivery at time t it resets to the age of the sensor's newest sample plus a fixed forwarding delay (2 s) and, for acoustic links, the propagation delay. Sensors start with E_0 = 600 J and die when depleted.

4 Problem Formulation

In each slot, every AUV k selects an action a_k ∈ {(i, acoustic), (i, optical), hold}, where i ranges over the sensors. The team objective is

$$\max \; E\Big[\, \Sigma_t \big(- \bar{A}_w(t)/600 - 0.5\,(1 - \rho_{alive}(t)) - \lambda\, E_{used}(t)/100 \big) \Big], \quad \bar{A}_w = \Sigma_i w_i A_i \,/\, \Sigma_i w_i \tag{4}$$

where Ā_w is the priority-weighted mean AoI, ρ_alive the fraction of live sensors, E_used the sensor energy spent in the slot and λ the energy weight (λ = 0.5 by default; λ ∈ {0, 0.5, 2} traces the AoI–energy trade-off). Joint action spaces grow as (2N+1)^K, and decisions interact through AUV conflicts and acoustic interference, which makes centralised exact optimisation intractable.

5 The AquaFresh Policy

5.1 Lyapunov drift-plus-penalty prior

Treating each A_i as a virtual queue, a one-slot drift-plus-penalty argument (Neely 2010) gives, for AUV k, sensor i and modality m, the score

$$S_{ki}^{ac} = w_i A_i p_{ki} - V \bar{e}_{ac} / E_i, \quad S_{ki}^{op} = w_i A_i e^{(-\tau_{ki}/300)} - V \bar{e}_{op} / E_i \tag{5}$$

where p_ki = exp(−γ_th / γ̄_ki) is the Rayleigh-fading success probability, τ_ki is the current-aware travel time, ē_ac = 20 J and ē_op = 0.1 J are nominal energy costs, and V = 400. The baseline Lyapunov scheduler picks the maximum score sequentially over the AUVs. In AquaFresh, the top-M = 10 sensors by max(S^ac, S^op) form each AUV's candidate set; the current target is pinned to slot 0 so that multi-slot optical approaches can be committed to. The scores, normalised per AUV, form a 2M+1 prior-logit vector ℓ^{0}_k (hold receives −6).

5.2 Graph-attention residual

Each candidate j of AUV k is described by a 12-dimensional feature vector x_kj: relative position, distance, AoI, priority, residual energy, mean acoustic SNR, current-aware travel time, a flag for being targeted by another AUV, distance to the nearest other AUV, current velocity, and a pinned-target flag. Each AUV node q is described by its position, current target offset and local current (u_q). With two-layer encoders φ_c, φ_a,

$$h_{kj} = \varphi_c(x_{kj}), \quad n_{kq} = \varphi_a(u_q), \quad z_{kj} = h_{kj} + \Sigma_{v \in \mathcal{N}_k} \mathrm{softmax}_v\big((W_Q h_{kj})^{\top} (W_K v)/\sqrt{H} \big)\, W_V v \tag{6}$$

where $\mathcal{N}_k$ = {h_k1,…,h_kM} ∪ {n_k1,…,n_kK}, i.e. each candidate attends jointly to the AUV's other candidates and to all AUVs. The final logits are

$$\ell_k = \beta\, \ell^{0}_k + [\, z_k \cdot w_{ac}\,,\; z_k \cdot w_{op}\,,\; \mathrm{mean}_j(z_{kj}) \cdot w_{hold} \,] \tag{7}$$

and each AUV executes argmax ℓ_k, with a learnable prior weight β (initialised to 3). Parameters are shared across AUVs, so the policy applies unchanged to any N and K. With H = 16 it has about $1.5 \cdot 10^3$ parameters.

5.3 Training by evolution strategies

Let F(θ) be the mean return (4) of the deterministic policy over two ocean realisations drawn afresh each generation and shared by all perturbations (common random numbers). Antithetic ES (Salimans et al. 2017) estimates the gradient from P = 10 mirrored perturbation pairs with fitness-rank shaping:

$$\hat{g} = \frac{1}{2P\sigma} \sum_{p=1}^{P} \big[\, u(F(\theta + \sigma\varepsilon_p)) - u(F(\theta - \sigma\varepsilon_p)) \,\big]\, \varepsilon_p, \quad \varepsilon_p \sim \mathcal{N}(0, I) \tag{8}$$

where u(·) maps fitness to centred ranks in [−0.5, 0.5]. θ is updated by Adam (step 0.03, weight decay 0.005) with σ = 0.05 for 50 generations (2000 episodes). Because the fitness is the team's episode return, no per-agent credit assignment is required.

Algorithm 1 AquaFresh training and deployment

1: Initialise θ (graph-attention residual ≈ 0, β = 3) so that the initial policy follows the Lyapunov prior.
2: for generation g = 1 … 50 do
3: draw two ocean realisations (sensor layout, hotspot, initial AUV positions, AoI, fading seeds); sample ε_1 … ε_10 ~ $\mathcal{N}$(0, I).

4: for each p and sign s ∈ {+1, −1}: run the deterministic policy θ + sσε_p on both realisations; F_p^s ← mean return (4).
5: rank-shape the fitness values; form ĝ by (8); θ ← Adam(θ, ĝ).
6: end for
7: Deployment, every Δt on each AUV: build candidates from Lyapunov scores (5) → compute prior ℓ^0 → graph attention (6) → a_k = argmax ℓ_k (7) → execute the approach, acoustic transfer or optical transfer.

6 Experimental Set-up

All models were implemented in Python/NumPy with JAX for the policy network and run on a single CPU core. Table 2 lists the parameters. Every controller is evaluated on the same 10 held-out ocean realisations (seeds 100–109), never seen in training. Learned policies are reported as mean ± standard deviation over independently trained seeds (5 for AquaFresh; 3 for each ablation; 2 per λ value).

Table 2 Simulation and training parameters

Parameter	Symbol	Value
Field / depth	—	1.5 km × 1.5 km / 100 m
Sensors / AUVs	N / K	60 / 3 (sensitivity: 40–100 / 2–5)
Hotspot priority / radius	w_i	3 / 400 m
Sampling period / slot / horizon	T_s / Δt / T	60 s / 20 s / 1 h
AUV speed	v_AUV	1.5 m/s
Current peak / model	U_c	0.4 m/s (0–0.8) / double gyre
Acoustic carrier / bandwidth	f_c / B	12 kHz / 4 kHz, 2 FDMA channels
Acoustic Tx / circuit power / efficiency	P_ac / P_circ / η_t	1 W / 0.5 W / 0.2
Spreading / noise	—	spherical + Thorp / Wenz (ship 0.7, wind 10 m/s)
Fading / SINR threshold	—	Rayleigh / 8 dB
Residual Doppler / subcarrier spacing	—	10% / 10 Hz
Packet size	L	64 kbit
Optical rate / power / threshold	R_op / P_op	1 Mb/s / 1 W / −45 dB
Beam attenuation c	—	0.151 / 0.398 / 2.195 m^{-1} (clear / coastal / turbid)
Sensor battery	E_0	600 J
Energy weight	λ	0.5 (sweep 0, 0.5, 2)
Lyapunov weight / candidates	V / M	400 / 10
Policy hidden size / parameters	H	16 / ≈$1.5 \cdot 10^3$
ES pairs / σ / step / generations	P / σ / η	10 / 0.05 / 0.03 / 50
Episodes per fitness	—	2 (common random numbers)

Baselines. Four baselines are compared. (i) Greedy-NN (acoustic): each AUV serves the nearest sensor whose AoI exceeds T_s, acoustically. (ii) TSP tour (optical): k-means partitioning, a nearest-neighbour tour improved by 2-opt, and optical collection at each sensor. (iii) AoI-greedy hybrid: maximise w_i(A_i + τ_ki)/(τ_ki + Δt), using optical if the sensor is within two slots of travel or its battery is below 30%. (iv) Lyapunov max-weight: sequential argmax of (5). The untrained AquaFresh policy (prior only, parallel decisions) is also reported. A cooperative-MARL baseline (MAPPO with the same network and prior, 200 iterations) was trained but did not improve on its initialisation; it is discussed in Section VIII.

Metrics. Priority-weighted mean AoI over the hour, mean peak AoI at delivery, number of delivered packets, share of optical deliveries, failed transmissions, sensor energy per sensor per hour, energy per delivered packet, total AUV travel distance and episode return (4).

7 Results

7.1 Training behaviour

Fig. 3 shows the mean ES fitness over three training seeds. Starting from the prior-only policy, the fitness rises steeply during the first 20–30 generations and then plateaus, overtaking the Lyapunov and AoI-greedy baselines early. For comparison, a cooperative MAPPO agent with the same network and prior, trained for 200 iterations (≈800 episodes), did not improve on its initial return; Section VIII discusses why.

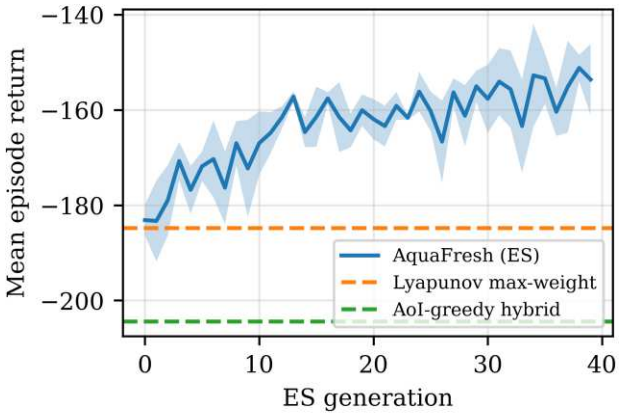


Fig. 3 ES training curve of AquaFresh (mean ± std of population fitness over 3 seeds) against the episode return of the two strongest baselines

7.2 Comparison with baselines

Table 3 Performance on 10 held-out ocean realisations (AquaFresh: mean ± std over 3 training seeds)

Policy	Weighted AoI (s)	Peak AoI (s)	Delivered pkts	Optical share	Failed tx	Energy (J/sensor/h)	Energy per pkt (J)	AUV distance (km)	Return
Greedy-NN (acoustic)	1713	162	395	0.00	138	114.6	18.8	0.0	-551.8
TSP tour (optical)	1154	1999	80	1.00	0	0.1	0.1	15.0	-346.3
AoI-greedy hybrid	587	534	362	0.25	125	94.5	15.8	3.2	-204.4
Lyapunov max-weight	483	845	244	0.15	207	131.4	32.7	8.2	-184.8
Prior only (untrained)	579	1010	190	0.10	18	33.6	10.6	7.4	-183.7
AquaFresh	423 ± 18	768 ± 17	271 ± 7	0.07 ± 0.02	122 ± 12	93.5 ± 6.9	20.9 ± 0.8	5.1 ± 0.4	-154.8 ± 3.7

Table 3 reports the 10 held-out realisations. AquaFresh achieved a weighted mean AoI of 423 ± 18 s, against 483 s for the Lyapunov scheduler and 587 s for AoI-greedy, i.e. 12.5% and 28.0% lower. Sensor energy was 93.5 J per sensor per hour, 29% below Lyapunov and on a par with AoI-greedy (94.5 J), so AquaFresh improves freshness without paying for it in energy. Failed acoustic transmissions fell from 207 to 122. The episode return improved from -184.8 to -154.8 ± 3.7.

The single-modality baselines show why joint decisions matter. The acoustic-only greedy policy never moves, because every sensor is within acoustic range, so it starves distant sensors (AoI 1713 s) and spends 115 J per sensor. The optical TSP tour spends almost no sensor energy (0.09 J) but, needing 15.0 km of AUV travel, delivers only 80 packets and its AoI grows to 1154 s (Fig. 4). A limitation: AquaFresh uses optical links for only 7% of deliveries, below AoI-greedy (25%). Under the one-hour horizon and the default energy weight, the optimiser mostly prefers acoustic transfers from well-chosen positions.

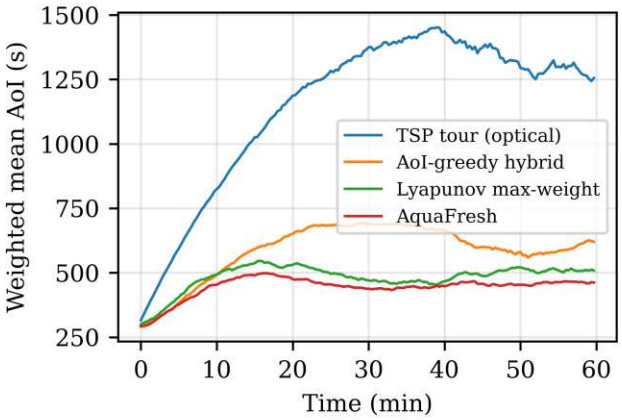


Fig. 4 Weighted mean AoI over the one-hour mission (average over the 10 held-out realisations)

7.3 AoI–energy trade-off

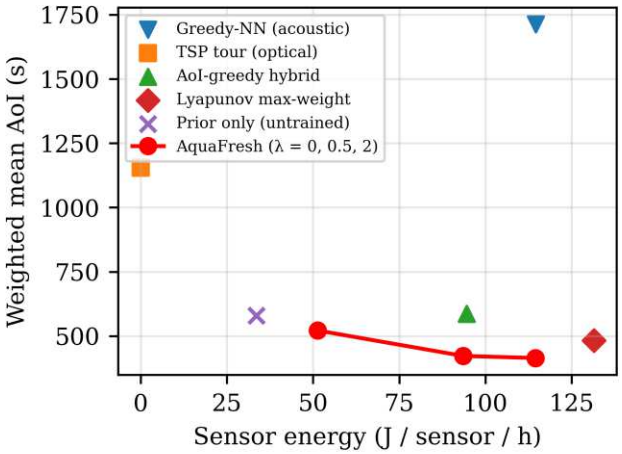


Fig. 5 AoI versus sensor energy. The AquaFresh curve connects policies trained with energy weights λ = 0, 0.5 and 2; baseline policies are single points

Training AquaFresh with λ = 0, 0.5 and 2 gives the operating points (114 J, 414 s), (94 J, 423 s) and (51 J, 522 s) in Fig. 5. The λ = 0 point dominates the Lyapunov scheduler on both axes. The λ = 0.5 point dominates AoI-greedy (lower AoI at equal energy). The λ = 2 point halves energy relative to AoI-greedy while still achieving lower AoI. The single weight λ therefore lets an operator choose the freshness–lifetime balance, which a fixed heuristic cannot provide. The λ = 0 and λ = 2 points each come from a single training seed.

7.4 Ablation study

Table 4 Ablations (2 seeds each) and energy-weight sweep (1 seed per λ)

Policy	Weighted AoI (s)	Peak AoI (s)	Delivered pkts	Optical share	Failed tx	Energy (J/sensor/h)	Energy per pkt (J)	AUV distance (km)	Return
AquaFresh	423 ± 18	768 ± 17	271 ± 7	0.07 ± 0.02	122 ± 12	93.5 ± 6.9	20.9 ± 0.8	5.1 ± 0.4	-154.8 ± 3.7
w/o graph attention	445 ± 20	793 ± 17	260 ± 9	0.10 ± 0.02	97 ± 7	79.4 ± 2.7	18.6 ± 1.2	5.8 ± 0.3	-157.4 ± 6.9
w/o Lyapunov prior	386 ± 6	723 ± 3	296 ± 1	0.00	174 ± 12	124.4 ± 6.1	25.3 ± 1.3	3.4 ± 0.2	-153.2 ± 3.5
w/o hybrid (acoustic only)	411 ± 19	749 ± 10	280 ± 7	0.00	150 ± 4	110.9 ± 2.9	23.8 ± 0.1	3.5 ± 0.1	-156.5 ± 4.7
w/o current awareness	426 ± 5	777 ± 3	270 ± 1	0.08 ± 0.02	120 ± 3	92.0 ± 1.8	20.7 ± 0.6	4.9 ± 0.3	-155.3 ± 2.0

AquaFresh (λ=0)	414	735	286	0.01	160	114.5	24.3	3.0	-158.7
AquaFresh (λ=2)	522	936	208	0.06	47	51.3	14.9	6.2	-171.9

Table 4 isolates each component. The results do not support a claim that every architectural component improves the objective. Without the Lyapunov prior (training from scratch by ES), AquaFresh reached a lower AoI (386 s) and a statistically indistinguishable return (-153.2 vs -154.8), but used 25% more energy and never used optical links. Restricting to acoustic only similarly gave AoI 411 s at 111 J (19% more energy) with a similar return. Removing current awareness changed nothing measurable (return -155.3), which suggests that, at 27% of AUV speed, the current mainly perturbs trajectories that are re-planned every slot anyway. Removing graph attention gave a slightly higher AoI (445 s) at lower energy (79 J), with a return within one standard deviation.

The ablation therefore attributes most of AquaFresh's advantage over the baselines to end-to-end optimisation of the energy-aware team objective by ES, rather than to any individual architectural element. The prior and the optical option move the learned policy towards lower energy at similar return, i.e. along the Pareto front, rather than improving the scalar objective. With 2–3 seeds per configuration, differences of a few return units are not statistically resolved.

7.5 Zero-shot generalisation

Table 5 Zero-shot generalisation (weighted AoI / sensor energy; 6 realisations; AquaFresh averaged over 3 seeds)

Setting	AoI-greedy hybrid	Lyapunov max-weight	AquaFresh
Water = clear	630 s / 91 J	510 s / 140 J	416 s / 87 J
Water = coastal	618 s / 96 J	467 s / 129 J	422 s / 87 J
Water = turbid	657 s / 83 J	492 s / 123 J	447 s / 80 J
Current (m/s) = 0.0	623 s / 96 J	483 s / 134 J	417 s / 93 J
Current (m/s) = 0.4	618 s / 96 J	467 s / 129 J	422 s / 87 J
Current (m/s) = 0.8	625 s / 96 J	515 s / 136 J	443 s / 80 J
Sensors N = 40	485 s / 117 J	362 s / 195 J	310 s / 113 J
Sensors N = 60	618 s / 96 J	467 s / 129 J	422 s / 87 J
Sensors N = 100	802 s / 62 J	767 s / 80 J	625 s / 55 J
AUVs K = 2	642 s / 40 J	426 s / 45 J	409 s / 41 J
AUVs K = 3	618 s / 96 J	467 s / 129 J	422 s / 87 J
AUVs K = 5	523 s / 150 J	810 s / 315 J	445 s / 182 J

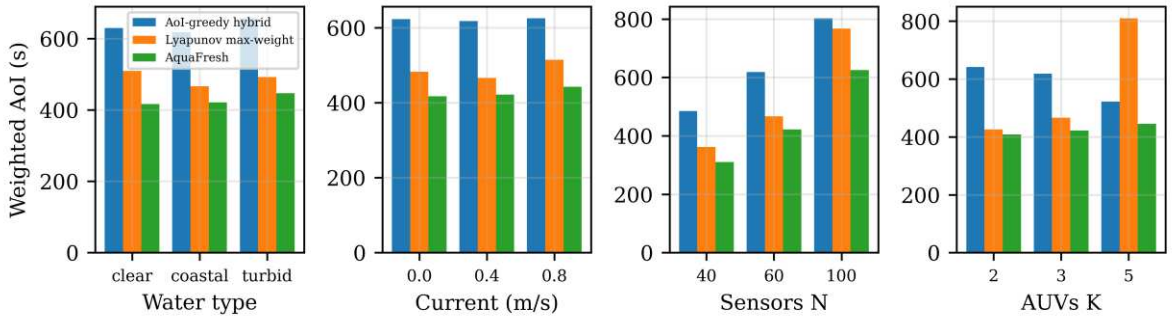


Fig. 6 Zero-shot AoI across water types, current strengths, network sizes and fleet sizes (policies trained only on coastal water, 0.4 m/s, N = 60, K = 3)

Table 5 and Fig. 6 evaluate the policies, without retraining, on unseen conditions. AquaFresh achieved the lowest AoI and best return in all twelve settings. Its AoI rose only from 416 to 447 s from clear to turbid water, and from 417 to 443 s as the current doubled to 0.8 m/s. The largest margin appeared with five AUVs, where the Lyapunov scheduler degraded to 810 s because of co-channel interference and conflicting targets, whereas AquaFresh reached 445 s. At N = 100, AquaFresh reached 625 s versus 767 s.

Table 6 Role of graph attention in zero-shot scaling (2 seeds each, 6 realisations)

Policy	N = 100: AoI (s)	N = 100: return	K = 5: AoI (s)	K = 5: return
AquaFresh	639	-219.1	452	-192.3
w/o graph attention	727	-239.8	446	-178.8

Table 6 asks whether graph attention drives this scaling. At N = 100 the attention policy clearly outperformed its attention-free ablation (AoI 639 vs 727 s; return -219.1 vs -239.8). With K = 5 AUVs, however, the attention-free policy was equal in AoI and better in return because it spent less energy. Attention thus helps generalisation to denser sensor fields, but not, in these experiments, to larger fleets.

7.6 Computational cost

A full decision for all three AUVs, including candidate construction, the prior and the network forward pass, took 0.92 ms on one CPU core. That is negligible compared with the 20 s slot and well within an AUV's onboard computing budget.

8 Discussion and Limitations

Why cooperative MARL failed and ES succeeded. With a shared team reward, the effect of one AUV's decision on the weighted AoI is delayed by several slots and mixed with the decisions of the others. In our MAPPO experiments the advantage estimates were dominated by this noise, and the policy drifted rather than improved. ES optimises the episode return of the deterministic joint policy directly. With common random numbers and antithetic sampling, its gradient estimate was reliable at the $\approx 1.5 \cdot 10^3$-parameter scale used here. Counterfactual baselines (Foerster et al. 2018) or difference rewards may rescue policy-gradient MARL and are worth investigating.

What the components contribute. The ablations indicate that the objective and the optimiser matter more than the architectural components in the studied setting. The Lyapunov prior and the optical option are most useful as levers towards energy efficiency. Current awareness had no measurable effect at a current speed of 27% of AUV speed, and graph attention helps mainly for larger sensor fields. Optical links became more valuable when energy was weighted more heavily; longer missions in which batteries actually deplete would likely raise their share.

Model fidelity. The acoustic channel uses standard analytic loss and noise models with Rayleigh fading rather than ray-traced multipath (e.g. Bellhop), and the optical link uses a simplified link budget without pointing errors. The ocean field is a double-gyre benchmark rather than reanalysis data. Seed counts are modest (3 for AquaFresh, 2 per ablation, 1 per λ). Validation with Bellhop channel traces, ocean-model current fields (e.g. HYCOM) and sea trials is needed before operational claims can be made.

9 Conclusion

This paper presented AquaFresh, an energy-aware, freshness-driven policy for multi-AUV hybrid acoustic–optical data harvesting in ocean-monitoring networks. It combines a Lyapunov drift-plus-penalty prior with a graph-attention residual trained by evolution strategies. On held-out ocean realisations it reduced weighted AoI by 12.5% and sensor energy by 28.8% relative to a Lyapunov max-weight scheduler, generalised zero-shot across water types, currents and network sizes, and exposed a tunable AoI–energy Pareto front. Ablations attribute the gains mainly to optimising the energy-aware team objective end to end, with the prior and the optical option shifting the policy towards energy efficiency and attention aiding scaling to denser fields. Future work will extend the horizon to multi-day missions with battery depletion, use Bellhop and HYCOM-driven environments, and explore credit-assignment-aware multi-agent RL.

Statements and Declarations

Funding No funding was received for conducting this study.

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval Not applicable. This study did not involve human participants or animals.

Data and code availability The simulation environment, training and evaluation code, trained policies and raw results that support the findings of this study are openly available at https://github.com/GArulDalton/AquaFresh.git under the MIT License.

Author contributions Marisargunam S and Arul Dalton G contributed to the study conception and design. Methodology, software, experiments and analysis were carried out by Marisargunam S and Arul Dalton G. The first draft of the manuscript was written by Marisargunam S and Arul Dalton G, and both authors read and approved the final manuscript.

References

Akyildiz IF, Pompili D, Melodia T (2005) Underwater acoustic sensor networks: research challenges. Ad Hoc Netw 3(3):257–279. https://doi.org/10.1016/j.adhoc.2005.01.004

Farr N, Bowen A, Ware J, Pontbriand C, Tivey M (2010) An integrated, underwater optical/acoustic communications system. In: Proceedings of OCEANS 2010 IEEE Sydney

Foerster J, Farquhar G, Afouras T, Nardelli N, Whiteson S (2018) Counterfactual multi-agent policy gradients. In: Proceedings of the AAAI Conference on Artificial Intelligence 32(1)

Han G, Jiang J, Bao N, Wan L, Guizani M (2015) Routing protocols for underwater wireless sensor networks. IEEE Commun Mag 53(11):72–78

Hema R, Diana DC, Nair RR, Poovannan E (2025) Depth estimation of nodes in underwater wireless sensor networks using mobile localization techniques. Franklin Open 13:100387. https://doi.org/10.1016/j.fraope.2025.100387

Hollinger GA, Choudhary S, Qarabaqi P, Murphy C, Mitra U, Sukhatme GS, Stojanovic M, Singh H, Hover F (2012) Underwater data collection using robotic sensor networks. IEEE J Sel Areas Commun 30(5):899–911

Jahanbakht M, Xiang W, Hanzo L, Rahimi Azghadi M (2021) Internet of underwater things and big marine data analytics: a comprehensive survey. IEEE Commun Surv Tutor 23(2):904–956

Jerlov NG (1976) Marine optics. Elsevier, Amsterdam Kadota I, Sinha A, Uysal-Biyikoglu E, Singh R, Modiano E (2018) Scheduling policies for minimizing age of information in broadcast wireless networks. IEEE/ACM Trans Netw 26(6):2637–2650

Kaul S, Yates R, Gruteser M (2012) Real-time status: how often should one update? In: Proceedings of IEEE INFOCOM, pp 2731–2735

Kaushal H, Kaddoum G (2016) Underwater optical wireless communication. IEEE Access 4:1518–1547

Lowe R, Wu Y, Tamar A, Harb J, Abbeel P, Mordatch I (2017) Multi-agent actor-critic for mixed cooperative-competitive environments. In: Advances in Neural Information Processing Systems 30

Neely MJ (2010) Stochastic network optimization with application to communication and queueing systems. Morgan & Claypool, San Rafael

Salimans T, Ho J, Chen X, Sidor S, Sutskever I (2017) Evolution strategies as a scalable alternative to reinforcement learning. arXiv:1703.03864. https://doi.org/10.48550/arXiv.1703.03864

Shadden SC, Lekien F, Marsden JE (2005) Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. Physica D 212(3–4):271–304. https://doi.org/10.1016/j.physd.2005.10.007

Stojanovic M (2007) On the relationship between capacity and distance in an underwater acoustic communication channel. ACM SIGMOBILE Mob Comput Commun Rev 11(4):34–43

Stojanovic M, Preisig J (2009) Underwater acoustic communication channels: propagation models and statistical characterization. IEEE Commun Mag 47(1):84–89

Swathy TM, Isabels KR, Rebecca AS, Rachapudi V, Baker El-Ebiary YA, Gorintla S, Muniyandy E (2025) Game theory-optimized attention-based temporal graph convolutional network for spatiotemporal forecasting of sea level rise. Int J Adv Comput Sci Appl 16(8):574–585. https://doi.org/10.14569/IJACSA.2025.0160857

Urick RJ (1983) Principles of underwater sound, 3rd edn. McGraw-Hill, New York

Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y (2018) Graph attention networks. In: Proceedings of the International Conference on Learning Representations (ICLR)

Wenz GM (1962) Acoustic ambient noise in the ocean: spectra and sources. J Acoust Soc Am 34(12):1936–1956. https://doi.org/10.1121/1.1909155

Yates RD, Sun Y, Brown DR, Kaul SK, Modiano E, Ulukus S (2021) Age of information: an introduction and survey. IEEE J Sel Areas Commun 39(5):1183–1210

Yu C, Velu A, Vinitsky E, Gao J, Wang Y, Bayen A, Wu Y (2022) The surprising effectiveness of PPO in cooperative multi-gent games. In: Advances in Neural Information Processing Systems 35, Datasets and Benchmarks Track

Zeng Z, Fu S, Zhang H, Dong Y, Cheng J (2017) A survey of underwater optical wireless communications. IEEE Commun Surv Tutor 19(1):204–238. https://doi.org/10.1109/COMST.2016.2618841