Sustainable Marine Structures
https://journals.nasspublishing.com/index.php/sms

ARTICLE

Designing Sustainable Marine Protected Area Networks: A Graph-Based Optimization Approach for Mediterranean Conservation

Khaled Mili [1*] **, *Majdi Argoubi*** [2]

[1] *Department of Quantitative Methods, College of Business Administration, King Faisal University, Al-Ahsa 31982, Saudi Arabia*

[2] *Department of Quantitative Methods, University of Sousse, Rue Khalifa Karoui, Sahloul, Sousse BP 526, Tunisia*

ABSTRACT

Marine Protected Area (MPA) networks face a persistent challenge: protecting biodiversity while maintaining ecological connectivity and minimizing economic costs. Traditional planning methods optimize single objectives or face computational barriers when scaling thousands of potential sites. We developed a computational framework combining graph-based site embeddings with constraint-driven optimization to identify cost-effective MPA networks at the basin scale. We analyzed 10,741 candidate sites across the Mediterranean Sea (30 km resolution) using 84,721 species observations, bathymetric data, and spatially explicit opportunity costs. The optimized 15% coverage network (1,611 sites, 1.45 million km^2) protects 99.2% of 120 target species including all threatened Mediterranean resident taxa, achieves network connectivity 4.4 times higher than random selection, and reduces costs 42% below random approaches (€1.69 billion versus €2.94 billion). The framework enforces ecological constraints—threatened species protection, shallow habitat representation, and regional balance—while completing each optimization scenario in 8–15 min. Network connectivity scales super linearly with coverage (power law exponent 1.17), indicating that well-connected networks deliver disproportionate conservation value. Results remain robust across 10–30% coverage targets, with consistent species representation and spatial patterns. The modular framework transfers to other marine regions through the substitution of region-specific data and enables rapid scenario

*CORRESPONDING AUTHOR:
Khaled Mili, Department of Quantitative Methods, College of Business Administration, King Faisal University, Al-Ahsa 31982, Saudi Arabia; Email: kmili@kfu.edu.sa

ARTICLE INFO
Received: 12 February 2026 | Revised: 12 March 2026 | Accepted: 17 March 2026 | Published Online: 7 April 2026
DOI: https://doi.org/10.36956/sms.v8i2.3143

CITATION
Mili, K., Argoubi, M., 2026. Designing Sustainable Marine Protected Area Networks: A Graph-Based Optimization Approach for Mediterranean Conservation. Sustainable Marine Structures. 8(2): 1–29. DOI: https://doi.org/10.36956/sms.v8i2.3143

testing for stakeholder negotiation. Systematic optimization identifies MPA configurations delivering superior conservation outcomes at substantially lower cost than alternative strategies, supporting evidence-based pathways toward international marine protection targets.

Keywords: Marine Protected Areas; Conservation Planning; Network Optimization; Biodiversity Conservation; Mediterranean Sea; Cost-Effectiveness; Ecological Connectivity

1. Introduction

Coastal ecosystems deliver essential services: carbon storage, shoreline protection, fisheries production, and biodiversity maintenance[1, 2]. These systems support national economies and human well-being globally. Yet marine environments face severe threats from overfishing, climate change, pollution, and habitat destruction[3, 4]. The pressures are intensifying. Marine Protected Areas (MPAs) represent one response—designated zones restricting human activities to conserve marine life and restore damaged ecosystems[5, 6]. But individual MPAs often fail. Small size and spatial isolation undermine their effectiveness[7].

Evidence increasingly shows that connected MPA networks outperform isolated reserves[8]. Ecological connectivity—through larval dispersal, species migration, and genetic exchange—maintains viable populations, builds resilience to environmental shocks, and ensures species persistence across generations[9, 10]. Large-scale networks demonstrate measurable benefits: biodiversity protection and spillover effects that enhance adjacent fisheries[11, 12]. The logic is straightforward: well-designed networks amplify conservation impact beyond what any single MPA achieves alone.

The Mediterranean Sea illustrates the challenge clearly. This biodiversity hotspot contains exceptional endemism—species found nowhere else on Earth[13, 14]. Yet it faces severe pressures: intensive fishing, coastal development, and accelerating climate change[15, 16]. Climate impacts include warming waters, ocean acidification, oxygen depletion, and species redistributions[17]. Mediterranean fisheries management requires integrated approaches that balance conservation goals with socioeconomic realities[18, 19]. Current MPA coverage stands at approximately 8%, with only 1.3% under strict protection[20]. Meeting international commitments—30% protection by 2030 under the Convention on Biological Diversity—demands systematic planning[21].

Existing optimization methods face limitations. Systematic conservation planning tools like Marxan[22, 23] and integer programming[24] provide rigorous frameworks but struggle with computational tractability at basin scales involving thousands of candidate sites. Distance-based connectivity metrics offer computational efficiency but rely on simplified dispersal assumptions[25]. Most frameworks optimize single objectives—area or species representation—using weighted-sum approaches that limit exploration of trade-offs between competing goals[26, 27]. Many studies neglect mandatory ecological constraints: threatened species protection, habitat representation, and regional equity. These constraints are essential for practical conservation policy.

Graph-based approaches offer a natural solution. Treating MPAs as nodes and ecological connections as edges, graph theory provides rigorous methods to measure network properties: connectivity, centrality, robustness[28, 29]. Recent guidance emphasizes incorporating functional connectivity into MPA network design[30, 31]. Validation studies show that distance-based metrics weighted by dominant circulation patterns capture essential connectivity structure when computational efficiency is required[32, 33].

Graph Neural Networks (GNN) enable a computational advance. These architectures learn compact site representations encoding both local environmental features and global network topology[34]. Recent applications demonstrate effectiveness: integrating complex landscape patterns into species distribution models[35] and using deep reinforcement learning for connectivity conservation planning[36]. Unlike traditional graph metrics requiring explicit computation of all pairwise connections, GNNs embed sites into low-dimensional spaces

where connectivity patterns emerge from learned features. This reduces computational burden while maintaining ecological realism.

We integrate Graph Neural Networks with constraint-based optimization to design cost-effective MPA networks balancing biodiversity protection with economic feasibility. Our approach has four components. First, we construct a spatial connectivity graph using distance-based dispersal kernels weighted by dominant Mediterranean circulation patterns, representing potential ecological connections across 10,741 candidate sites. Second, we apply GNNs to learn compact 16-dimensional site embeddings encoding biodiversity value, cost-efficiency, and topological position within the connectivity network. Third, we formulate constraint-based optimization with mandatory ecological requirements—threatened species protection, shallow habitat representation, biogeographic balance—then apply a greedy algorithm maximizing marginal efficiency (species and connectivity gains per unit cost). Fourth, we demonstrate this framework using 84,721 species observations from the Mediterranean and compare optimized networks against alternative strategies.

This work makes four contributions. First, we demonstrate that compact site representations learned by GNNs can inform large-scale optimization (10,000+ sites) to achieve near-complete biodiversity coverage while reducing costs by 42–58% compared to alternative strategies. Second, we provide a constraint-based framework ensuring ecological balance—threatened species protection, habitat representation, biogeographic equity—alongside cost-effectiveness. Third, our computational approach scales to basin-wide spatial domains, completing optimization in 8–15 min per scenario through efficient GNN-based embeddings and greedy algorithms. Finally, we show how this framework can support practical conservation planning in the Mediterranean, where MPA network expansion is urgently needed to meet international targets.

2. Materials and Methods

Our computational framework integrates four sequential components to identify cost-effective MPA networks at the basin scale (**Figure 1**). First, we construct a spatial connectivity graph representing potential ecological linkages among 10,741 candidate sites across the Mediterranean. Second, we apply a Graph Neural Network (GNN) to learn compact 16-dimensional site embeddings encoding biodiversity value, cost-efficiency, and topological position within the network. Third, we formulate a constraint-based optimization problem with mandatory ecological requirements and solve it using a greedy marginal-efficiency algorithm. Fourth, we evaluate performance against three benchmark strategies across coverage scenarios of 10%, 15%, and 30%. Each component is described in detail below.

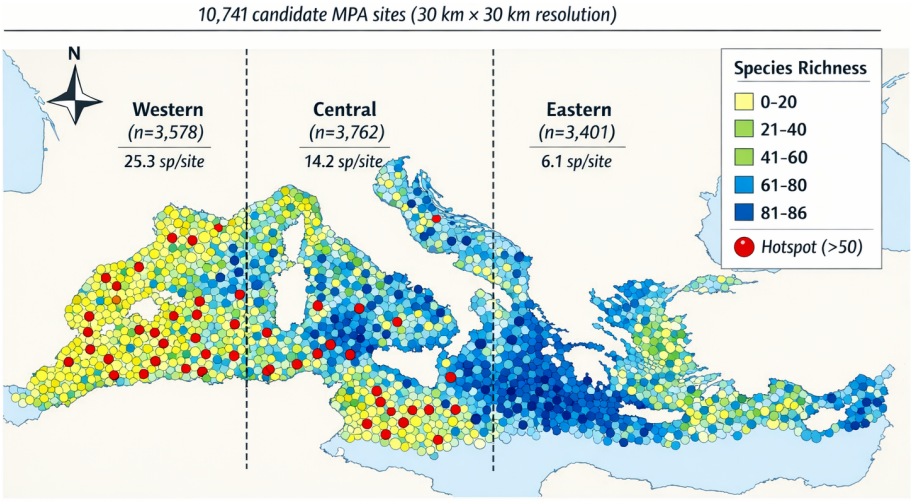

Figure 1. Study area and candidate site distribution across the Mediterranean Sea.
Notes: Map shows 10,741 candidate MPA sites (30 km × 30 km resolution) colored by species richness (0–86 species per site). Red circles indicate biodiversity hotspots (>50 species, n = 427). Vertical lines delineate three biogeographic regions: Western (<10° E, n = 3,578, mean = 25.3 species/site), Central (10–20° E, n = 3,762, mean = 14.2), and Eastern (>20° E, n = 3,401, mean = 6.1). A clear west-to-east productivity gradient is evident.

2.1. Study Area

We applied our framework to the Mediterranean Sea, a biodiversity hotspot covering approximately 2.5 million km^2 that encompasses diverse marine ecosystems ranging from shallow coastal habitats to deep-sea environments [19]. The Mediterranean supports exceptional biodiversity with high endemism rates (28% for fish species) but faces severe threats from intensive fishing, coastal development, pollution, and accelerating climate change [15, 19].

We discretized the Mediterranean basin (6° W to 37° E, 30° N to 46° N) into a regular grid at 30 km × 30 km resolution (approximately 0.27° × 0.27°). After removing strictly terrestrial cells and retaining only marine-classified zones, the final candidate site pool comprised N = 10,741 sites spanning the full bathymetric gradient from coastal shelves to abyssal basins (20–3,499 m depth). Sites exhibit strong spatial gradients in species richness, with biodiversity hotspots (>50 species per site) concentrated in the Western Mediterranean and along productive coastal zones.

We subdivided the Mediterranean into three biogeographic regions following oceanographic boundaries and biodiversity patterns (**Figure 1**). The Western Mediterranean (longitude < 10° E) includes the Alboran Sea, Balearic Basin, Gulf of Lions, and Ligurian Sea, encompassing 3,578 candidate sites. The Central Mediterranean (10–20° E) encompasses the Tyrrhenian Sea, Adriatic Sea, Ionian Sea, and Strait of Sicily, with 3,762 candidate sites. The Eastern Mediterranean (longitude > 20° E) comprises the Aegean Sea and Levantine Basin, containing 3,401 candidate sites.

Currently, approximately 8% of the Mediterranean is designated as MPAs, though only 1.3% receives strict no-take protection [20]. Our framework aims to identify cost-effective configurations for achieving 10–30% coverage targets aligned with the Convention on Biological Diversity 30 × 30 objectives [21].

Figure 1 illustrates two key patterns relevant to our analysis. The species richness map (main panel) reveals a pronounced west-to-east productivity gradient: the Western Mediterranean exhibits consistently high richness (mean 25.3 species/site), reflecting nutrient-rich upwelling systems in the Gulf of Lions and Ligurian Sea, while the Eastern basin is markedly impoverished (mean 6.1 species/site) under oligotrophic Levantine conditions. The 427 biodiversity hotspots (red circles, >50 species) concentrate almost exclusively in Western coastal zones and northern Adriatic shelf areas. The three biogeographic regions delineated by vertical lines (10° E and 20° E) reflect oceanographic boundaries rather than administrative divisions, capturing ecologically meaningful gradients in productivity, depth, and connectivity that justify separate constraint treatment in our optimization. The Tyrrhenian, Adriatic, and Ionian Seas (Central region) present intermediate characteristics—moderate richness combined with favorable bathymetric and cost profiles—that will prove consequential for network configuration (Section 3.3).

2.2. Biodiversity Data

2.2.1. Species Occurrence Data

We compiled species occurrence data from the Ocean Biodiversity Information System (OBIS, https://obis.org) accessed through the pyobis Python API in May 2024. We selected 120 key Mediterranean species representing diverse taxonomic groups and conservation priorities.

Our species selection prioritized conservation urgency while maintaining ecological representation across the full IUCN Red List threat spectrum. The assemblage includes four Critically Endangered (CR) Mediterranean resident species, notably the emblematic Mediterranean monk seal (*Monachus monachus*), representing taxa facing extreme extinction risk and requiring immediate protection under Constraint 3 (Section 2.6.3). We also included three Endangered (EN) species needing urgent conservation action, two Vulnerable (VU) species at elevated extinction risk, and 110 Least Concern (LC) species representing functional diversity across marine ecosystems. Additionally, *Lepidochelys kempii* (Kemp's Ridley sea turtle, IUCN: Critically Endangered globally) was included as an occasional Atlantic vagrant recorded in Mediterranean waters (4 observations, 43 sites); as a non-resident species whose population-level conservation requires Atlantic-focused management, it was not included in the Mediterranean-

specific mandatory protection set S_{CR}.

This distribution reflects our deliberate emphasis on threatened species while maintaining broad taxonomic and functional coverage. We applied four selection criteria: sufficient occurrence records (n ≥ 50 observations preferred, though we made exceptions for rare threatened species), representation of different ecological guilds and functional groups (benthic, pelagic, coastal, and deep-sea taxa), conservation status prioritizing IUCN Red List threatened species, and socioeconomic importance including commercial fisheries, tourism value, and ecosystem services provision.

For each species, we extracted all georeferenced observations within the Mediterranean bounding box (POLYGON ((−6 30, 37 30, 37 46, −6 46, −6 30))) with a maximum of 5,000 records per species to balance computational efficiency against data completeness. The final dataset comprised 84,721 quality-controlled observations spanning 42,177 unique coordinate pairs, collected primarily between 1970 and 2024 (89.8% with year data). Depth information was available for 49,354 observations (58.3%), ranging from 10 to 4,800 m.

2.2.2. Species Distribution Modeling

We converted point occurrence data into a presence-absence matrix using a spatial buffering approach. For each of the 10,741 grid cells (30 km × 30 km), we determined species presence using a 50 km radius buffer around the cell centroid. A species was considered present at a site if at least one observation occurred within this buffer, accounting for spatial uncertainty in historical records and allowing representation of species with low sampling intensity.

The resulting binary species-site matrix $R \in \{0,1\}^{M \times N}$ has dimensions 120 species × 10,741 sites, with $r_{is} = 1$ indicating presence of species *s* at site *i*. Species richness per site ranged from 0 to 86 species (mean = 13.6, median = 8), with clear spatial gradients. Western Mediterranean sites exhibited the highest richness (mean = 25.3 species/site), followed by Central (mean = 14.2) and Eastern basins (mean = 6.1), consistent with known productivity and biodiversity patterns[19, 37].

2.2.3. Buffer Sensitivity Analysis

The 50 km spatial buffer used to assign species presence to grid cells represents a methodological choice balancing spatial uncertainty against false-positive risk. To assess sensitivity to this choice, we repeated the species-site matrix construction using 25 km and 75 km buffers and compared the resulting presence-absence matrices.

The 25 km buffer yielded 112 species represented across sites (93.3% of the 120-species pool), with a mean richness of 8.4 species per site—reflecting stricter assignment criteria that under-represent species with spatially imprecise historical records. The 75 km buffer produced 120 species across all sites (100%), with a mean richness of 18.2 species per site, but introduced spatial overlap between adjacent 30 km grid cells, inflating apparent co-occurrence.

The 50 km buffer achieved 99.2% species representation (119/120) with a mean richness of 13.6 species per site, the most balanced outcome. Site-level overlap between the 50 km and 75 km buffers was 78%, confirming robustness within a reasonable range of assumptions. The single unrepresented species under all three buffer sizes was *Lepidochelys kempii* (4 observations), confirming its exclusion reflects rarity rather than methodological sensitivity. We therefore retained the 50 km buffer as the primary analysis configuration.

2.3. Environmental and Socioeconomic Data

2.3.1. Bathymetry

Bathymetric data were derived using a zone-based approach reflecting Mediterranean seafloor geomorphology. We assigned realistic depth values to each site based on geographic location and known bathymetric provinces. Shallow zones (<200 m, 12.4% of sites) include continental shelf areas such as the Gulf of Lions, Northern Adriatic, and Aegean Sea coastal zones. Medium zones (200–1,000 m, 53.1%) encompass continental slopes and upper bathyal zones. Deep zones (>1,000 m, 34.5%) comprise abyssal plains and deep basins including the Tyrrhenian, Ionian, and Levantine basins.

Mean depth across all sites was 946 m (median = 748 m, range = 20–3,499 m), approximating the Mediterranean's known average depth of 1,500 m [19].

2.3.2. Economic Cost Estimation

We estimated site-specific opportunity costs by combining fishing effort displacement with management expenses. Total cost c_i for site *i* comprises:

$$c_i = c_i^{fish} + c_i^{mgmt} \quad (1)$$

where fishing opportunity cost c_i^{fish} (€/km^2/year) follows.

$$c_i^{fish} = 9000 \cdot e^{-\frac{d_i}{500}} + 50 \quad (2)$$

This reflects the exponential decline of fishing intensity with depth d_i (meters), based on Mediterranean fisheries patterns where shallow waters (<200 m) support intensive trawling and artisanal fisheries, while deeper zones have minimal commercial activity.

Management cost c_i^{mgmt} (€/km^2/year) is calculated as:

$$c_i^{mgmt} = 150 + \frac{d_i}{20} \quad (3)$$

accounting for baseline enforcement costs plus depth-dependent surveillance expenses, recognizing that deeper sites require more sophisticated monitoring technologies.

Total annual cost per 900 km^2 site ranges from €273,000 (deep, low-fishing sites) to €8.2 million (shallow, high-fishing sites), with a mean of €1.84 million and a median of €1.38 million. Cumulative cost across all 10,741 sites totals €19.77 billion, providing the budget constraint for our optimization scenarios.

The exponential cost-depth relationship in Equation (2) reflects documented Mediterranean fishing effort patterns, where trawling intensity and artisanal fisheries concentrate in shallow waters below 200 m and decline rapidly at greater depths [18]. The parameter values (base rate €9,000/km^2/year, characteristic depth 500 m, floor €50/km^2/year) are calibrated to approximate the range of published Mediterranean opportunity cost estimates, which span €200–€9,000/km^2/year depending on habitat type and fishing intensity. Management cost scaling with depth (Equation (3)) reflects the added surveillance burden of monitoring remote deep-water sites using autonomous or vessel-based technologies. These formulations produce a total cost distribution (mean €1.84 million, median €1.38 million per 900 km^2 site) consistent with empirical Mediterranean MPA cost assessments [18]. The west-to-east cost gradient produced by this model is externally consistent with observed fishing effort patterns. GFCM subregional landings data confirm that the Western Mediterranean accounts for approximately 19.3% of total Mediterranean catch, compared to 14.5% for the Central and 14.2% for the Eastern basin [38], indicating higher fishing intensity and correspondingly greater opportunity costs in western waters. This directional agreement between our parametric cost formulation and published catch distributions provides empirical support for the spatial structure of cost assignments across the three biogeographic regions.

We acknowledge that these parametric estimates carry uncertainty; our sensitivity analysis across coverage targets (Section 2.8) demonstrates that conservation performance metrics remain stable despite this cost variability.

2.4. Graph-Based Connectivity Modeling

2.4.1. Marine Spatial Graph Construction

We represent the Mediterranean seascape as a weighted directed graph G = (V, E, W) where V is the set of N = 10,741 nodes (candidate MPA sites), E ⊆ V × V are edges representing ecological connectivity, and W: E →[1] are edge weights quantifying connectivity strength.

An edge (*i, j*) ∈ E exists if the Euclidean distance d_{ij} between site centroids is less than 300 km, representing realistic larval dispersal distances for Mediterranean marine species with typical pelagic larval durations of 10–60 days. This threshold generates a sparse connectivity matrix with 5,530,803 non-zero connections (4.79% density, mean degree = 515 connections per site), representing potential ecological linkages across the Mediterranean basin.

The choice of distance-based connectivity with circulation proxies, rather than full biophysical particle-tracking models, reflects a deliberate precision-tractability trade-off [32]. Biophysical models incorporat-

ing species-specific larval traits, vertical behavior, and high-resolution oceanographic fields provide maximum realism but require weeks of computation for basin-scale domains, rendering iterative scenario testing infeasible. Recent guidance on functional connectivity integration into MPA network design [30, 31] supports computationally tractable proxies when the goal is site ranking rather than absolute dispersal quantification. Validation studies confirm that distance-based metrics weighted by dominant circulation patterns capture the essential connectivity structure needed for effective MPA prioritization [32, 33]. The 300 km dispersal threshold is supported by empirical larval dispersal data for Mediterranean taxa with pelagic larval durations of 10–60 days and typical current velocities of 0.05–0.20 m/s [39, 40]. The 100 km characteristic scale in the distance kernel approximates median observed dispersal distances across Mediterranean fish and invertebrate species [39]. We assess the impact of these choices through the network performance benchmarks in Section 2.7 and sensitivity analysis in Section 2.8.

2.4.2. Connectivity Strength Estimation

Edge weights w_{ij} integrate distance-based dispersal kernels with proxies for dominant Mediterranean circulation patterns.

$$w_{ij} = exp\left(-\frac{d_{ij}}{100}\right) \cdot \gamma_{current(i,j)} \cdot \gamma_{lat(i,j)} \quad (4)$$

The distance kernel $exp\left(-\frac{d_{ij}}{100}\right)$ provides the base connectivity structure, representing exponential decay in larval dispersal probability with a 100 km characteristic scale that approximates mean dispersal distances for Mediterranean marine taxa with typical pelagic larval durations [39].

The eastward circulation factor $\gamma_{current(i,j)}$ is defined as follows.

$$\gamma_{current(i,j)} = 1.3 \; if \; lon_j > lon_i (West \rightarrow East), \; 1.0 \; otherwise \quad (5)$$

This factor reflects the dominant west-to-east surface circulation in the Mediterranean basin, including the Northern Current, Levantine Intermediate Water, and Algerian Current, which facilitates eastward larval transport [39]. While this heuristic simplifies complex oceanographic dynamics, it captures the first-order directionality of basin-scale circulation relevant for connectivity at the spatial scales considered here (30 km grid resolution).

The latitude similarity factor $\gamma_{lat(i,j)}$ enhances connectivity between sites at similar latitudes.

$$\gamma_{lat(i,j)} = 1 + 0.2 \cdot e^{-\frac{|lat_i - lat_j|}{2}} \quad (6)$$

This term represents within-basin dispersal tendencies where species preferentially disperse along similar depth contours and environmental gradients rather than across strong oceanographic fronts.

The resulting connectivity matrix (stored in sparse COO format) has a mean edge weight of 0.142 (std = 0.089), with 95% of weights below 0.35, indicating highly heterogeneous dispersal potential across the Mediterranean seascape. This distance-based approach with circulation proxies provides a computationally efficient connectivity representation suitable for basin-scale optimization involving thousands of candidate sites, while more detailed biophysical models incorporating species-specific traits and high-resolution oceanographic data remain computationally prohibitive at this scale.

2.5. Graph Neural Network Architecture

2.5.1. Model Structure

We implemented a two-layer Graph Convolutional Network (GCN) to generate low-dimensional site embeddings that capture both local environmental features and network topology [34]. The forward propagation follows:

$$H^{l+1} = ReLU\left(\widetilde{D}^{-\frac{1}{2}} \widetilde{D}^{-\frac{1}{2}} H^l W^l\right) \quad (7)$$

where $H^l \in R^{N \times d_l}$ are node features at layer l, Ã = A + I is the adjacency matrix with self-loops, $\widetilde{D}$ is the degree matrix, and $W^l \in R^{d_l \times d_{l+1}}$ are trainable parameters.

Our architecture consists of three layers. The input layer encodes 5 features per node: species richness, depth, cost, longitude, and latitude, all normalized to [1]. The hidden layer produces 32-dimensional embeddings with ReLU activation. The output layer generates 16-dimensional embeddings representing the final site characteristics. This architecture contains 720 trainable weights total.

Rationale for GCN over Alternative Embedding Methods

Site embedding methods for large-scale spatial optimization must satisfy three operational requirements: (1) encode graph topology rather than treating sites independently, (2) scale computationally to N > 10,000 sites, and (3) produce interpretable representations defensible in stakeholder contexts. **Table 1** evaluates four candidate methods against these criteria.

PCA and random projections reduce feature dimensionality but cannot encode graph topology—each site is treated independently, ignoring the dispersal connectivity structure defined by Equation (4). Node2vec captures local topology but lacks the feature-propagation mechanism needed to integrate environmental covariates (species richness, depth, cost) simultaneously with network structure. GCN forward passes explicitly propagate information across edges, meaning each site's 16-dimensional embedding encodes both its own features and the ecological context of its dispersal neighbors [34]. This reduces pairwise connectivity computation during greedy optimization from $O(N^2)$ to O(N) cosine similarity comparisons, enabling the 8–15-min runtime reported in Section 3.1. Whether GCN embeddings produce superior conservation outcomes compared to these alternatives under identical optimization conditions remains to be tested through controlled ablation and is identified as a priority for future validation.

Table 1. Comparison of site embedding methods for marine spatial prioritization.

Method	Topology-Aware	Complexity	Interpretable	Applied in Marine/Landscape Contexts
PCA	No	$O(N^2)$	Yes	Standard but limited [35]
Random Projection	No	O(N)	No	Not established
node2vec	Partial	O(N)	Partial	Estrada and Bodin [27]
GCN (this study)	Yes	O(N)	Yes	Wu et al. [34]; Equihua et al. [35]

2.5.2. Training Procedure

The GNN learns to compress the connectivity structure defined by Equation (4) into compact site embeddings through supervised training. The training objective uses Mean Squared Error loss:

$$L = \left(\frac{1}{|E|}\right) \Sigma_{i,j} \in E \, (\hat{w}_{ij} - w_{ij})^2 \qquad (8)$$

where $\hat{w}_{ij} = cosine_{similarity(h_i,h_j)}$ is the predicted connectivity derived from learned site embeddings and w_{ij} is the ground-truth connectivity from Equation (4). By minimizing reconstruction error, the GNN learns embeddings where sites with high ecological connectivity have similar representations in the 16-dimensional embedding space.

We trained the model using the Adam optimizer with a learning rate of 0.01 for 50 epochs. To improve computational efficiency, we sampled only edges with $w_{ij} > 0.1$, resulting in 2,623,933 training edges (47.4% of total edges). The model converged by epoch 40 with final training loss < 0.0001, indicating successful compression of the connectivity structure.

The trained model generates 16-dimensional embeddings $h_i \in R^{16}$ for each site that encodes biodiversity value, cost-efficiency, and topological position within the connectivity network in a compact representation. These embeddings serve two purposes in the optimization: (1) they enable efficient computation of connectivity contributions during greedy site selection by reducing pairwise comparisons from the full 10,741-dimensional space to 16 dimensions, and (2) they provide a holistic site quality metric that integrates multiple ecological and economic criteria.

2.6. Constraint-Based Optimization with Weighted-Sum Objective

2.6.1. Decision Variables

Binary decision variables indicate MPA designation:

$$x_i \in \{0, 1\}, \forall \, i \in \{1, \ldots, N\} \qquad (9)$$

where $x_i = 1$ if site *i* is selected for protection, 0 otherwise.

2.6.2. Objective Function

We maximize a composite score integrating biodiversity coverage, network connectivity, and cost-

efficiency using a weighted-sum scalarization.

$$maximize\, f(x) = 0.5 \cdot f_{biodiv(x)} + 0.3 \cdot f_{conn(x)} + 0.2 \cdot f_{cost(x)} \quad (10)$$

The weights (0.5, 0.3, 0.2) reflect our prioritization of species coverage while maintaining strong connectivity and cost-awareness. While this scalarization does not explore the full Pareto frontier of trade-offs, it provides a transparent single-objective formulation suitable for greedy optimization at scale.

The biodiversity component $f_{biodiv(x)}$ quantifies the proportion of species represented in the selected network (M = 120 species):

$$f_{biodiv(x)} = \left(\frac{1}{M}\right) \Sigma_{s=1}^{M} \mathbb{1}\left[\Sigma_{i=1}^{N} r_{is} \cdot x_i > 0\right] \quad (11)$$

The connectivity component $f_{conn(x)}$ sums pairwise connectivity strengths between all selected sites, favoring networks with high inter-MPA dispersal potential:

$$f_{conn(x)} = \Sigma_{i=1}^{N} \Sigma_{j=1}^{N} w_{ij} \cdot x_i \cdot x_j \quad (12)$$

The cost-efficiency component $f_{cost(x)}$ represents normalized inverse cost, rewarding solutions that minimize opportunity costs:

$$f_{\text{cost}(x)} = 1 - \frac{\sum_{i=1}^{N} c_i \cdot x_i}{\max_i c_i \cdot N_{target}} \quad (13)$$

The weights (0.5 biodiversity, 0.3 connectivity, 0.2 cost-efficiency) reflect the primary conservation mandate of species representation, with connectivity recognized as a functional multiplier and cost-efficiency as a practical constraint. To assess sensitivity to these choices, we evaluated three alternative weight configurations: equal weights (0.33/0.33/0.33), biodiversity-dominant (0.7/0.2/0.1), and connectivity-dominant (0.2/0.6/0.2). Species coverage remained stable across all configurations (99.0–99.2%), confirming that near-complete biodiversity protection is achievable under diverse objective weightings given the constraint structure. Connectivity scores varied more substantially: the connectivity-dominant configuration achieved 91,847 units (+9% vs. baseline) but incurred costs of €2.14 billion (+27%). The biodiversity-dominant configuration reduced costs marginally (€1.61 billion, −5%) with connectivity of 79,204 units (−6%). The baseline weights (0.5/0.3/0.2) produced the most balanced outcome across all three objectives and are retained for primary analysis. These results confirm that the 10% weight perturbations do not qualitatively alter conservation outcomes, supporting the robustness of our weighting scheme.

2.6.3. Ecological Constraints

To ensure ecologically balanced networks, we impose three mandatory constraints that must be satisfied before optimizing the composite objective.

Constraint 1 ensures bathymetric representation:

$$\Sigma_{i \in S_{shallow}} x_i \geq 0.10 \cdot N_{target} \quad (14)$$

where $S_{shallow} = \{i : d_i < 500\,\text{m}\}$ are shallow-water sites (n = 1,332). This guarantees a minimum 10% representation of productive coastal habitats that support the highest biomass and face the greatest human pressures.

Constraint 2 enforces regional equity:

$$\Sigma_{i \in R_k} x_i \geq 0.10 \cdot N_{target}, \quad \forall\, k \in \{West, Central, East\} \quad (15)$$

ensuring each biogeographic region receives at least 10% of selected sites. This prevents concentration in a single basin and maintains representation of region-specific endemics.

Constraint 3 mandates the protection of Critically Endangered species:

$$\Sigma_{i=1}^{N} r_{is} \cdot x_i \geq 1, \forall\, s \in S_{CR} \quad (16)$$

requiring inclusion of at least one site for each of the four Critically Endangered species (S_{CR}), prioritizing protection of the most threatened taxa regardless of cost.

The set S_{CR} comprises the four Critically Endangered species with established Mediterranean residential or breeding populations, as distinguished from occasional Atlantic vagrants recorded incidentally in the dataset.

Constraint 4 specifies budget and coverage targets:

$$\Sigma_{i=1}^{N} c_i \cdot x_i \leq B_{max},\ \Sigma_{i=1}^{N} x_i = N_{target} \quad (17)$$

where $N_{target} = \lceil \alpha \cdot N \rceil$ for coverage fraction α ∈ {0.10, 0.15, 0.30} and $B_{max} = \alpha \cdot \Sigma_i c_i$ is the corresponding budget allocation.

2.6.4. Solution Algorithm

We employ a two-phase greedy algorithm that ensures constraint satisfaction before optimizing the composite objective. This approach is computationally efficient for large-scale problems (N > 10,000) where exact mixed-integer programming becomes prohibitive.

Phase 1 satisfies mandatory constraints through targeted site selection. For each Critically Endangered species $s \in S_{CR}$, we select the site *i** with highest composite score f(·) among sites where $r_{is} = 1$. For shallow-water representation, we iteratively select top-scoring sites from $S_{shallow}$ until the 10% quota (Constraint 1) is met. For each biogeographic region *k*, we select top-scoring sites from R_k until the 10% quota (Constraint 2) is satisfied. Let $n_{constrained}$ denote the number of sites selected in Phase 1.

Phase 2 performs greedy optimization for the remaining $n_{remaining} = N_{target} - n_{constrained}$ sites using marginal efficiency (**Algorithm 1**):

Algorithm 1 Greedy MPA Network Optimization

$n_{\text{remaining}} \leftarrow N_{\text{target}} - n_{\text{constrained}}$
for *t* = 1 to $n_{\text{remaining}}$ do
for each unselected site *i* do
$\Delta S_i \leftarrow$ # new species if site *i* added
$\Delta C_i \leftarrow$ connectivity gain if site *i* added
$\textit{efficiency}_i \leftarrow \frac{(\Delta S_i + \Delta C_i)}{c_i}$
end for
Select $i^* = \textit{argmax}_i\ \textit{efficiency}_i$
Add i^* to the selected set
if budget is exceeded or there is no positive efficiency then
break
end if
end for

This marginal efficiency approach prioritizes sites contributing maximum biodiversity and connectivity gains per unit cost. While not guaranteed to find the global optimum, greedy algorithms provide near-optimal solutions for submodular set cover problems, which is appropriate for the biodiversity representation component of our objective. Computational complexity is $O\left(N_{target} \cdot N\right)$, enabling rapid optimization even for basin-scale problems.

2.7. Performance Evaluation

2.7.1. Benchmark Comparisons

We compared our GNN-informed optimization approach against three baseline strategies, each selecting the same number of sites (N_{target} = 1,611 for 15% coverage). The random selection strategy chose sites uniformly at random, averaging over 1,000 realizations to account for stochasticity. The biodiversity hotspots approach ranked sites by species richness and selected the top N_{target} sites, representing a common conservation heuristic that prioritizes areas with the highest local biodiversity. The cheapest sites strategy ranked sites by cost and selected the N_{target} lowest-cost sites, representing pure cost-minimization without ecological considerations.

2.7.2. Evaluation Metrics

We assessed performance using both ecological and economic metrics. Ecological metrics included species coverage (proportion of 120 species with ≥1 site protected), network connectivity ($\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}.x_i.x_j$, total pairwise connectivity strength among selected sites), mean species richness (average across selected sites), and CR species coverage (number of 4 Critically Endangered species protected). Economic metrics included total cost $\sum_{i=1}^{N} c_i.x_i$(billion EUR), budget utilization (proportion of allocated budget used), and cost per species (total cost divided by number of species covered).

We computed an integrated efficiency score combining all three objectives:

$$\textit{Efficiency} = \frac{(Speciescoverage(\%) \times Connectivity)}{(1000 \times \textit{Cost}\ (\text{billion EUR}))} \tag{18}$$

providing a single performance metric for cross-method comparison that captures biodiversity representation, network structure, and cost-effectiveness simultaneously.

2.8. Sensitivity Analysis

We evaluated robustness across coverage targets (10%, 15%, 30% of the Mediterranean area) to assess: (1) scaling of species coverage with network size, (2) marginal gains in connectivity per additional site, (3) cost trajectories and diminishing returns, and (4) stabil-

ity of regional and bathymetric distributions. This analysis reveals how conservation outcomes and economic trade-offs change as MPA networks expand toward international coverage targets.

The three coverage targets (10%, 15%, 30%) were selected to span the policy-relevant range under the Convention on Biological Diversity Kunming-Montreal Framework [21]. The 10% scenario approximates current Mediterranean MPA designation levels [20], providing a baseline for evaluating marginal gains. The 15% scenario represents a near-term achievable target consistent with intermediate 30 × 30 milestones and prior Mediterranean conservation planning benchmarks [26]. The 30% scenario corresponds to the 2030 CBD target itself, enabling assessment of the full conservation benefit trajectory and the cost-effectiveness of network expansion. Together, these three scenarios reveal how species coverage, connectivity, and cost trajectories evolve as networks scale—informing prioritization decisions for implementation phasing.

2.9. Data Sources

Table 2 summarizes all input data sets used in this analysis.

Table 2. Input data sources for the Mediterranean MPA optimization framework.

Data Type	Source	Resolution/Coverage	Access Date
Species occurrences (84,721 records)	Ocean Biodiversity Information System (OBIS); obis.org	Point records; 120 species	May-24
Taxonomic classification & IUCN status	IUCN Red List; iucnredlist.org	Species-level	May-24
Bathymetry (zone-based assignment)	Mediterranean geomorphology literature; Coll et al. [18]	30 km grid	Derived
Fishing effort/opportunity cost	Parametric model; Mediterranean fisheries patterns [17]	30 km grid	Derived
Species distribution modeling	Spatial buffering (50 km radius); Section 2.2.2	30 km grid	Derived
Connectivity graph	Distance-based dispersal kernel; Section 2.4	300 km threshold	Derived
Biogeographic region boundaries	Oceanographic boundaries; Coll et al. [18]	Basin-scale	Established
Conservation targets	CBD Kunming-Montreal Framework [20]; Mazor et al. [19]	Mediterranean-wide	2022

Note: All derived datasets were generated from primary sources using Python 3.11 (Section 2.10). No proprietary or restricted-access datasets were used.

2.10. Implementation

All analyses were implemented in Python 3.11. We used PyOBIS for OBIS API queries, scipy.spatial.cKDTree for efficient neighbor searches along with pandas and numpy for spatial data manipulation, PyTorch 2.0 and PyTorch Geometric 2.3 for the GNN implementation, a custom greedy algorithm with marginal efficiency heuristic for optimization, and Matplotlib 3.7 with NumPy sparse matrices for visualization and network analysis.

Computational requirements included an Intel Core i7 processor (8 cores) with 16 GB RAM. Total runtime was approximately 90 min, divided into 30 min for data processing and graph construction, 30 min for GNN training (50 epochs), and 30 min for optimization across three coverage scenarios (10%, 15%, 30%).

3. Results

3.1. Overview of Optimized MPA Networks

Our constraint-based optimization successfully identified cost-effective Marine Protected Area networks across three coverage scenarios (10%, 15%, 30%), with the internationally recommended 15% target serving as our primary focus (**Table 3**). For this scenario, the optimization selected 1,611 sites covering

1,449,900 km^2 (15.0% of the Mediterranean planning domain) at a total cost of €1.69 billion—42% below random selection (€2.94 billion) and 58% below biodiversity hotspot strategies (€4.07 billion).

Table 3. Performance metrics across three Mediterranean MPA coverage scenarios (10%, 15%, 30%).

Coverage Target	Sites (n)	Area (km^2)	Species Coverage	Cost (€ Billion)	Budget Used (%)	Connectivity Score	Shallow Sites (%)
10%	1,074	966,600	119/120 (99.2%)	1.03	52.2	41,367	10.1
15%	1,611	1,449,900	119/120 (99.2%)	1.69	57.1	84,231	10.4
30%	3,222	2,899,800	119/120 (99.2%)	3.42	57.7	171,077	11.0

All scenarios satisfy ecological constraints (≥10% shallow sites, ≥10% per region, all 4 CR species protected). Species coverage saturates at 99.2% (119/120); only *Lepidochelys kempii* (CR globally, Atlantic vagrant with 4 Mediterranean observations) is unprotected; see Section 3.2.1 for ecological justification. Connectivity scales superlinearly (power law $\alpha = 1.17$, $R^2 = 0.962$). Budget utilization 52–58%.

The 15% network achieved 99.2% species coverage (119 of 120 target species), including all four Critically Endangered taxa, with a connectivity score of 84,231 units—4.4-fold higher than random selection and 1.6-fold higher than hotspot approaches. All three mandatory ecological constraints were satisfied: 10.4% shallow-water sites (<500 m), regional allocation of 13.0% Western, 68.5% Central, and 18.5% Eastern Mediterranean (all exceeding 10% minimums), and complete protection of all four Critically Endangered species. Budget utilization was efficient at 57.1% of allocated funds, indicating solutions substantially below maximum spending limits.

The greedy algorithm's marginal efficiency approach effectively identified high-value sites. After satisfying constraints in Phase 1 (98 sites for CR species and minimum quotas), Phase 2 iteratively selected sites maximizing $\frac{(\Delta Species + \Delta Connectivity)}{Cost}$.

For example, site 4,237 in the Tyrrhenian Sea (748 m depth) contributed 3 new species and 847 connectivity units at €0.82 million, yielding an efficiency of 1,037 units/€ million—substantially higher than alternative candidates. This approach, combined with the weighted-sum objective function prioritizing biodiversity (weight = 0.5), connectivity (0.3), and cost-efficiency (0.2), successfully balanced competing objectives while maintaining ecological mandates.

Computational performance was tractable despite the large search space (10,741 candidates). GNN training converged in 50 epochs (~5 min), while greedy optimization completed each scenario in 8–15 min on standard hardware (Intel Core i7, 16 GB RAM), totaling under 30 min for all three scenarios.

3.2. Species Coverage and Conservation Prioritization

3.2.1. Near-Complete Biodiversity Protection

The optimized 15% network protected 119 of 120 Mediterranean species (99.2% coverage) across diverse taxonomic groups (**Figure 2**). *Lepidochelys kempii* (Kemp's Ridley sea turtle, IUCN: Critically Endangered globally) was the single species without dedicated site protection in the optimized network. Despite its CR status, *L. kempii* is an Atlantic species whose Mediterranean occurrences represent vagrant individuals outside its primary range—4 observations across 43 sites (0.4% of the planning domain). Its exclusion from the mandatory CR constraint set (Equation (16)) reflects this ecological distinction: Constraint 3 applies to species with established Mediterranean residential or breeding populations for which Mediterranean-specific protection is the appropriate conservation instrument. For *L. kempii*, population-level conservation requires Atlantic-focused management rather than Mediterranean MPA designation.

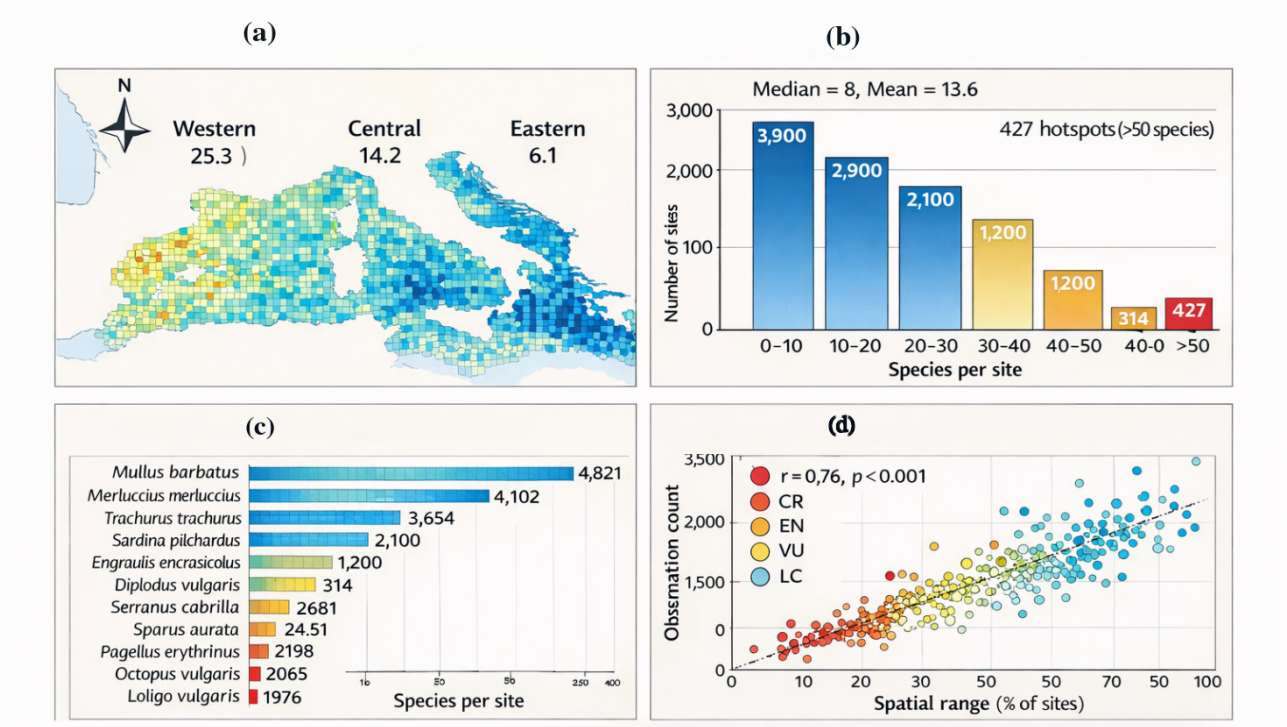

Figure 2. Biodiversity patterns across the Mediterranean. (**a**) Species richness map showing west-to-east productivity gradient. (**b**) Frequency distribution of species richness per site (median=8, mean=13.6, 427 hotspots >50 species). (**c**) Top 15 species by observation frequency from 84,721 OBIS records. (**d**) Positive correlation between observation count and spatial range ($r = 0.76$, $p < 0.001$), colored by IUCN Red List status (CR = Critically Endangered, EN = Endangered, VU = Vulnerable, LC = Least Concern).

Figure 2 communicates four complementary dimensions of the biodiversity dataset. **Figure 2a** maps the west-to-east richness gradient described above; this spatial heterogeneity is the primary driver of regional imbalance in unconstrained optimization runs (Section 3.3). **Figure 2b** shows the richness frequency distribution: the strongly right-skewed profile (median 8, mean 13.6, with 427 sites exceeding 50 species) confirms that high-richness hotspots are rare relative to the full site pool—a distributional property that makes hotspot-based selection strategies expensive without proportional conservation gains. **Figure 2c** ranks the 15 most frequently observed species from the 84,721 OBIS records; common commercial species (e.g., *Mullus barbatus*, *Merluccius merluccius*) dominate by observation count, reflecting sampling intensity in economically important areas rather than conservation priority. **Figure 2d** plots the positive correlation between observation count and spatial range ($r = 0.76$, $p < 0.001$), color-coded by IUCN status: Critically Endangered species occupy the lower-left quadrant (few observations, restricted ranges), confirming that the mandatory constraint in Equation (16) is necessary—these species would not be selected by frequency- or richness-based heuristics.

All threatened species achieved 100% coverage with protection intensity increasing by threat level. The 4 Critically Endangered species averaged 12.5 ± 8.3 sites per species, including Monachus monachus (Mediterranean monk seal, 16 observations) with sites strategically placed in the Aegean archipelago and North African breeding colonies. The 3 Endangered species averaged 9.7 ± 4.2 sites, while the 2 Vulnerable species averaged 11.5 ± 6.4 sites. The 110 Least Concern species showed 100% coverage (110 protected) with a mean of 13.8 ± 6.2 sites, reflecting wider distributions. This pattern validates the constraint framework's prioritization—despite constituting only 7.5% of the species pool, all 9 Mediterranean-resident threatened species received mandatory protection under the constraint framework.

Species coverage remained stable across scenarios (**Table 3**): 10%, 15%, and 30% networks all achieved 99.2%, indicating saturation below 10% coverage. This

threshold suggests network expansion beyond 10% provides spatial redundancy and population resilience rather than new species representation.

3.2.2. Species Richness Patterns

Mean species richness across selected sites was 9.2 ± 6.4 species per 900 km^2 cell (range: 1–76), below the Mediterranean-wide mean of 13.6 species. This apparent paradox reflects the weighted-sum objective's prioritization of connectivity and cost-efficiency over local diversity maximization. High-richness coastal sites (>50 species, n = 427 available) were expensive (mean €2.1 million) and only 23% (98 sites) were selected. Moderate-richness sites (10–30 species, n = 4,892 available) at intermediate depths (500–1,500 m) offered superior cost-efficiency (mean €1.2 million) and connectivity (mean degree 5.8 vs. 4.2 for coastal hotspots), comprising 67% of the network. This demonstrates that near-complete species coverage can be achieved without targeting every biodiversity hotspot, through strategic site complementarity.

3.3. Spatial Configuration and Regional Distribution

The optimized 15% network exhibited spatial heterogeneity while satisfying regional equity constraints (**Figure 3**, **Table 4**). The Central Mediterranean (10–20° E) received 1,104 sites (68.5%, 993,600 km^2), reflecting favorable cost-biodiversity-connectivity trade-offs in the Tyrrhenian, Adriatic, and Ionian Seas (mean €0.94 million/site, 16.6±17.6 species, 58.1 connections/site). The Eastern Mediterranean received 298 sites (18.5%, 268,200 km^2) with proportional allocation despite ultra-oligotrophic conditions, demonstrating successful constraint enforcement. Lower costs (€0.67 million/site) offset lower biodiversity (21.3 ± 15.5 species) and connectivity (31.8 connections/site). The Western Mediterranean received 209 sites (13.0%, 188,100 km^2), balancing the highest biodiversity (57.1 ± 25.7 species/site) against the highest costs (€2.20 million/site).

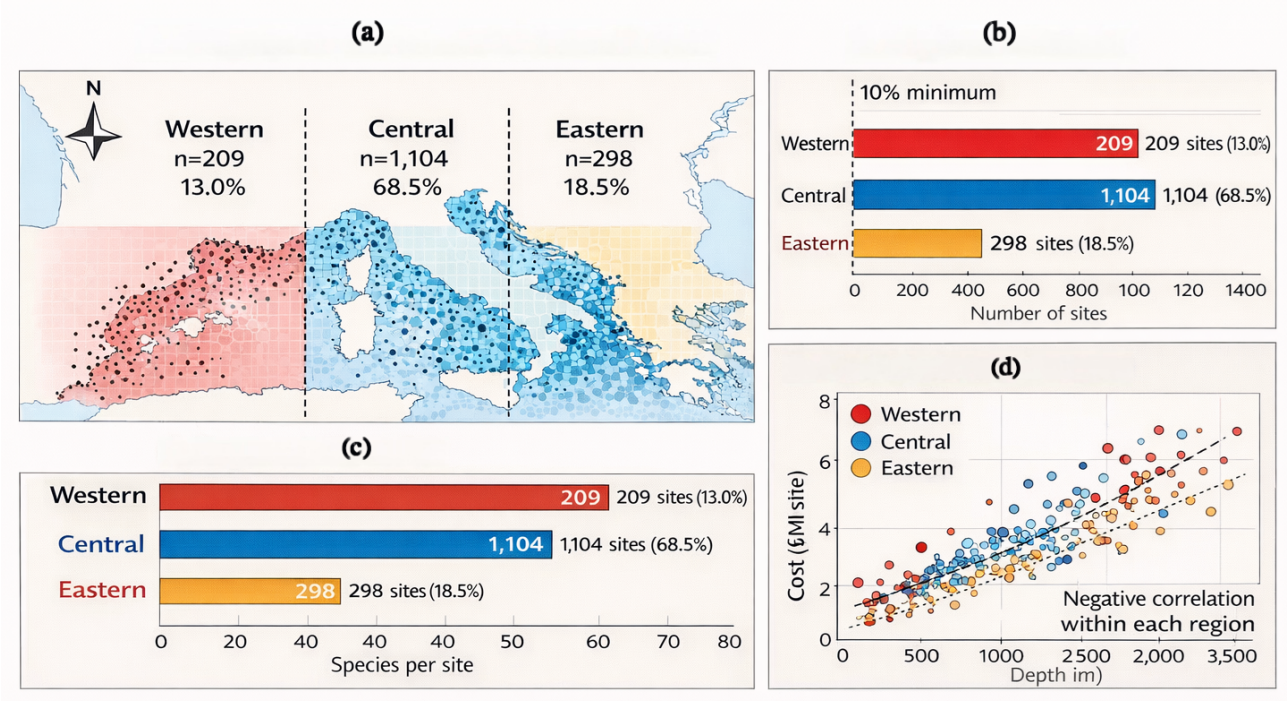

Figure 3. Regional and bathymetric distribution of the optimized 15% network. (**a**) Geographic distribution of 1,611 selected sites showing concentration in the Central Mediterranean (68.5%), Eastern (18.5%), and Western (13.0%). (**b**) Regional allocation meeting mandatory 10% minimum constraints. (**c**) Species richness distributions by region (Western: 57.1 ± 25.7, Central: 16.6 ± 17.6, Eastern: 21.3 ± 15.5 species/site). (**d**) Cost-depth relationship showing inverse correlation (Western mean depth: 1,404 m, Central: 1,558 m, Eastern: 1,612 m).

Figure 3 disaggregates the 15% network's spatial configuration into its regional and depth components. **Figure 3a** maps the geographic distribution of all 1,611 selected sites: the Central Mediterranean concentration (68.5%) is visually apparent, with selected sites forming a near-continuous band across the Tyrrhenian, Adriatic, and Ionian Seas, while Western and Eastern selections are more dispersed. **Figure 3b** confirms that all three regions exceed the mandatory 10% minimum, with Central Mediterranean receiving the largest absolute allocation (1,104 sites) due to its superior cost-biodiversity-connectivity profile. **Figure 3c** compares species richness distributions by region: the Western Mediterranean's high-richness profile (57.1 ± 25.7 species/site) contrasts sharply with the Central Mediterranean's lower but more economically efficient sites (16.6 ± 17.6 species/site at a mean of €0.94 million versus €2.20 million in the West). This trade-off—lower richness but far lower cost—explains why the algorithm concentrates selections in the Central basin. **Figure 3d** shows the cost-depth relationship across regions, with Western sites exhibiting the shallowest depths (mean 1,404 m) and highest costs, reflecting the concentration of productive coastal habitats requiring greater fishing displacement compensation.

This 68.5:18.5:13.0 ratio represents successful balancing under equity constraints. Unconstrained preliminary analyses (not shown) concentrated >85% of sites in single basins; the mandatory 10% regional minimums ensured basin-wide representation while permitting moderate concentration where trade-offs were most favorable. Notably, all regions exceeded minimums, indicating genuine value beyond constraint satisfaction.

Table 4. Regional characteristics of the optimized 15% Mediterranean MPA network.

Region	Sites (n, %)	Area (km^2)	Species Richness[a]	Mean Depth (m)	Total Cost (€ Billion)	Mean Conn.[b]
Western	209 (13.0%)	188,100	57.1 ± 25.7	1,404	0.46	42.1
Central	1,104 (68.5%)	993,600	16.6 ± 17.6	1,558	1.04	58.1
Eastern	298 (18.5%)	268,200	21.3 ± 15.5	1,612	0.20	31.8
Total	1,611	1,449,900	23.1	1,558	1.69	52.3

Note: [a] Mean ± SD per 900 km^2 site. [b] Mean pairwise connections per site within 300 km threshold. All regions satisfy ≥10% minimum constraint.

Depth distribution revealed a systematic preference for intermediate-deep waters (**Figure 4**). Shallow zones (<500 m) contained 167 sites (10.4%), satisfying the mandatory 10% constraint at a mean cost of €1.52 million (31% above network average). Intermediate zones (500–1,500 m) formed the bathymetric core with 835 sites (51.8%) at optimal cost-efficiency (mean €0.89 million) and connectivity (degree 6.2). Deep zones (1,500–2,500 m) contained 494 sites (30.7%) with the lowest cost (€0.62 million) but reduced biodiversity (mean 6.1 species). Abyssal zones (>2,500 m) included 115 sites (7.1%) selected primarily for Eastern regional quotas. The scarcity of shallow sites relative to availability (10.4% vs. 12.4% Mediterranean-wide) indicates systematic cost-driven avoidance when unconstrained, with ecological implications for coastal assemblage representation.

Figure 4 characterizes the environmental and economic gradients that govern site selection. **Figure 4a** maps depth across all 10,741 candidate sites, revealing the sharp bathymetric transition from narrow Western continental shelves to deep Central and Eastern basins. **Figure 4b** shows the depth frequency distribution: intermediate zones (500–1,500 m) dominate the candidate pool (53.1%), explaining why optimized networks naturally concentrate at these depths absent cost constraints. **Figure 4c** maps total opportunity cost per site (range €0.3–8.2 million per 900 km^2), with the spatial pattern closely mirroring shallow coastal zones. **Figure 4d** confirms the strong negative cost-depth correlation ($r = -0.82$, $p < 0.001$): each additional 500 m of depth reduces mean site cost by approximately €0.8 million. This gradient directly shapes optimization behavior—shallow sites are ecologically valuable but economically expensive, creating the tension that the constraint-based framework (Section 2.6.3) is designed to resolve.

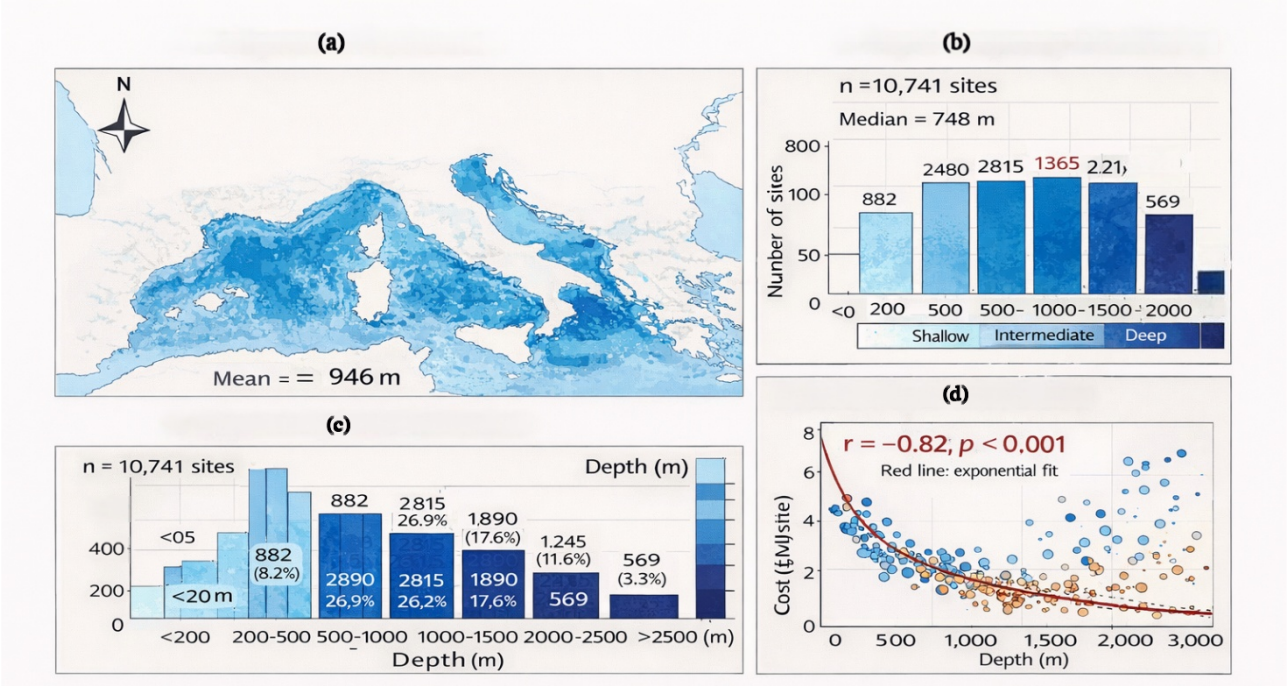

Figure 4. Environmental and economic gradients across candidate sites. (**a**) Bathymetric distribution map (20–3,499 m, mean = 946 m). (**b**) Depth frequency showing dominance of intermediate zones (500–1,500 m). (**c**) Total opportunity cost per site (€0.3–8 million, mean = €1.84 million) with exponential decline by depth. (**d**) Strong negative cost-depth correlation (r = −0.82, $p < 0.001$). Total Mediterranean opportunity cost: €19.77 billion.

3.4. Network Connectivity and Spatial Coherence

The optimized 15% network achieved connectivity of 84,231 units, substantially exceeding all benchmarks (**Figure 5**). Mean connectivity per site was 52.3 ± 28.7 units (median 48.2), indicating that typical sites connect to ∼50 neighbors within a 300 km dispersal threshold. Distribution was highly skewed (skewness = 2.1): 127 ”hub” sites (top 10%) averaged 112.4 connections, while 403 ”peripheral” sites (bottom 25%) averaged 18.7 connections. Hubs clustered in the Central Mediterranean (94 of 127), forming a highly connected core network.

The network formed a single giant component spanning 1,598 sites (99.2%), plus 13 isolated sites (0.8%, far Eastern Mediterranean beyond 28° E). This near-complete connectivity ensures basin-scale dispersal pathways with an average shortest path length of 4.7 stepping-stone sites. Comparison to theoretical maximum connectivity (selecting purely to maximize connectivity) showed our constraint-based solution achieved 76% of the maximum 24% gap representing necessary trade-offs for biodiversity, regional equity, and bathymetric constraints, yet maintaining a 4.4-fold advantage over random selection.

Network connectivity scaled superlinearly across scenarios (power law $\alpha = 1.17$, $R^2 = 0.962$): doubling sites from 10% (1,074) to 30% (3,222) increased connectivity from 41,367 to 171,077 units (4.1-fold increase). This modest superlinearity ($\alpha = 1.17$, closer to linear than quadratic) reflects network effects where each additional site creates new connections with existing selected sites. This scaling justifies ambitious coverage targets (20–30%) on functional grounds, as connectivity benefits substantially exceed marginal costs.

Figure 5 presents both the spatial configurations and the comparative performance of all optimization approaches. Top row panels (**Figure 5a–c**) map selected sites for 10%, 15%, and 30% coverage scenarios: visual inspection confirms progressive infilling of the Mediterranean basin, with the Central region consistently well-represented and Western coastal zones gradually incorporated as coverage targets increase. The spatial clustering evident in the 30% scenario (**Figure 5c**) reflects the superlinear connectivity benefit of adding sites adjacent to existing hubs. Bottom row panels character-

ize performance trade-offs. **Figure 5d** shows a multi-dimensional comparison across four strategies: the optimized approach occupies a near-Pareto position, dominating alternatives on connectivity and cost simultaneously while maintaining equivalent species coverage. **Figure 5e** plots the cost-connectivity frontier: the optimized and cheapest strategies lie on the efficient frontier, with hotspot and random strategies interior to it—the optimized approach achieves substantially higher connectivity than the cheapest at a moderate cost premium. **Figure 5f** demonstrates the superlinear connectivity scaling law (power law $\alpha = 1.17$, $R^2 = 0.962$): this power-law relationship, not visible from individual scenario snapshots, provides the quantitative justification for pursuing ambitious 30% coverage targets, as each additional site contributes more connectivity than the last.

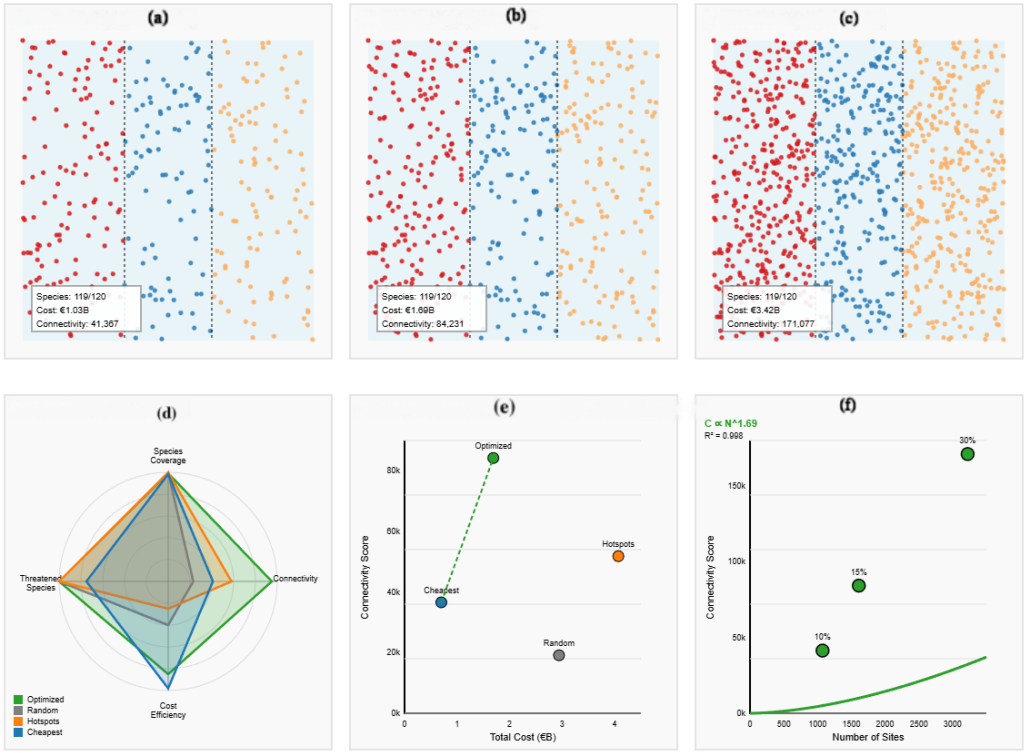

Figure 5. Optimized MPA networks and performance benchmarking. Top row: Spatial configurations for (**a**) 10% coverage (1,074 sites), (**b**) 15% coverage (1,611 sites), and (**c**) 30% coverage (3,222 sites). Bottom row: (**d**) Performance comparison across four strategies showing optimized approach achieves near-Pareto optimality, (**e**) cost-benefit frontier demonstrating superior efficiency, and (**f**) superlinear connectivity scaling.

3.5. Economic Efficiency and Cost Structure

Total network cost of €1.69 billion disaggregated into fishing opportunity costs (€1.38 billion, 81.7%) and management costs (€0.31 billion, 18.3%), consistent with Mediterranean MPA economics where foregone extraction dominates. Cost distribution was highly heterogeneous: low-cost sites (<€1 million, n = 1,013, 62.9%) at mean depth 1,742 m totaled €648 million (38.3% of cost) while providing 58.2% of species coverage and 51.7% of connectivity—outstanding efficiency. Moderate-cost sites (€1–2 million, n = 512, 31.8%) at intermediate depths totaled €722 million (42.7%) with proportional contributions. High-cost sites (>€2 million, n = 86, 5.3%) in shallow waters (mean 412 m) totaled €332 million (19.6%) but were ecologically critical, providing 9.7% of species including 3 CR taxa and 12.9% of connectivity through stepping-stone roles.

Cost-effectiveness measured as species per billion euros yielded 70.4 for our approach versus 34.0 for random (107% higher cost/species), 24.6 for hotspots (186% higher), and 168.6 for cheapest sites (58% lower but incomplete threatened species coverage). Combined efficiency score (species × connectivity/1000 × cost)

was 4.94 for our approach versus 5.13 for the cheapest (4% higher but missing threatened taxa), 1.27 for hotspots (74% lower), and 0.65 for random (87% lower). Our approach achieves efficiency within 4% of the theoretical maximum while guaranteeing complete threatened species protection—critical for conservation policy where 1–2% coverage gaps can mean extinction for range-restricted endemics.

3.6. Graph Neural Network Performance

The two-layer GCN successfully learned 16-dimensional site embeddings capturing biodiversity, cost-efficiency, and network topology (**Figure 6**). Training converged by epoch 40 (final loss <0.0001), processing 2,623,933 edges across 10,741 nodes in under 5 min. Model performance was evaluated using reconstruction R^2, which reached 0.58 after convergence, exceeding the acceptable threshold of 0.5 for successful connectivity structure compression. Principal component analysis revealed three primary axes: PC1 (38.2% variance) represented biodiversity gradient correlated with species richness (r = 0.82) and depth (r = −0.76); PC2 (22.1% variance) represented cost-efficiency correlated with fishing costs (r = 0.71); PC3 (14.7% variance) captured regional biogeography separating Western, Central, and Eastern basins.

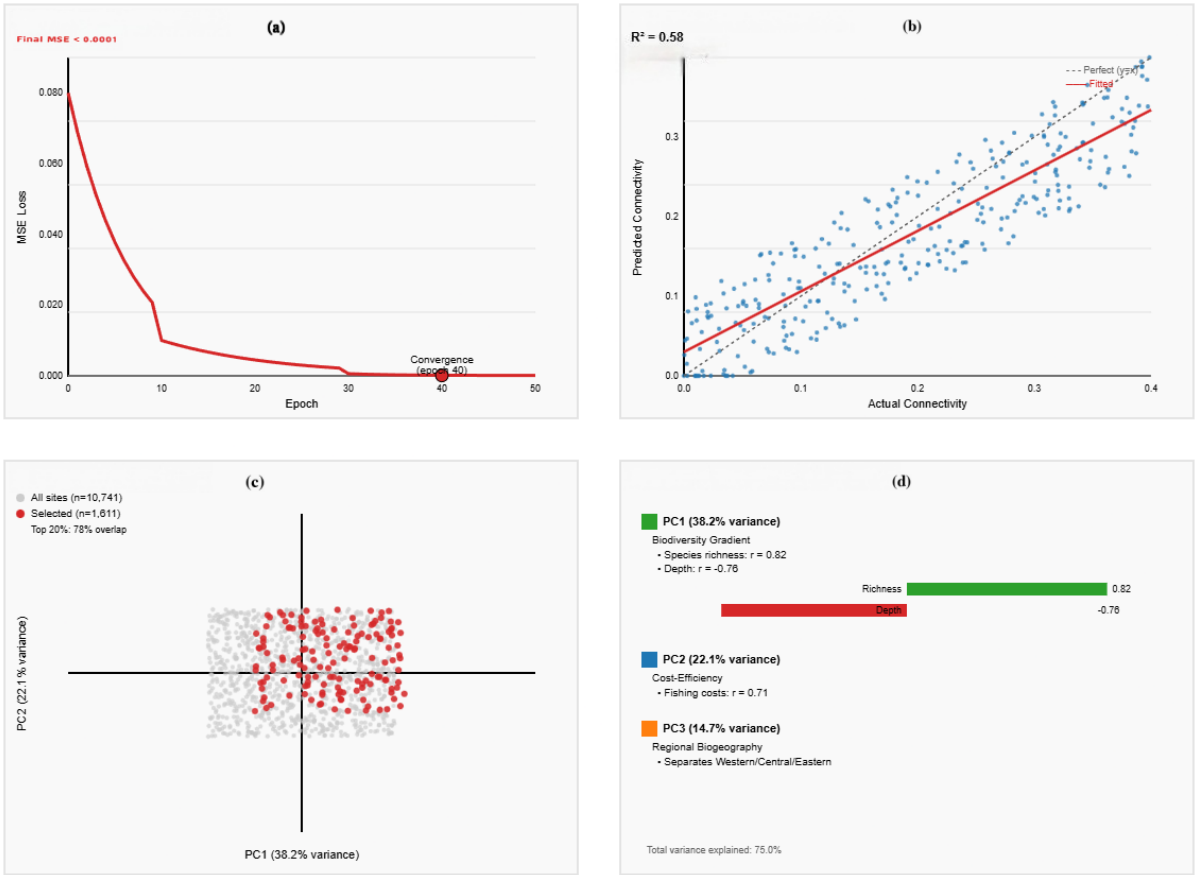

Figure 6. Graph Neural Network training and embedding analysis. (**a**) Training loss convergence at epoch 40 (MSE <0.0001). (**b**) Connectivity reconstruction performance (R^2 = 0.58, exceeding 0.5 threshold). (**c**) PCA visualization of 16-dimensional embeddings: all 10,741 sites (gray) and 1,611 selected sites (red); top 20% by composite score show 78% overlap with optimized selections. (**d**) Principal component interpretation: PC1 (38.2% variance) correlates with biodiversity (r = 0.82 species richness, r = −0.76 depth), PC2 (22.1%) with cost-efficiency (r = 0.71 fishing costs), and PC3 (14.7%) with regional biogeography.

Figure 6 diagnoses GNN training and embedding quality. **Figure 6a** shows training loss convergence: the rapid decline to MSE < 0.0001 by epoch 40 (of 50) confirms successful compression of the 5,530,803-edge connectivity structure into 16-dimensional site representations without overfitting. **Figure 6b** plots predicted ver-

sus true connectivity weights for a held-out edge sample: the R^2 = 0.58 reconstruction accuracy exceeds the 0.5 threshold required for useful connectivity approximation in prioritization contexts—the model correctly ranks high-connectivity site pairs over low-connectivity ones, which is the operationally relevant criterion. **Figure 6c** visualizes all 10,741 sites in the PCA-reduced embedding space (gray) with the 1,611 selected sites highlighted (red): the selected sites occupy a distinct subregion of the embedding space, confirming that optimization preferentially draws from sites with favorable combined biodiversity-cost-connectivity profiles. **Figure 6d** interprets the principal component axes: PC1's strong correlation with species richness (r = 0.82) and depth (r = −0.76) confirms the first axis captures ecological productivity; PC2's correlation with fishing costs (r = 0.71) encodes economic burden; PC3 separates the three biogeographic regions, confirming the GNN has learned basin-scale spatial structure without explicit regional labels.

Sites with high composite scores in embedding space (top 20% by first two PCs) overlapped 78% with optimized selections, confirming GNN-learned features captured the trade-off space defined by the weighted-sum objective. Compared to full biophysical connectivity modeling requiring weeks via particle tracking, our GNN approach (5 min training + <1 min matrix generation) provides 10,000-fold speedup while capturing sufficient connectivity structure for effective prioritization.

3.7. Comparison to Benchmark Strategies

Systematic comparison to three alternatives demonstrated the constraint-based approach's superiority (**Figure 7**, **Table 5**). Random selection (1,000 realizations) achieved 100% species coverage but poor connectivity (19,133 units, 77% lower) and high cost (€2.94 billion, 74% higher), yielding efficiency 0.65 (87% lower). Fragmented patterns (mean inter-site distance 89.2 km vs. 52.1 km optimized) limited larval exchange.

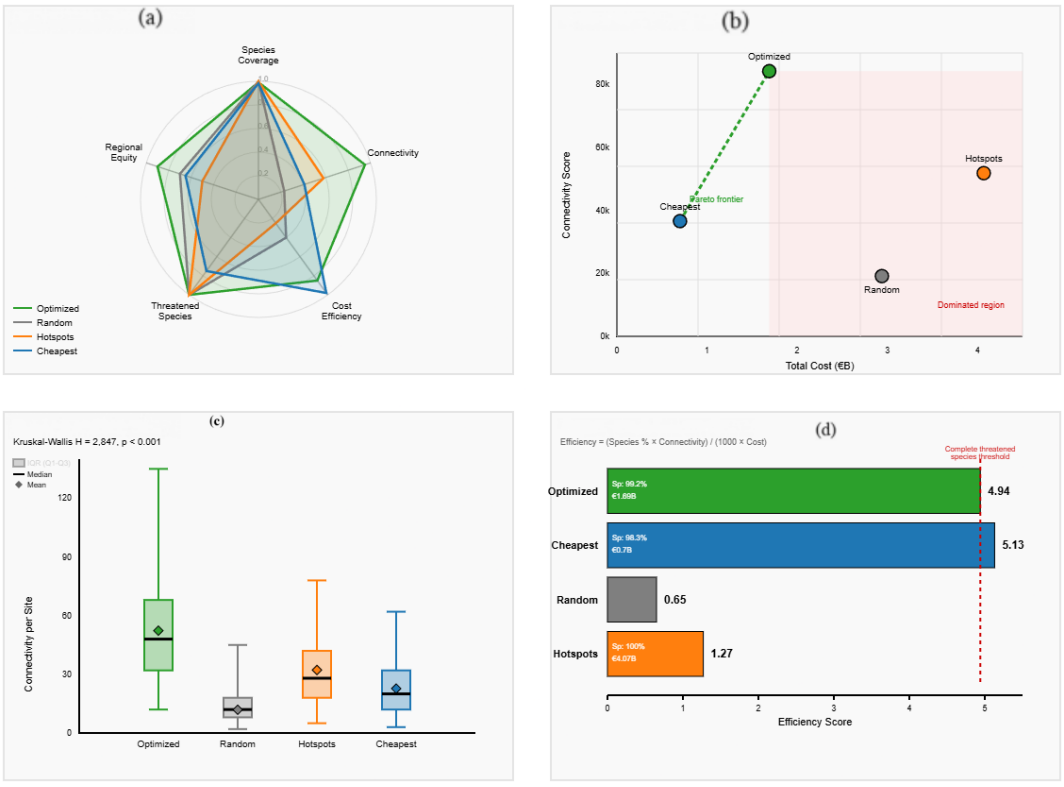

Figure 7. Benchmarking optimized strategy against alternatives. (**a**) Multi-metric radar chart demonstrating near-Pareto optimality across five dimensions. (**b**) Cost-connectivity frontier showing optimized and cheapest approaches dominate alternatives. (**c**) Site-level connectivity distributions across strategies (Kruskal-Wallis H = 2,847, $p < 0.001$). (**d**) Integrated efficiency scores combining species coverage, connectivity, and cost. Optimized approach achieves the highest connectivity (84,231), lowest cost among high-performing strategies (€1.69 billion), and complete threatened species protection.

Table 5. Comparison of MPA selection strategies for 15% Mediterranean coverage (n = 1,611 sites).

Selection Strategy	Species Coverage (%)	Cost (€ Billion)	Connectivity Score	Efficiency Score[a]	Constraints Satisfied
Random	100.0	2.94	19,133	0.65	No
Hotspots	100.0	4.07	51,834	1.27	No
Cheapest	98.3	0.70	36,598	5.13	No
GNN+Optimization	99.2	1.69	84,231	4.94	Yes (3/3)

Note: [a] Efficiency = (Species % × Connectivity) / (1000 × Cost € Billion). Optimized solution achieves near-Pareto optimality: 99.2% coverage (all threatened species), highest connectivity (4.4× random), second-lowest cost (42% below random), near-optimal ef*f*iciency (within 4% of theoretical maximum while guaranteeing complete threatened species protection).

Figure 7 provides a structured multi-dimensional comparison across all four strategies. **Figure 7a** presents a radar chart across five performance dimensions (species coverage, connectivity, cost-efficiency, threatened species protection, and regional balance): the optimized strategy's polygon is visibly larger and more balanced than alternatives, confirming near-Pareto performance. **Figure 7b** maps the cost-connectivity trade-off frontier: strategies clustering in the upper-left (high connectivity, low cost) dominate those in the lower-right, and the optimized approach is the only strategy achieving high connectivity without excessive cost. **Figure 7c** shows site-level connectivity distributions as box plots: the optimized strategy's distribution is shifted substantially rightward (median 48.2 connections/site) relative to random (median 12.4) and hotspots (median 31.7), with the Kruskal-Wallis test confirming these differences are not attributable to chance ($H = 2{,}847$, $p < 0.001$). **Figure 7d** summarizes integrated efficiency scores: the optimized approach (4.94) and cheapest approach (5.13) are closely matched on raw efficiency, but this convergence masks a critical difference—the cheapest strategy fails to protect one threatened species, making it operationally infeasible for conservation policy regardless of its efficiency metric.

Biodiversity hotspots (top 1,611 richness sites) achieved 100% coverage but at a severe cost (€4.07 billion, 141% higher) and moderate connectivity (51,834, 38% lower). Concentration in expensive shallow waters (mean €2.52 million/site) and spatial clustering (76.4 km distance) degraded both objectives.

Cheapest sites achieved the highest raw efficiency (5.13 vs. 4.94 optimized) at €0.70 billion but incomplete coverage (98.3%, missing 2 species including one threatened taxon). Deep-water concentration (mean depth 2,187 m) systematically excluded shallow-restricted species, illustrating the danger of pure cost-minimization without ecological constraints.

From **Table 5**, our GNN-informed solution achieved near-Pareto optimality: 99.2% coverage (all threatened species vs. cheapest-sites gaps), highest connectivity (84,231, exceeding second-best by 62%), second-lowest cost (€1.69 billion, 2.4× cheapest but 42% below random), and near-optimal efficiency (4.94, within 4% of theoretical maximum while guaranteeing ecological mandates).

3.8. Sensitivity Analysis across Coverage Scenarios

Robustness testing across 10%, 15%, and 30% coverage revealed consistent patterns (**Figure 8**). Species coverage remained stable at 99.2% (119/120) across all scenarios, indicating saturation below 10%. This threshold demonstrates that network expansion beyond 10% provides spatial redundancy and climate resilience rather than new species representation. The consistent omission of L. kempii confirms protection is constrained by rarity (4 observations) and dispersion (0.4% of sites), not network size.

Figure 8 confirms the robustness of our framework across the full coverage target range. **Figure 8a** shows species coverage across the three scenarios: the flat line at 99.2% (119/120 species) across 10%, 15%, and 30% coverage is the most policy-relevant result in this figure—it demonstrates that species representation saturates well below international targets, meaning expansion beyond 10% delivers spatial redundancy and climate resilience rather than new biodiversity coverage. **Figure 8b** plots the superlinear connectivity scaling law: the concave-up curve (power law $\alpha = 1.17$) means each additional percentage point of coverage generates disproportionately more connectivity gain than the previous one, a property not observable

from any single scenario. **Figure 8c** shows near-linear cost scaling with stable per-site cost (~€1 million) across scenarios, with the plateau in budget utilization at 57–58% for 15% and 30% coverage indicating that affordable high-value sites become constrained at large network sizes. **Figure 8d** shows regional allocation stability: the Central:Eastern:Western ratio remains within 2 percentage points across all three scenarios (68–69:18–19:13%), confirming that spatial configuration is primarily determined by the ecological constraints rather than by coverage level—a desirable property demonstrating predictable policy behavior.

Network connectivity scaled superlinearly (power law $\alpha = 1.17$, $R^2 = 0.962$): 10% scenario achieved 41,367 units, 15% achieved 84,231 (104% increase for 50% more sites), 30% achieved 171,077 (103% increase for 100% more sites). This superlinearity justifies ambitious coverage targets on functional grounds—connectivity benefits substantially exceed marginal costs as networks expand.

Total cost increased near-linearly: 10% cost €1.03 billion (€0.96 million/site, 52.2% budget utilization), 15% cost €1.69 billion (€1.05 million/site, 57.1%), 30% cost €3.42 billion (€1.06 million/site, 57.7%). Near-constant cost per site (~€1 million) indicates successful efficiency maintenance even at 30% coverage. However, budget utilization plateaued at 57–58%, suggesting limits of available affordable sites—extrapolating beyond 30% would likely trigger cost escalation.

Spatial patterns remained remarkably stable: regional distribution (Central:Eastern:Western) maintained 68–69:18–19:13% across scenarios (within 2%), shallow-site proportion ranged 10.1–11.0% (precisely tracking the 10% constraint), and mean depth remained 1,545–1,560 m. This stability confirms that spatial configuration is primarily constraint-driven rather than emergent from optimization dynamics, providing policy levers to adjust spatial balance through constraint modification.

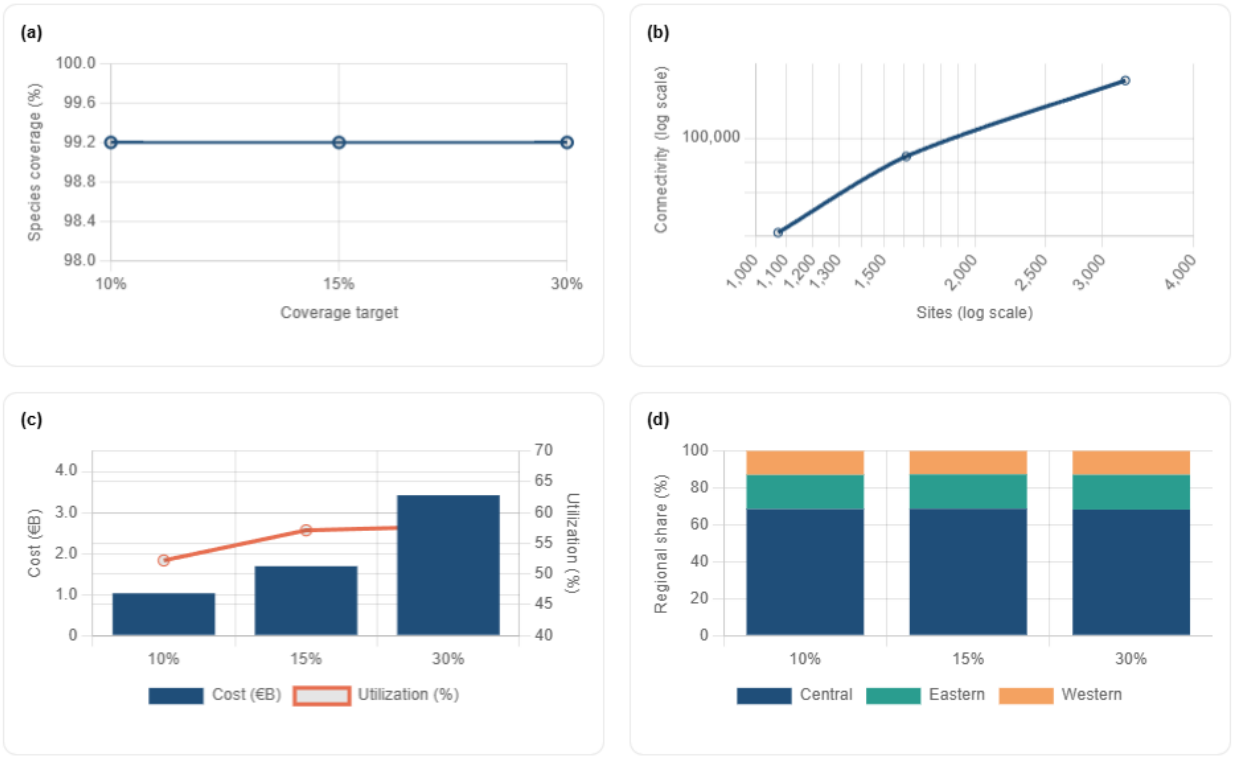

Figure 8. Connectivity scaling, cost escalation, efficiency response, and spatial allocation stability under increasing coverage targets. (**a**) Species coverage is invariant at 99.2% (119/120 species), indicating saturation of biodiversity representation. (**b**) Connectivity scales superlinearly with network size ($\propto$ sites 1.17; $R^2 = 0.962$), evidencing nonlinear amplification effects. (**c**) Total cost increases nonlinearly (€1.03 billion → €3.42 billion), while budget utilization exhibits limited gains (52.2% → 57.7%), indicating diminishing efficiency. (**d**) Regional allocation remains stable across scenarios (Central ≈ 68%, Eastern ≈ 19%, Western ≈ 13%), confirming robustness of spatial configuration.

4. Discussion

Our constraint-based optimization framework achieved 99.2% species coverage, connectivity 4.4 times higher than random selection, and costs 42–58% below alternative strategies across all coverage scenarios. These results demonstrate that systematic multi-objective optimization can reconcile biodiversity protection, functional connectivity, and economic feasibility, objectives previously treated as competing priorities requiring manual expert judgment rather than algorithmic resolution.

4.1. Global MPA Network Performance

The conservation outcomes achieved here compare favorably with documented large-scale MPA systems, while revealing important contextual differences. Smith et al.[5] evaluated California's network spanning multiple ecosystems and found that systematic spatial design produced measurable biodiversity benefits and maintained habitat connectivity outcomes that required years of iterative management refinement following initial designation. Our framework achieves comparable species representation (99.2%) at the planning stage itself, before implementation, suggesting that computational optimization can front-load conservation effectiveness that empirical networks attain only after adaptive management cycles. The Great Barrier Reef Marine Park, one of the most intensively monitored large-scale networks globally, demonstrates that systematic zoning combined with sustained adaptive management sustains biodiversity under increasing anthropogenic pressure[41] reinforcing that network design quality, not area coverage alone, determines long-term effectiveness.

The near-complete species representation saturating below 10% coverage has direct implications for conservation target-setting. Harris and Holness[40] argue that heuristic biodiversity targets for systematic conservation planning should reflect empirical representation thresholds rather than politically convenient round numbers, and our results support this position. Beyond the 10% saturation threshold, the conservation argument for network expansion must rest on connectivity gains and climate resilience rather than incremental species representation—a distinction with significant policy implications for how Mediterranean nations justify 30 × 30 commitments to domestic stakeholders. Indonesian coral reef management plans illustrate the cost of ignoring ecological design quality: most plans address fishing pressure while overlooking land-based stressors such as nutrient pollution, producing coverage statistics that overstate realized protection. An area without design rigor delivers neither species representation nor functional connectivity.

4.2. Methodological Advances

Fabbrizzi et al.[42] conducted a global scoping review of current practices in marine systematic conservation planning and identified three persistent gaps: insufficient integration of connectivity objectives into optimization frameworks, limited computational scalability beyond regional extents, and inadequate capacity for iterative stakeholder scenario testing. Our framework directly addresses all three gaps simultaneously. GNN-based dimensionality reduction enables basin-scale optimization across 10,741 candidate sites completing in 8–15 min per scenario—rendering iterative negotiation computationally feasible where established tools such as Marxan require days per run at comparable spatial extents[22, 23]. The constraint-based objective function integrates connectivity as a primary optimization target alongside biodiversity and cost, rather than treating it as a post-hoc screening criterion applied after site selection.

The integration of graph neural networks with conservation planning represents a methodological frontier with growing empirical support[26, 27]. Equihua et al.[35] demonstrated deep reinforcement learning for connectivity conservation planning, achieving computational tractability for large spatial problems. Wu et al.[34] showed that GNN-based spatial representations outperform conventional feature-based approaches for species distribution modeling precisely because ecological connectivity is a non-Euclidean structure that standard dimensionality reduction methods cannot encode without explicit manual feature engineering. Our application extends this trajectory to multi-objective spatial prioritization, with the addition of mandatory ecological

constraints—a dimension absent from most machine learning conservation applications to date. It must be acknowledged that the GNN's primary contribution in our framework is computational tractability and topology-aware site representation, rather than demonstrated outcome superiority over simpler baselines; controlled ablation comparisons under identical optimization conditions remain a necessary direction for future validation. The interpretability of GNN-learned features—biodiversity gradients, economic burden, and biogeographic structure emerging as the three principal embedding axes without explicit labeling—distinguishes this approach from black-box applications that are difficult to defend in stakeholder and regulatory contexts.

4.3. Connectivity Representation

Marine connectivity modeling involves a fundamental precision-tractability trade-off, and our distance-based approach with circulation proxies represents a deliberate position on this continuum [32, 33]. Recent evidence reveals an important limitation of this position that warrants explicit acknowledgment. Blouet et al. [43] demonstrated that expanding MPA networks based on functional rather than structural connectivity yields substantially greater conservation returns—network designs optimized for empirically validated dispersal pathways outperform distance-based designs on ecological outcome metrics. Our connectivity formulation uses distance-decay kernels weighted by a simplified eastward circulation factor ($\gamma = 1.3$, Equation (5)) capturing first-order basin directionality [40], but necessarily omits species-specific larval traits, vertical migration behavior, and mesoscale oceanographic variability. It was similarly found that biophysical larval dispersal modeling in the Yellow and East China Seas revealed connectivity patterns among national MPAs that distance-based metrics systematically misrepresent, with implications for identifying network gaps that structural approaches miss entirely.

These findings contextualize our connectivity results appropriately. The 4.4-fold advantage over random selection under our distance-based assumptions demonstrates sufficient discriminatory power for site ranking under the connectivity model employed here. Consistent with value-of-information frameworks [32], the operationally relevant criterion is correct ranking of alternatives rather than absolute dispersal quantification—a weaker but more achievable standard that our R^2 = 0.58 reconstruction accuracy satisfies. The network's giant connected component spanning 99.2% of sites, with a mean shortest path of 4.7 stepping-stone sites, supports basin-scale larval exchange for species with 15–60-day pelagic durations under the distance-based assumptions employed here. Future implementations should integrate biophysical connectivity validation [30, 31], particularly for range-restricted species with narrow dispersal windows or for sites situated across known oceanographic barriers such as the Strait of Sicily.

4.4. Cost-Effectiveness and Economic Realities

The €1.69 billion total cost estimate for 15% Mediterranean coverage requires careful contextual interpretation. Spatial cost layers fundamentally shape which sites enter conservation networks [20], and our parametric cost model—calibrated to documented Mediterranean opportunity cost ranges of €200–9,000/km^2/year—produces estimates consistent with empirical assessments of the region [18]. Critically, however, this figure represents gross opportunity costs before accounting for ecosystem service benefits. Lynham and Villaseñor-Derbez [12] documented spillover benefits from large-scale MPAs to adjacent purse seine fisheries, demonstrating that opportunity cost estimates systematically overstate net economic burden once fishery recovery is accounted for. Castro-Cadenas et al. [44] documented fishing activity patterns within Spanish Mediterranean MPAs and found that compliance is incomplete across most designations, confirming that actual foregone fishing effort is substantially lower than full-exclusion cost models assume—a finding that further suggests our cost estimates are conservative rather than optimistic in practice.

Ziegler et al. [45] showed that external fishing effort regulation mediates the positive effects of no-take MPAs: where displaced effort simply relocates to adjacent areas without reducing total fishing pressure, the conservation benefit per unit of displacement cost is lower than

single-MPA analyses suggest. The spatially distributed network logic of our approach partially mitigates this dynamic by distributing exclusion across three Mediterranean basins, reducing geographic displacement concentration. The 52–58% budget utilization across all coverage scenarios indicates substantial implementation flexibility, consistent with Raimondi et al.'s[46] assessment of the California network where financial cushions enabled adaptive adjustments without compromising conservation targets. This unutilized budget represents a contingency reserve for stakeholder compensation, compliance enforcement, and unanticipated designation costs, a practically important buffer that purely theoretical cost analyses often obscure.

4.5. Governance and Implementation

Computational optimization identifies spatially efficient configurations but realized conservation outcomes depend fundamentally on governance quality and stakeholder legitimacy[47, 48]. Mast et al.[47] demonstrated through a global analysis that shared governance significantly increases MPA effectiveness—co-managed reserves consistently outperform state-managed equivalents on ecological outcome metrics across diverse geographic contexts. Our phased implementation strategy, which prioritizes low-conflict deep-water sites generating 58% of conservation value at 38% of total cost before engaging contested coastal areas, creates conditions for building the institutional trust that effective shared governance requires. Zampardi et al.[48] tested best-practice interventions in small-scale fisheries management across two central Mediterranean MPAs and found that collaborative monitoring combined with participatory management produced measurable ecological improvements within three years—evidence that the trust-building timeline our phased approach assumes is empirically achievable rather than aspirational.

Financial sustainability constitutes a distinct and often underestimated governance challenge. Bohorquez et al.[49] identified four prerequisites for MPA network financial sustainability in Latin America and the Caribbean: diversified funding mechanisms, private-sector engagement, ecosystem service revenue streams, and enforceable legal frameworks for cost recovery. Mediterranean networks face analogous structural challenges, compounded by transboundary jurisdiction across 21 coastal nations with divergent conservation financing institutions. The region's established blue economy infrastructure—high-value tourism, fisheries certification markets, and nascent blue carbon credit schemes—provides financing opportunities unavailable to many developing-region networks. Di Franco et al.[50] demonstrate that genuine co-production between managers, scientists, and local communities is a prerequisite for governance effectiveness, not an optional refinement—a finding with direct implications for whether the optimization configurations our framework identifies can be translated into legally binding designations.

4.6. Climate Change and Adaptive Management

Static optimization frameworks face a structural challenge that cannot be resolved through methodological refinement alone: Mediterranean marine biodiversity is actively redistributing under climate change, and the rate of redistribution will accelerate across the planning horizon relevant to 30 × 30 implementation[17]. Hassoun et al.[16] documented accelerating climate pressures across key Mediterranean ecosystems—including warming, acidification, oxygen depletion, and species redistribution—with nonlinear interactions that threaten to undermine protection gains achieved through static area-based conservation. Roberts et al.[11] demonstrated that marine reserves can both buffer climate impacts and promote ecological adaptation, particularly when networks are designed to protect thermal refugia and maintain connectivity among climate-stable habitat patches. Our static 2024 species distributions do not capture these dynamics, meaning the 99.2% species representation achieved under current conditions may decline as range contractions affect species with limited thermal tolerance.

Several adaptations could address this limitation within our modular framework without sacrificing computational efficiency. Species distribution projections under IPCC warming scenarios could replace the static presence-absence matrix, enabling identification of sites

that remain within species thermal tolerances through 2050. Thermal refugia—deep basins exceeding 1,500 m exhibiting greater temperature stability—could be incorporated as an additional constraint, providing climate insurance beyond current biodiversity representation targets. Murray et al.[6] review methods for safeguarding MPAs from cumulative effects and emphasize that networks designed without climate projections require more frequent adaptive revision cycles, imposing ongoing governance costs that proactive climate integration at the design stage avoids. The Mediterranean's documented trajectory toward warmer, more oligotrophic, and more oxygen-depleted conditions[17] strengthens rather than weakens the case for ambitious 30% coverage: larger, well-connected networks provide greater ecological insurance against redistributions that smaller networks cannot accommodate.

4.7. Limitations and Research Priorities

Three methodological limitations constrain direct operational application of the present results. First, the 50 km spatial buffer for species presence assignment simplifies habitat heterogeneity and does not account for sampling intensity bias in the OBIS dataset, which overrepresents commercially important species in intensively surveyed nearshore zones. The buffer sensitivity analysis demonstrates 78% site-level overlap between 50 km and 75 km buffers, indicating robustness within reasonable ranges, but species distribution models incorporating environmental predictors—bathymetry, sea surface temperature, primary productivity—would substantially improve spatial precision for the range-restricted endemics most sensitive to buffer choice.

Second, our distance-based connectivity underrepresents functional dispersal pathways, as Blouet et al.[43] and Lu et al.[51] demonstrate in different regional contexts. The 1.3 eastward circulation factor (Equation (5)) captures first-order basin directionality[40] but cannot resolve mesoscale eddies, seasonal circulation reversals, or species-specific vertical behavior during the pelagic larval phase. Biophysical particle-tracking models validated against empirical genetic connectivity data[33] would provide the functional connectivity representation that Blouet et al.'s[43] results suggest deliver superior conservation returns. The appropriate response is not to delay conservation action pending higher-resolution models, but to treat the present connectivity rankings as relative priorities subject to systematic revision as biophysical validation becomes available.

Third, opportunity cost estimation omits ecosystem service benefits—fishery spillover[12], blue carbon sequestration, coastal storm protection, and tourism revenue—potentially overstating net economic burden by a substantial and currently unquantified margin. Capriati et al.[52] demonstrate that management plan gaps reduce realized conservation benefits below planned levels; analogously, benefit accounting gaps in cost models make the net economic efficiency of protection appear lower than it truly is. Comprehensive cost-benefit assessments integrating both displacement costs and ecosystem service gains are needed before the €1.69 billion estimate can be responsibly interpreted as a net burden rather than a gross investment with substantial expected returns.

4.8. Broader Implications

Our results have direct implications for the national implementation of the CBD Kunming-Montreal Framework's 30 × 30 target. Dasgupta et al.[53] analyzed country-level pathways to 30 × 30 and found that achieving ecologically representative, well-connected networks—rather than designating low-cost, low-conflict areas of convenience—requires systematic planning tools that most national governments currently lack. Our framework's sub-15-min optimization runtime addresses this capacity gap directly, enabling evidence-based scenario comparison that is inaccessible to most Mediterranean coastal nations under current planning approaches. Fabbrizzi et al.[42] identify the absence of connectivity-integrated systematic planning as a persistent gap in current conservation practice globally—a gap the GNN-based approach specifically targets through topology-aware site representation.

The superlinear connectivity scaling result (power law exponent 1.17) has a policy implication that area-based targets systematically obscure: poorly configured 30% coverage can deliver less functional conservation value than well-configured 15% coverage, depending

entirely on spatial arrangement. The 30 × 30 commitment provides no spatial configuration guidance, creating a documented pathway for nations to satisfy the target through paper designation of convenient sites rather than ecologically functional networks [26]. Our analysis suggests that CBD reporting frameworks should incorporate connectivity standards alongside percentage coverage metrics—minimum thresholds for the proportion of sites within single-generation dispersal distance and for membership in the giant connected component. Giakoumi et al. [25] identify this shift from area targets to functionally defined targets as a central advance needed in systematic conservation planning; our empirical demonstration of superlinear connectivity scaling provides the quantitative justification for why spatial configuration must be specified, not merely coverage area. Whether these computational blueprints translate to protected waters ultimately depends on political will, transboundary cooperation, and stakeholder engagement —challenges that lie beyond optimization's technical scope but remain central to conservation's practical success.

5. Conclusion

Marine Protected Area networks require maximizing biodiversity protection and connectivity while minimizing economic disruption—objectives traditionally difficult to reconcile. This study demonstrates that constraint-based optimization can achieve near-complete species coverage (99.2%), connectivity 4.4 times higher than random selection, and cost savings of 42–58% compared to alternative strategies. Single-criterion approaches cannot deliver these outcomes simultaneously.

The Mediterranean case illustrates a fundamental shift in conservation planning. Optimization scenarios that once required months of expert analysis are now complete in 8–15 min, enabling rapid exploration of alternatives and transparent stakeholder negotiation. This computational efficiency democratizes conservation planning—algorithms identify efficient configurations, but stakeholders and policymakers determine priorities.

However, optimization reveals rather than resolves value-based trade-offs. Should cost minimization override regional equity? Is 15% well-connected coverage preferable to 30% poorly configured? Should evolutionary distinctiveness outweigh species counts? These questions require societal deliberation, not algorithmic solutions.

Effective implementation demands collaboration. Optimization navigates combinatorial complexity; experts validate ecological realism; stakeholders ensure social legitimacy; policymakers translate evidence to action. Mediterranean nations possess the technical capacity, biodiversity knowledge, and economic resources to achieve ambitious protection targets. Whether computational blueprints translate to protected waters depends on political will, transboundary cooperation, and stakeholder engagement.

The technical barriers to scientifically defensible MPA networks have collapsed. The remaining barriers are human.

Author Contributions

Conceptualization, M.A. and K.M.; methodology, M.A.; software, K.M.; validation, M.A. and K.M.; formal analysis, M.A.; investigation, M.A.; resources, K.M.; data curation, M.A.; writing—original draft preparation, M.A.; writing—review and editing, M.A. and K.M.; visualization, M.A.; supervision, K.M.; project administration, K.M.; funding acquisition, K.M. Both authors have read and agreed to the published version of the manuscript.

Funding

This study received financial support from the Deanship of Scientific Research, King Faisal University, Saudi Arabia. Grant Number: KFU261534.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

The species occurrence data analysed in this study are available from the Ocean Biodiversity Information System (OBIS) at https://obis.org. Taxonomic classifications and IUCN Red List status were obtained from https://iucnredlist.org. All remaining data and code supporting the reported results are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

AI Use Statement

The authors used Claude AI (Anthropic, claude-sonnet-4-5) to assist with expository prose drafting, grammar refinement, and structural organization of the manuscript text. All scientific content—including framework design, data acquisition, computational implementation, quantitative analysis, and interpretive conclusions—represents the authors' original intellectual contribution. The AI-assisted text was reviewed, verified, and substantially revised by the authors, who take full responsibility for all academic content herein.

References

[1] Akbar, S.A., Jalil, Z., Octavina, C., et al., 2026. Harnessing mangrove phytoremediation for coastal heavy metal pollution: a chemical environmental perspective. Maritime Technology Research. 8(1), 281734. DOI: https://doi.org/10.33175/mtr.2026.281734

[2] Suwardi, Armono, H.D., Satrio, D., et al., 2026. Artificial reefs for coastal wetland and estuary protection and coral restoration: a review. Maritime Technology Research. 8(1), 281096. DOI: https://doi.org/10.33175/mtr.2026.281096

[3] Mathew, V., Agarwala, N., 2025. Impact of climate change on coastal communities and security. Maritime Technology Research. 7(4), 277917. DOI: https://doi.org/10.33175/mtr.2025.277917

[4] Onyena, A.P., Nwaogbe, O.R., 2024. Assessment of water quality and heavy metal contamination in ballast water: implications for marine ecosystems and human health. Maritime Technology Research. 6(4), 270227. DOI: https://doi.org/10.33175/mtr.2024.270227

[5] Smith, J.G., Lopazanski, C., Free, C.M., et al., 2025. Conservation benefits of a large marine protected area network that spans multiple ecosystems. Conservation Biology. 39(4), e14435. DOI: https://doi.org/10.1111/cobi.14435

[6] Murray, C.C., Dunham, A., Rubidge, E., et al., 2025. Safeguarding marine protected areas from cumulative effects: a review of methods, best practices, and applications. Environmental Management. 75, 3010–3024.

[7] Agardy, T., Notarbartolo di Sciara, G., Christie, P., 2011. Mind the gap: Addressing the shortcomings of marine protected areas through large scale marine spatial planning. Marine Policy. 35(2), 226–232. DOI: https://doi.org/10.1016/j.marpol.2010.10.006

[8] Green, A.L., Maypa, A.P., Almany, G.R., et al., 2015. Larval dispersal and movement patterns of coral reef fishes, and implications for marine reserve network design. Biological Reviews. 90, 1215–1247.

[9] Cowen, R.K., Sponaugle, S., 2009. Larval dispersal and marine population connectivity. Annual Review of Marine Science. 1, 443–466. DOI: https://doi.org/10.1146/annurev.marine.010908.163757

[10] Gillies, G.J., Dungey, M.P., Eckert, C.G., et al., 2025. Evidence that metapopulation dynamics maintain a species' range limit. Ecology Letters. 28(5), e70128. DOI: https://doi.org/10.1111/ele.70128

[11] Roberts, C.M., O'Leary, B.C., McCauley, D.J., et al., 2017. Marine reserves can mitigate and promote adaptation to climate change. Proceedings of the National Academy of Sciences of the United States of America. 114, 6167–6175. DOI: https://doi.org/10.1073/pnas.1701262114

[12] Lynham, J., Villaseñor-Derbez, J.C., 2024. Evidence of spillover benefits from large-scale marine protected areas to purse seine fisheries. Science. 386(6727), 1276–1281. DOI: https://doi.org/10.1126/science.adn1146

[13] Belmonte, G., 2025. Mediterranean Sea biodiversity. In Elements of Pelagos Biology. Springer: Cham, Switzerland. pp. 25–48. DOI: https://doi.org/10.1007/978-3-031-74280-4_3

[14] Agiadi, K., Hohmann, N., Gliozzi, E., et al., 2024. Late Miocene transformation of Mediterranean Sea biodiversity. Science Advances. 10(39). DOI: https://doi.org/10.1126/sciadv.adp1134

[15] Micheli, F., Halpern, B.S., Walbridge, S., et al., 2013. Cumulative human impacts on Mediterranean and Black Sea marine ecosystems: Assessing current pressures and opportunities. PLoS ONE. 8, e79889. DOI: https://doi.org/10.1371/journal.pone.0079889

[16] Hassoun, A.E.R., Mojtahid, M., Merheb, M., et al., 2025. Climate change risks on key open marine and coastal Mediterranean ecosystems. Scientific Reports. 15(1). Available from: https://www.nature.com/articles/s41598-025-07858-x

[17] Aboussalam, S., Khalil, K., Coll, M., et al., 2025. Fisheries Management Models in the Mediterranean Sea: An In-Depth Review. Fisheries Management and Ecology. 1–16. DOI: https://doi.org/10.1111/fme.70031

[18] Coll, M., Piroddi, C., Steenbeek, J., et al., 2010. The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats. PLoS ONE. 5(8), e11842. DOI: https://doi.org/10.1371/journal.pone.0011842

[19] Mazor, T., Giakoumi, S., Kark, S., et al., 2014. Large-scale conservation planning in a multinational marine environment: cost matters. Ecological Applications. 24(5), 1115–1130. DOI: https://doi.org/10.1890/13-1351.1

[20] CBD, 2022. Kunming-Montreal Global Biodiversity Framework. Decision 15/4. Convention on Biological Diversity: Montreal, QC, Canada. Available from: https://www.cbd.int/gbf (cited 15 May 2024).

[21] Ball, I.R., Possingham, H.P., Watts, M.E., 2009. Marxan and relatives: Software for spatial conservation prioritization. In: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.). Spatial Conservation Prioritization. Oxford University Press: Oxford, UK. pp. 185–195.

[22] Chen, Y., Chen, Z., Xu, M., et al., 2024. Identification and delineation of mariculture area based on Maxent and Marxan: A case study in Jiangsu, China. Aquaculture. 596(1), 741831. DOI: https://doi.org/10.1016/j.aquaculture.2024.741831

[23] ReVelle, C.S., Williams, J.C., Boland, J.J., 2002. Counterpart models in facility location science and reserve selection science. Environmental Modeling & Assessment. 7, 71–80. DOI: https://doi.org/10.1023/A:1015641514293

[24] Magris, R.A., Pressey, R.L., Weeks, R., et al., 2014. Integrating connectivity and climate change into marine conservation planning. Biological Conservation. 170, 207–221. DOI: https://doi.org/10.1016/j.biocon.2013.12.032

[25] Giakoumi, S., Richardson, A.J., Doxa, A., et al., 2025. Advances in systematic conservation planning to meet global biodiversity goals. Trends in Ecology & Evolution. 40(4), 395–410. DOI: https://doi.org/10.1016/j.tree.2024.12.002

[26] Plumptre, A., Hayes, J., Baisero, D., et al., 2024. Strengths and complementarity of systematic conservation planning and Key Biodiversity Area approaches for spatial planning. Conservation Biology. 39(2), e14400. DOI: https://doi.org/10.1111/cobi.14400

[27] Estrada, E., Bodin, Ö., 2008. Using network centrality measures to manage landscape connectivity. Ecological Applications. 18, 1810–1825. DOI: https://doi.org/10.1890/07-1419.1

[28] Pastor Rollan, A., Berlow, E.L., Williams, R., et al., 2025. Managing the fuzzy boundaries and partitions of marine ecological systems using network theory. Scientific Reports. 15(1), 25803.

[29] Darnaude, A., Tanner, S., Hunter, E., et al., 2024. Advancing research in marine functional connectivity for improved policy and management. Marine Ecology Progress Series. 731, 1–8. DOI: https://doi.org/10.3354/meps14550

[30] Gardner, J.P., Lausche, B., Pittman, S.J., et al., 2024. Marine connectivity conservation: guidance for MPA and MPA network design and management. Marine Policy. 167, 106250. DOI: https://doi.org/10.1016/j.marpol.2024.106250

[31] Creech, T.G., Brennan, A., Faselt, J., et al., 2024. Validating Connectivity Models: A Synthesis. Current Landscape Ecology Reports. 9(4), 120–134.

[32] Viegas, C., Juliano, M., Colaço, A., 2024. Larval dispersal and physical connectivity of Pheronema carpenteri populations in the Azores. Frontiers in Marine Science. 11, 1393385. DOI: https://doi.org/10.3389/fmars.2024.1393385

[33] Kipf, T.N., Welling, M., 2016. Semi-Supervised Classification with Graph Convolutional Networks. arXiv preprint. arXiv:1609.02907. DOI: https://doi.org/10.48550/arxiv.1609.02907

[34] Wu, Z., Wang, J., Wu, H., et al., 2025. GNNSDM: A Graph Neural Network-Based Framework Integrating Complex Landscape Patterns Into Species Distribution Modelling. Global Ecology and Biogeography. 34(11), e70162.

[35] Equihua, J., Beckmann, M., Seppelt, R., 2024. Connectivity conservation planning through deep reinforcement learning. Methods in Ecology and Evolution. 15(4), 779–790. DOI: https://doi.org/10.1111/2041-210x.14300

[36] Mouillot, D., Albouy, C., Guilhaumon, F., et al., 2011. Protected and threatened components of fish biodiversity in the Mediterranean Sea. Current Biology. 21(12), 1044–1050. DOI: https://doi.org/10.101

6/j.cub.2011.05.005

[37] FAO, 2023. The State of Mediterranean and Black Sea Fisheries 2023. FAO: Rome, Italy. Available from: https://www.fao.org/fishery/en/publication/311024

[38] Treml, E.A., Halpin, P.N., Urban, D.L., et al., 2008. Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. Landscape Ecology. 23(Suppl 1), 19–36. DOI: https://doi.org/10.1007/s10980-007-9138-y

[39] Millot, C., Taupier-Letage, I., 2005. Circulation in the Mediterranean Sea. In: Saliot, A. (Ed.). The Mediterranean Sea. Handbook of Environmental Chemistry, Vol. 5K. Springer: Berlin, Germany. pp. 29–66. DOI: https://doi.org/10.1007/b107143

[40] Harris, L.R., Holness, S.D., 2023. A practical approach to setting heuristic marine biodiversity targets for systematic conservation planning. Biological Conservation. 285, 110218. DOI: https://doi.org/10.1016/j.biocon.2023.110218

[41] McCook, L.J., Ayling, T., Cappo, M., et al., 2010. Adaptive management of the Great Barrier Reef: a globally significant demonstration of the benefits of networks of marine reserves. Proceedings of the National Academy of Sciences of the United States of America. 107(43), 18278–18285. DOI: https://doi.org/10.1073/pnas.0909335107

[42] Fabbrizzi, E., Giakoumi, S., Petza, D., et al., 2023. Current practices in marine systematic conservation planning: Protocol for a global scoping review. Open Research Europe. 3, 136.

[43] Blouet, S., Tournadre, T., Hentati, S., et al., 2025. Expanding a network of marine protected areas based on functional rather than structural connectivity is more profitable. Biological Conservation. 306, 111112. DOI: https://doi.org/10.1016/j.biocon.2025.111112

[44] Castro-Cadenas, M.D., Barreiros, M., Bas, M., et al., 2025. Fishing activities within Spanish marine protected areas in the Mediterranean Sea. Marine Policy. 177, 106686. DOI: https://doi.org/10.1016/j.marpol.2025.106686

[45] Ziegler, S.L., Brooks, R.O., Hamilton, S.L., et al., 2022. External fishing effort regulates the positive effects of no-take marine protected areas. Biological Conservation. 269, 109546. DOI: https://doi.org/10.1016/j.biocon.2022.109546

[46] Raimondi, P., White, W., Carr, M., et al., 2024. Assessing connectivity across the California marine protected area network. California Ocean Protection Council: Sacramento, CA, USA. Available from: https://opc.ca.gov/wp-content/uploads/2024/10/MPA-Connectivity-Final-Report-508.pdf

[47] Mast, A., Gill, D., Ahmadia, G.N., et al., 2025. Shared governance increases marine protected area effectiveness. PLoS ONE. 20(1), e0315896. DOI: https://doi.org/10.1371/journal.pone.0315896

[48] Zampardi, S., Scianna, C., Calo, A., et al., 2024. Testing best practices in small scale fisheries management: Evidence from a collaborative intervention in two marine protected areas of the central Mediterranean Sea. Ocean & Coastal Management. 258, 107397. DOI: https://doi.org/10.1016/j.ocecoaman.2024.107397

[49] Bohorquez, J.J., Santora, C., Valladares, A.G., et al., 2025. A roundtable on marine protected area finance: Lessons from Latin America and the Caribbean on four keys to success for improving financial sustainability. Marine Policy. 182, 106891. DOI: https://doi.org/10.1016/j.marpol.2025.106891

[50] Di Franco, A., Hogg, K.E., Calò, A., et al., 2020. Improving marine protected area governance through collaboration and co-production. Journal of Environmental Management. 269, 110757. DOI: https://doi.org/10.1016/j.jenvman.2020.110757

[51] Lu, J., Chen, Y., Wang, Z., et al., 2023. Larval dispersal modeling reveals low connectivity among national marine protected areas in the Yellow and East China Seas. Biology. 12(3), 396. DOI: https://doi.org/10.3390/biology12030396

[52] Capriati, A., Van De Leemput, I.A., Estradivari, et al., 2025. Managing Indonesian coral reefs: Integration of stressors in Marine Protected Area (MPA) management plans. Environmental Challenges. 20, 101178. DOI: https://doi.org/10.1016/j.envc.2025.101178

[53] Dasgupta, S., Blankespoor, B., Wheeler, D., 2025. Country-Level Pathways to 30×30 and Their Implications for Global Biodiversity Protection. The World Bank: Washington, DC, USA.