Designing Floating Reef Habitats for Coastal Aquaculture: A Data-Driven Approach to Balancing Structural Stability and Marine Biodiversity

Rosalie Heidrich
School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR, USA.
heidrich331@oregonstate.edu

Leone Gargallo
Department of Computer Science, Binghamton University, Binghamton, NY, USA.
contactleone@binghamton.edu

Rosaria Scalfaro
School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
rosariascalfaro@uc.edu

Annemarie Fliegner
Department of Computer Science, University of Central Florida, Orlando, FL, USA.
annemariemail@ucf.edu

Abstract

Floating reef habitats offer a means of adding structural complexity and ecological function to coastal aquaculture systems without occupying additional benthic area, but their design is constrained by coupled hydrodynamic, mooring, material, and ecological requirements. Existing artificial-reef practice tends to treat structural safety and biodiversity enhancement as separate problems, which can produce overly conservative modules with limited habitat value or ecologically elaborate modules with unacceptable loading and maintenance burdens. This manuscript develops a data-driven design framework for floating reef habitats in coastal aquaculture that formalizes the design problem as a reliability-constrained, multi-objective optimization under uncertainty. The framework represents each module through a design vector covering displaced volume, projected area, mooring geometry, module spacing, habitat complexity, substrate rugosity, and siting. Structural performance is described through wave loading, buoyancy equilibrium, mooring tension, and fatigue-related reliability constraints, while biodiversity performance is represented as an expected, taxa-weighted function of habitat complexity, substrate condition, spacing, and connectivity. Bayesian surrogate models and global sensitivity analysis are used to link sparse field, laboratory, and numerical information to design objectives, and a Pareto-based search identifies non-dominated designs rather than a single nominal optimum. The central insight is that biodiversity gains from added complexity are not separable from drag, buoyancy, and mooring penalties; therefore, effective floating reef design must negotiate these trade-offs explicitly and update them as monitoring data accumulate. The framework provides a technically defensible basis for designing modular floating habitats that are both safer to moor and more ecologically purposeful in aquaculture settings.

Keywords

floating artificial reefs; coastal aquaculture; structural reliability; habitat complexity; surrogate modeling; multi-objective optimization

1. Introduction

Coastal aquaculture is expanding into environments where operations must coexist with sensitive habitats, competing marine uses, and variable hydrodynamic exposure. Artificial reefs have long been used to modify local habitat structure, enhance fish aggregation, and support rehabilitation objectives, yet their design has often followed site-specific experience rather than a formal treatment of coupled engineering and ecological performance [1]. Evaluation frameworks for artificial reefs have emphasized the need to relate design attributes to ecological outcomes and to account for the age and developmental trajectory of deployed structures [2]. Structural features, material choice, and surface texture influence colonization and community assembly, but the evidence also shows that artificial reefs do not automatically mimic natural reef communities [3]. These observations make it insufficient to treat habitat value as a simple by-product of adding hard substrate.

Ecological research has established that structural complexity is a major determinant of reef-associated biodiversity, influencing refuge availability, predator-prey interactions, and the diversity of microhabitats available to fish and invertebrates [4]. For reef fish assemblages, the relative importance of live coral cover, habitat complexity, and structural characteristics varies among taxa and life stages, which means that a single complexity metric cannot represent all ecological objectives [5]. In parallel, engineering work on marine artificial structures has called for multifunctional designs that incorporate ecological objectives into structural planning rather than retrofitting habitat features after stability calculations are complete [6]. Floating reef habitats sharpen this requirement because buoyancy, mooring loads, and habitat geometry are directly coupled: additional crevice volume, surface rugosity, or module density may improve refuge value while increasing drag, mass, and inspection burden.

The specific design gap addressed here is the absence of an integrated formulation that treats floating reef habitats for coastal aquaculture as reliability-constrained ecological engineering systems. Fixed artificial reefs are usually governed by placement, settlement, and benthic interactions. Floating modules introduce a different failure space: mooring overload, fatigue at connections, loss of freeboard or buoyancy reserve, excessive motion, and biofouling-induced load growth. At the same time, floating modules can be arranged around aquaculture infrastructure to provide refuge, substrate, and connectivity without covering additional seabed. The central contribution of this manuscript is a data-driven design framework that combines a structural-hydrodynamic assessment, a biodiversity performance model, and surrogate-assisted multi-objective optimization. The framework does not presume a single best geometry; it defines how to search for non-dominated designs and how to update that search as site-specific evidence becomes available.

2. Design Problem and Analytical Framework

A floating reef habitat can be described by a design vector that separates geometric, structural, and ecological attributes. Let the vector be written as

$$x = \left(\ell, V_0, A_p, L_m, C_D, s, H_c, R_s, \theta_s\right)$$

where ℓ is a characteristic module length, V_0 is displaced volume at the design draft, A_p is projected area normal to the dominant current or wave direction, L_m is a mooring line or tether length parameter, C_D is the drag coefficient of the assembled module, s is the spacing between modules, H_c is a habitat complexity index, R_s is substrate rugosity, and θ_s represents siting variables such as depth, exposure, and orientation relative to aquaculture infrastructure. This vector is not a complete finite-element description; it is a design-level parameterization intended to support rapid comparison before detailed analysis. Each component must be bounded by operational constraints, including navigation clearance, minimum depth under the module, exclusion zones around cages and service vessels, and maximum mooring footprint.

The framework assumes that the floating reef is a moored, compliant structure composed of modules that may be rigid, semi-rigid, or 3D-printed lattice units. For slender frame members and small components relative to the wavelength, wave-induced forces can be estimated with Morison-type formulations [7,8]; for large buoyancy elements comparable to the wavelength, diffraction effects require separate treatment. Standard environmental-load guidance provides a basis for load cases, safety factors, and combination rules [9]. Recent work on 3D-printed floating artificial reefs indicates that material-efficient modular units can be fabricated with internal voids and complex surfaces, but it also highlights the need to connect fabrication freedom to load-bearing and mooring performance [10]. The design vector therefore includes both ecological geometry and structural descriptors, because printing flexibility does not remove the need to constrain strength, buoyancy, and fatigue life.

Data enter the framework from four streams. Environmental data include wave spectra, current profiles, water depth, salinity, and temperature. Structural data include material tests, connection details, buoyancy measurements, and inspection records. Ecological data include colonization surveys, fish and invertebrate assemblages, substrate coverage, and juvenile-habitat observations. Operational data include maintenance access, vessel traffic, and aquaculture production cycles. The framework is deliberately formulated so that these streams can be incomplete and heterogeneous: some variables may be represented by probability distributions, others by bounded intervals, and others by calibrated numerical models. The purpose is not to force all information into a single statistical model prematurely, but to preserve the distinction between measured quantities, model predictions, and expert judgments.

3. Data-Driven Surrogate and Multi-Objective Formulation

The coupled design problem is expensive to evaluate directly because each candidate may require hydrodynamic simulation, mooring analysis, material checks, and ecological scoring. Surrogate modeling provides a practical bridge between limited high-fidelity evaluations and broad design-space search [11]. A Bayesian surrogate, such as a Gaussian process, can represent each performance output as a distribution rather than a point prediction. This is important because the designer needs both an estimate of expected biodiversity value and a credible interval around that estimate. Bayesian calibration further allows discrepancy between a numerical model and field observations to be represented explicitly, rather than absorbed into fitted parameters [12]. In practice, calibration data may include measured mooring tensions, accelerations, biofouling mass, and colonization trajectories.

Surrogate construction should be sequential. An initial design of experiments samples the feasible region across module size, spacing, complexity, and mooring length. High-fidelity evaluations are then concentrated in regions where predicted Pareto improvement is likely or

where predictive uncertainty is high. Global sensitivity analysis is used to identify which variables dominate structural and ecological outputs before the surrogate is trusted for optimization [13]. This step matters because a design framework can be misled by a surrogate that is accurate on average but poor in the tails, which is precisely where reliability constraints operate. Sensitivity indices also help distinguish variables that should be fixed by manufacturing or operational standards from those that should remain free in the optimization.

The optimization itself is Pareto-based. A fast elitist multi-objective genetic algorithm is appropriate for non-dominated sorting when the design space is mixed and constraints are nonlinear [14]. The method does not require a single weighted objective in advance, which is advantageous because the relative value of biodiversity gain and structural utilization is a legitimate stakeholder decision. Candidate designs are retained if they are non-dominated with respect to biodiversity performance, structural utilization, and cost or maintenance burden. Constraint handling must be explicit: designs that violate mooring safety, buoyancy reserve, or navigation clearance are not allowed to enter the feasible Pareto set merely because they score well ecologically.

4. Structural-Hydrodynamic Assessment

The first structural relationship is the wave load on a slender module component. Using a Morison-type formulation,

$$F_w(t) = 0.5\rho C_D A_p |u(t)|u(t) + \rho C_M V_0 a(t)$$

where F_w(t) is the time-dependent wave force, ρ is water density, C_D is the drag coefficient, A_p is the projected area, u(t) is the water-particle velocity, C_M is the inertia coefficient, V_0 is the displaced volume, and a(t) is the local fluid acceleration. The first term represents drag and the second represents inertia. The equation is appropriate for members whose diameter is small relative to the wavelength; for large buoyancy chambers or closely spaced arrays, the added mass and diffraction problem must be solved with a radiation-diffraction model. The design implication is that ecological features protruding into the flow increase A_p and may increase C_D, so habitat complexity cannot be evaluated only by its surface area.

Buoyancy and vertical equilibrium provide the second constraint. For a module or module group,

$$\rho g_0 V_0 - M g_0 - T_v - F_w v = 0$$

where g_0 is gravitational acceleration, M is total mass including structure, ballast, and biofouling, T_v is the vertical component of mooring tension, and F_wv is the vertical wave force. This relation is a design-balance condition, not a dynamic equation of motion. It ensures that the module maintains freeboard or submergence within its intended operating band. Biofouling enters through M and can also change V_0 if fouling displaces water or blocks voids; inspection and cleaning schedules therefore affect both habitat value and buoyancy reserve.

Mooring safety is expressed as a load-factor constraint:

$$\gamma_m \, max_t \, T(t) \leq T_allow$$

where γ_m is a mooring safety factor, T(t) is the time-dependent tension, and T_allow is the allowable tension derived from component strength, connector capacity, and anchor performance. The maximum is taken over the design load cases, including combined wave

and current conditions and degraded-buoyancy scenarios. A floating reef that is biologically attractive but has a short mooring fatigue life is not a viable habitat. Fatigue at shackles, splices, and printed connections deserves particular attention because oscillatory loading in shallow coastal water can produce many cycles at moderate stress.

Structural reliability is evaluated through a limit-state probability:

$$P_f(x) = P(max_t\,[\sigma(t)/\sigma_y] \geq 1) \leq P_f*$$

where P_f(x) is the probability of failure for design x, σ(t) is the governing stress, σ_y is the yield or allowable stress, and P_f is the target failure probability. This expression can be adapted to mooring ultimate limit states, fatigue damage accumulation, or loss-of-buoyancy events by replacing the stress ratio with the relevant limit-state function. The purpose is to make risk explicit rather than relying on a single deterministic safety factor. The target P_f must reflect consequence severity, inspection frequency, and redundancy; a module adjacent to active aquaculture infrastructure warrants stricter reliability than an isolated experimental unit.

5. Biodiversity and Habitat Performance

Biodiversity performance is not equivalent to total surface area. A habitat complexity index can be defined as

$$H_c = w_1\left(\frac{A_i}{A_0}\right) + w_2 R_s + w_3\left(\frac{V_i}{V_0}\right) + w_4 C_f$$

where A_i is interstitial or sheltered surface area, A_0 is a reference projected area, R_s is substrate rugosity, V_i is void or crevice volume, V_0 is displaced volume, C_f is a crevice-frequency or refuge-size descriptor, and w_1 to w_4 are nonnegative weights that sum to one. The weights encode management priorities: juvenile refuge may weight crevice volume and frequency more heavily, whereas invertebrate colonization may weight rugosity and interstitial area more heavily. This index is deliberately transparent so that ecological assumptions can be reviewed and changed. It also exposes a structural penalty: increasing A_i, R_s, or V_i tends to increase projected area, material mass, or connection complexity, which feeds back into the load and buoyancy equations.

Expected biodiversity performance can be written as

$$B(x) = \sum_k \pi_k \mathrm{E}\left[f_k\left(H_c, R_s, s, N_p, C\right)\right]$$

where k indexes taxa or functional groups, π_k is the management weight for group k, E denotes expectation over environmental and recruitment uncertainty, f_k is a response function linking habitat attributes to suitability for group k, s is module spacing, N_p is module density, and C is a connectivity descriptor. The response functions should be monotonic only where evidence supports monotonicity. For example, refuge size may improve juvenile survival up to a point and then favor larger predators. Spacing may enhance connectivity among modules but also create hydraulic wakes and mooring congestion. The expectation term is essential because colonization is stochastic and site-dependent; a design that performs well in one recruitment season may not do so in another.

Artificial reefs can support distinct assemblages rather than simply reproducing nearby natural reefs [15], so biodiversity performance should include community composition and functional diversity, not only fish abundance. A module that attracts large predators may

score highly on a naive abundance metric while offering poor juvenile habitat. Conversely, a low-profile module with high rugosity may support diverse invertebrates without producing strong fish aggregation. The framework therefore treats B(x) as a vector of group-specific scores before aggregation, allowing decision-makers to inspect which groups benefit and which are disadvantaged. This is especially relevant near aquaculture, where aggregating wild fish may create operational, biosecurity, or entanglement concerns. Biodiversity value and aquaculture compatibility are related but not identical objectives.

6. Integrated Optimization and Trade-off Analysis

The integrated design problem can be stated as

$$min_x\,\{-B(x), U_s(x), C_m(x)\}\, s.t.\, g_j(x) \le 0$$

where U_s(x) is structural utilization or a surrogate for failure probability, C_m(x) is a cost or maintenance burden, and g_j(x) ≤ 0 represents constraints such as mooring safety, buoyancy reserve, navigation clearance, and material limits. The negative sign converts biodiversity maximization into a minimization objective. The solution is not a single vector x but a Pareto set of designs that trade biodiversity against structural utilization and cost. A design with high H_c may be non-dominated because it provides refuge value unavailable in simpler modules, while a lower-complexity design may be preferred if it reduces mooring load and inspection frequency. The Pareto set makes those choices visible instead of hiding them in a weighted sum.

Decision rules should be applied after the Pareto set is generated. One rule selects the design with the largest biodiversity gain per unit increase in structural utilization relative to a baseline module. Another rule imposes a reliability threshold and then maximizes B(x). A third rule minimizes regret across plausible stakeholder weights. These rules are not equivalent, and the choice should be documented because it affects which habitat geometry is deployed. The surrogate should be updated after deployment using inspection, load monitoring, and ecological survey data. If measured biofouling mass or colonization differs from prediction, the posterior surrogate changes and the Pareto set may shift. This adaptive loop is the practical meaning of a data-driven approach: design decisions are revised as evidence accumulates, not fixed once and forgotten.

Spatial arrangement adds a further layer. Modules can be clustered to create connected habitat patches, distributed to reduce wake interference, or aligned to dominant current directions to improve feeding opportunities for filter feeders. However, spacing and orientation also change mooring interference and load shadowing. A design that is optimal for a single module may be suboptimal for an array. The framework can accommodate array-level variables by expanding x with spacing and orientation parameters and by evaluating array-level drag, mooring redundancy, and habitat connectivity. The key requirement is that array-level ecological benefits be compared against array-level structural and operational costs using the same reliability-constrained logic.

7. Discussion

The framework's main technical claim is that structural stability and biodiversity performance in floating reef habitats are coupled through geometry. Complexity increases surface area, crevice volume, and often projected area; projected area increases drag; drag increases mooring tension and fatigue; and heavier or more heavily fouled modules lose buoyancy reserve. At the same time, complexity can provide refuge and colonization substrate

that simple modules lack. This coupling means that ecological enhancement should not be treated as an additive feature applied after structural design. It should enter the design vector from the beginning and be evaluated against reliability constraints.

The data-driven component is equally important because the evidence base for floating reef habitats is necessarily heterogeneous. Hydrodynamic models may be well constrained by wave and current measurements, whereas recruitment and colonization responses are often sparse and site-specific. Bayesian surrogates are useful in this setting because they represent uncertainty explicitly and can incorporate multiple information sources without pretending that all observations are equally informative. Global sensitivity analysis helps avoid overfitting ecological detail in regions of the design space where structural constraints dominate. The approach is therefore appropriate for a design problem in which some relationships are physically derived and others are empirical or expert-informed.

Several boundary conditions limit interpretation. First, the framework is design-level and does not replace detailed finite-element analysis, mooring certification, or site-specific environmental impact assessment. Second, the biodiversity response functions must be calibrated to local species pools and aquaculture operations; a module designed for one coastal region may not transfer directly to another. Third, biofouling and colonization are time-varying, so a design that is stable at deployment may become unacceptable after substantial growth. Inspection and cleaning schedules should therefore be treated as design variables or constraints, not as afterthoughts. Fourth, aggregation of wild fish near aquaculture facilities raises ecological and operational questions that cannot be resolved by structural optimization alone. The framework can make those trade-offs explicit, but it cannot substitute for local ecological and regulatory judgment.

8. Conclusion

This manuscript has developed a data-driven design framework for floating reef habitats in coastal aquaculture that balances structural stability and marine biodiversity within a single reliability-constrained, multi-objective formulation. The framework parameterizes floating modules by geometric, structural, and ecological variables; represents wave loading, buoyancy, mooring tension, and reliability through explicit limit-state relationships; and represents biodiversity through a transparent habitat-complexity index and taxa-weighted expected performance functions. Bayesian surrogate modeling, sensitivity analysis, and Pareto-based optimization allow sparse and heterogeneous evidence to be used without collapsing the design problem into a single arbitrary score. The central conclusion is that habitat complexity and structural load are inseparable design variables. Designs that increase refuge and substrate value also alter drag, mass, buoyancy, and mooring demand, so ecological enhancement must be optimized jointly with reliability constraints. By producing non-dominated design options and updating them as monitoring data accumulate, the framework supports floating reef habitats that are both safer to moor and more ecologically purposeful in coastal aquaculture systems.

References

1. Baine, M. (2001). Artificial reefs: a review of their design, application, management and performance. Ocean & Coastal Management, 44(3-4), 241-259.

2. Seaman, W. (Ed.). (2000). Artificial reef evaluation: With application to natural marine habitats. CRC Press.

3. Perkol-Finkel, S., Shashar, N., & Benayahu, Y. (2006). Can artificial reefs mimic natural reef communities? The roles of structural features and age. Marine Environmental Research, 61(2), 121-135.

4. Graham, N. A. J., & Nash, K. L. (2013). The importance of structural complexity in coral reef ecosystems. Coral Reefs, 32(2), 315-326.

5. Komyakova, V., Munday, P. L., & Jones, G. P. (2013). Relative importance of coral cover, habitat complexity and structural characteristics in determining reef fish assemblages. PLoS ONE, 8(2), e55860.

6. Dafforn, K. A., Glasby, T. M., Airoldi, L., Rivero, N. K., Mayer-Pinto, M., & Johnston, E. L. (2015). Marine urbanization: an ecological framework for designing multifunctional artificial structures. Frontiers in Ecology and the Environment, 13(2), 82-90.

7. Faltinsen, O. M. (1990). Sea loads on ships and offshore structures. Cambridge University Press.

8. Morison, J. R., O'Brien, M. P., Johnson, J. W., & Schaaf, S. A. (1950). The force exerted by surface waves on piles. Petroleum Transactions, AIME, 189, 149-154.

9. Det Norske Veritas. (2010). Environmental conditions and environmental loads (DNV-RP-C205). Det Norske Veritas.

10. Luo, D., Wang, B., Bao, D. W., & Ward, S. (2022). Opportunities for further development of 3D-printed floating artificial reefs. Journal of Aquaculture and Marine Biology, 11(2), 58-63.

11. Forrester, A. I. J., Sóbester, A., & Keane, A. J. (2008). Engineering design via surrogate modelling: A practical guide. Wiley.

12. Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3), 425-464.

13. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). Global sensitivity analysis: The primer. Wiley.

14. Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182-197.

15. Folpp, H., Lowry, M., Gregson, M., & Suthers, I. M. (2013). Fish assemblages on estuarine artificial reefs: natural rocky-reef mimics or discrete assemblages? PLoS ONE, 8(6), e63505.