# Statistical Tools and machine learning Algorithms to uncover the ecological organization of Mediterranean Cold-Water Coral Reefs

FINAL REPORT OF THE 2025 MASTS VISITING FELLOWSHIP

ARIADNA MARTÍNEZ-DIOS

Supervised by:

Dr. Johanne Vad
The University of Edinburgh

Dr. Laurence De Clippele
Glasgow University

# 1. Background and Objectives

The Mediterranean Sea is home to several cold-water coral populations, with seven coral regions or provinces identified to date (Chimienti et al., 2019 and references therein). While most of Mediterranean Cold-Water Coral (hereafter, CWC) populations do not form reefs, the Alboran Sea hosts extensive and exceptionally thriving reefs growing on top of volcanic banks, such as the Cabliers and Catifas reefs (Corbera et al., 2019; Pardo et al., 2011). There, CWC communities are exceptionally dense, with reported densities, two orders of magnitude higher than other Mediterranean CWC assemblages (Corbera et al., 2019). The ability to construct massive 3D frameworks, allows CWCs to create oases of life in the deep sea (Roberts *et. al.,* 2006), as they do in the Alboran Sea. These communities self-organize to enhance food capture for many other species (Van Der Kaaden et al., 2023; Guill et al., 2021) like sponges, crinoids, sea spiders, crabs, sea urchins, which successfully exploit the cycling of nutrients and organic matter in coral communities and transfer energy to higher trophic levels (Maier et al., 2023; Gili and Coma 1998). Beyond size and extension, Cabliers and Catifas reefs are hotspots of biodiversity, providing habitat for up to 45 protected species, many of them indicator of Vulnerable Marine Ecosystems (Habitats Directive, FAO). In turn, the habitats they create support important fisheries in the region, with many commercially valuable fish observed to feed and nurse on the reefs (Corbera et al., 2019).

In this context, the concept of *size* becomes relevant, as it is a fundamental organismal trait and an important driver of ecosystem functioning that in turn influences community structure and ecological organization in a density-dependent way (Chase *et al.,* 2020). This is especially true when ecosystem functions involve the creation of biogenic habitats as CWC do in Cabliers and Catifas banks, here is where *size matters*. However, current knowledge on distribution and composition of the reefs is exclusively based on three ROV dives in the area (Corbera et al., 2019), and the factors driving their presence, diversity and exceptional state of conservation remain poorly understood. Moreover, the development of environmental regulations for conservation is hampered by profound gaps on the knowledge of the ecology and distribution patterns on multiple-timescales and the driving forces that favor the formation of these complex systems.

Therefore, understanding the demographic characteristics of species is essential for creating effective conservation plans and predicting their development in disturbed conditions. This is especially relevant now that the General Assembly of the General Fisheries Commission for the Mediterranean Sea (FAO, UN) declared in 2023 a Fisheries Restricted Area (FRA) of 400 $km^2$ to protect Cabliers reefs. Therefore, data on size and structure of the populations of Cabliers and Catifas banks will provide empirical support to drive where direct protection efforts (areas of highest density of mature/functional

organisms vs. areas of highest diversity, etc.), as well as to define adequate delimitation, making sure not leaving important areas unprotected.
In this sense, the project embraces three key concepts in Ecology: size, pattern and scale, in the study of complex ecological systems while providing new data about perhaps the last living CWC reefs in the Mediterranean Sea. We addressed these questions through two main objectives by using Machine Learning tools and advanced statistical approaches:

### Objective 1 – Extraction of habitat-structuring species' morphometrics using Machine Learning models.

Cutting-edge Machine Learning tools, such as RootPainter (Smith et al., 2022, Clark et al., 2024) were used to extract metrics related to the size and shape of habitat-structuring species, such as corals and sponges. The models were built in collaboration with Dr. Laurence De Clippele from Glasgow University.

### Objective 2 – Investigate the effects of density and size structure in the ecological structure of the community.

The learning, coding and execution of these methods will these analyses were carried out at the University of Edinburgh under the supervision of Dr. Johanne Vad.

## 2. Activities and main results

### Objective 1 – Extraction of habitat-structuring species' morphometrics using Machine Learning models.

In the context of advancing ecological monitoring, machine learning tools are increasingly being employed for the automatic detection of marine species from imagery. As part of this effort, I have been developing and training deep learning models using *RootPainter*, a Deep Learning tool that uses artificial neural networks to detect and classify objects from large and complex image datasets through semantic segmentation (Smith et al., 2022).

Specifically, I have trained three models to detect the presence of the glass sponge *Asconema setubalense*. After annotating and training on a set of 20 images, the best-performing model achieved an accuracy of 0.98, demonstrating strong potential for automatic detection. Nonetheless, model precision varied significantly (ranging from 0.4 to 1), particularly in regions corresponding to the edges of the organisms. This variability reflects the sensitivity of RootPainter's pixel-level classification to subtle shifts in color and texture, highlighting the need for additional training data and refinement to improve boundary recognition and overall model robustness.

A) B) C)

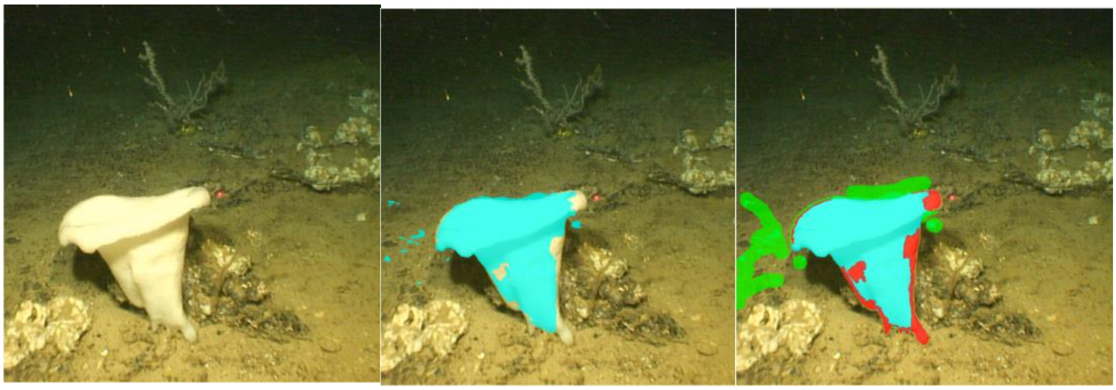

Figure 1. Stages of RootPainter training for the species of interest *A. setubalense*. From left to right A) Input image; B) Model's segmentation (i.e., prediction) after training 16 images; C) Corrective training. Blue color indicates model's prediction, green correction of background pixels, and red correction for foreground pixels. That is, areas where the model failed to predict the organism.

As a next step, I plan to extend this approach to other ecologically important ecosystem-engineer species, such as *D. pertusum* and *Acanthogorgia spp.*, in order to build a broader automated detection framework for benthic habitat assessments.

### Objective 2 – Investigate the effects of density and size structure in the ecological structure of the community

A total of 20,185 frames were analysed, resulting in 53.202 organisms annotated, and 153 species identified of major taxonomic marine Phyla including Cnidaria, Porifera, Echinodermata, Mollusca, Annelida, Brachiopoda, and Chordata.

Cnidaria emerged as the most dominant phylum of benthic megafauna, consistently representing the highest proportion of observed organisms and accounted for over 75% of the total fauna observed; although the relative contribution of species varied across dives (Fig.2). Among cnidarians, *Acanthogorgia spp.* overwhelmingly dominated the community, followed by scleractinian corals as *Desmophyllum dianthus*, *Desmophyllum dianthus* and *Desmophyllum pertusum*, and the black coral like *Parantipathes larix* (Fig.2). Within the phylum Porifera, *A. setubalense* was the most frequently recorded species, together with small representation of *Sympagella delazuei* (Fig.2).

Although the aforementioned results may suggest spatial heterogeneity in community assembly, the k-means clustering analysis based on species composition data identified four distinct faunal assemblages composed mainly by the same structuring species. This therefore suggests the presence of a single community structured by ecosystem engineer species *Acanthogorgia spp., D. dianthus, D. pertusum, A. setubalense and P. larix* (Fig.3).

The analyzed boxplots reveal clear differences in density distribution and spatial occupancy among the five taxa. *Acanthogorgia spp.* displays relatively high densities across dives compared to other taxa, with mean density values of 3 individuals $m^{-2}$ and several outliers reaching 20 individuals $m^{-2}$. Overall, indicating broad spatial occupancy but a strongly patchy distribution i.e., it's widespread, but not evenly distributed. The hexactinellid sponge *A. setubalense*, is present at all dives with variable densities, never greater than 1 individual $m^{-2}$. Still, it shows slightly higher densities in dive 11, suggesting certain habitat selectivity. Scleractinian corals *D. pertusum* but especially *M. oculata*, occur at low densities, meaning that individuals are separated by several meters. Occasionally, both species present extreme density values of 4 individuals $m^{-2}$, indicative of more localized populations and certain clustering. Finally, the antipatharian *P. larix* is a common species in Catifas Bank, occurring at densities lower than 1 individual per $m^2$, but larger variability. Certain locations, such as in dive 3 and 11, seem to favor the presence of *P. larix*, where 1-2 individuals per $m^2$ can be found. Overall, the benthic community of Catifas Bank reflects taxon-specific occupancy patterns**,** likely reflecting differences in ecological niches/environmental preferences, and interactions with co-occurring species.

Beyond differences in the relative contribution of structuring species, the truth is that species composition alone proved insufficient tool to characterize the benthic community of Catifas Bank. In fact, changes in the size of the individuals and subtle, intra-community transitions were observed thanks to the continuous analysis of video footage (Fig.4). For instance, sites that had similar species composition (i.e., *D. pertusum, Acanthogorgia spp.* and *A. setubalense*), clearly differed in the size structure. In short spatial intervals, big lobular *D. pertusum* colonies gave way to smaller, but spatially denser colonies (Fig. 4). Similarly, while dense gardens of medium to big *Acanthogorgia spp.* at certain locations, other sites distinctly were occupied by small colonies (<10cm in height) (Fig. 4).

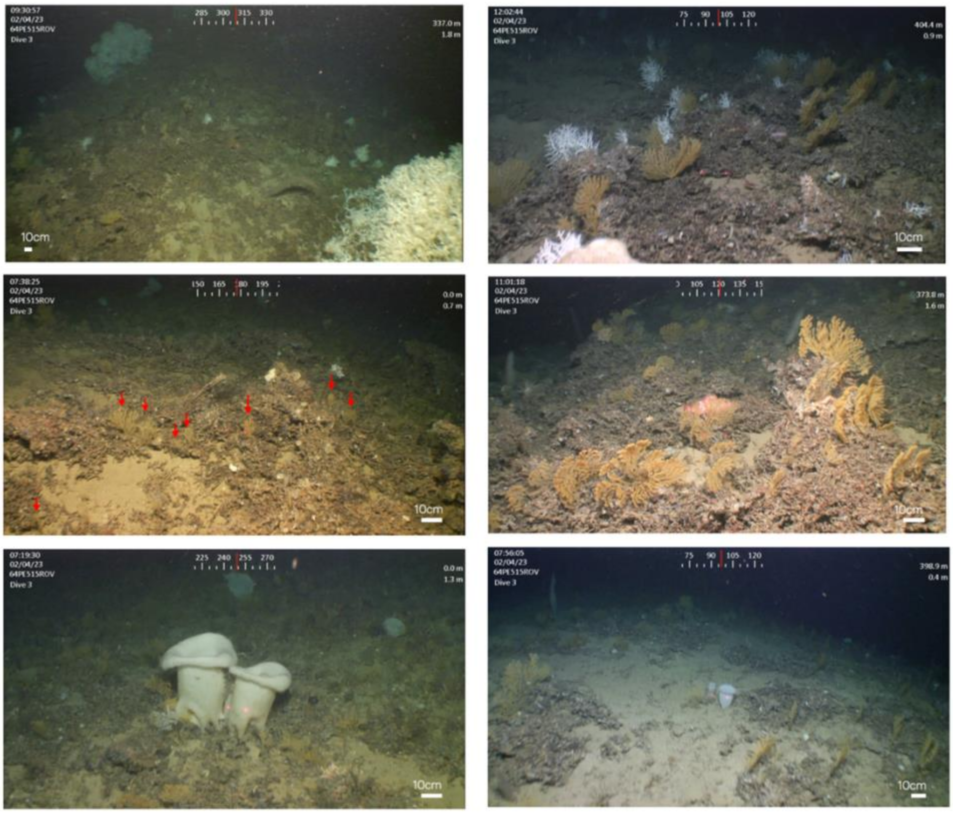


**Figure 4.** Examples of distinct size structures and intra-community transitions observed. Red arrows highlight the presence of a cluster containing small Acanthogorgia spp. individuals.

In this sense, one of the main tasks of the fellowship was to investigate which size classes are contributing more to the structure of the community at each site. To this, we computed Cumulative Abundance Profiles (CAPs) for all sessile species, encompassing both density and individual size data (height) following the approach developed for forest ecology by De Cáceres *et al.* (2013). By analyzing together species composition and size structure we were able to better characterize benthic diversity and organization. Cumulative Abundance Profiles proved to be a very useful tool, and effectively revealed marked interspecific variation in the distribution of organism sizes, allowing for the identification of five distinct size-structured assemblages or clusters (Fig. 5).

The dominant species — *Acanthogorgia spp.*, *Desmophyllum pertusum*, *A. setubalense*, and *Parantipathes larix* — exhibited contrasting CAP shapes, reflecting divergent growth stages and population structures. For instance, *Acanthogorgia spp.* and *D. pertusum* displayed a truncated profile consistent with size-skewed populations dominated by larger density of small to medium colonies, while *A. setubalense* exhibited longer, more gradual profiles, pointing to broader size distributions and a higher prevalence of both juveniles and large, mature individuals. On the contrary, of *P. larix* populations were distinctly dominated by a narrow range of large individuals. Overall, the resulting size clusters provided an ecologically meaningful classification of species into groups characterized by shared structural attributes, highlighting the potential of CAPs to uncover subtle differences in life-history traits and population dynamics within benthic communities.

CAPs were used with dissimilarity matrices and ordination analysis, in an attempt to understand how the scale-dependent distribution of species and sizes, and their interactions within and among species, foster or impede the co-existence of species; how they alter spatial patterns in species richness and composition. We dedicated special effort in developing distance-based Moran's eigenvector maps (dbMEMs, Borcard et al., 2004; Borcard et al., 2018) – a type of multiscale ordination analysis –to decompose the spatial relationships among study sites. This analysis uses the geographical coordinates of our sampling sites, to construct a matrix of Euclidean distances among the sites. In this procedure, the first eigenvectors describe broad spatial structures, encompassing the spatial variation in the whole sampled area (13Km), while the last eigenvectors (with lower eigenvalues) describe fine spatial structures, which may capture variation at the scale of sampling units ($5m^2$ and $1m^2$ respectively). Finally, we used Redundancy Analysis (RDA) with full suite of terrain variables (depth, slope, rugosity, eastness, northness and bpi) and significant MEMs, followed by partial redundancy analyses (pRDAs) controlling the influence of the spatial structure to investigate 1) to what extend the community assembly is spatially and environmentally mediated and 2) what terrain variables influence the distribution of species and sizes. This method is based in what is called *variation partitioning*, that allows the assessment of variation explained purely by environmental variation, spatially structured environmental variation or purely spatial variation (Legendre and Legendre 2012).

Regarding the spatial organization of such assemblages, the DbMEMs analysis identified 57 eigenvalues, the first 12 of which were statistically significant and explained 8% of the variation in community assembly, indicating that broad-scale patterns are explaining the distribution of species and sizes. When the effects of the spatial structure were excluded by partialling out their effects in the redundancy analysis (pRDA), community assembly was significantly related to environmental heterogeneity (4%). All terrain variables seem to have the same influence; however, certain patterns were observed. For instance, assemblages containing largest individuals across species (clusters A and D) were found at complex and sloping settings facing North, while cluster B containing the smallest individuals, was located in areas facing east. Overall, indicating that BPI, slope and terrain orientation are key geomorphologic attributes that create heterogeneous environments to support distinct faunal assemblages.

Nevertheless, full model only accounted for 12% of the variance explained. We hypothesize that other ecological processes that operate on a broad scale and that we have not incorporated into the model, might be explaining the distribution of the assemblages beyond environmental forcing. Two main mechanisms might be responsible for the spatial structure observed: biotic processes (recruitment, mortality, inter/intra-specific competition) or stochastic processes (abrupt fluctuations in environmental conditions).

## 3. Achievements and follow-up activities

To date, the MASTS Visiting Fellowship results have been presented at:

1) the Global Change seminar of the Faculty of Geosciences of the University of Edinburgh (Scotland, United Kingdom) on the 13th of May 2025, with a talk entitled: *"When size matters: Deciphering structure of deep-sea communities beyond species composition".*

2) III SIBECOL & XVII AEET MEETING, Pontevedra (Spain), as part of the poster session "Coping with the Change: Exploring the impacts of global change on marine biodiversity across spatial-temporal scales".

3) Submission of a Project Focus article to Deep-Sea Life Newsletter of the Deep-Sea Biology Society (September 2025).

Currently two publilcations are being prepared and during the course of this year, I expect to further disseminate at the MASTS Annual Science Meeting that will take place on 18-20 November 2025 at the at the Technology & Innovation Centre (TIC) of University of Strathclyde (Glasgow, UK) and the Ocean Sciences Meeting 2026 (Glasgow, UK).

## 4. References

Borcard, D., Legendre, P., Avois-Jacquet, C., Tuomisto, H., 2004. Dissecting the spatial structure of ecological data at multiple spatial scales. Ecology 85, 1826–1832. https://doi.org/10.1890/03-3111

Borcard, D., Gillet, F., Legendre, P., 2018. Numerical Ecology with R, Use R! Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-71404-2

Chase, J.M., Jeliazkov, A., Ladouceur, E., Viana, D.S., 2020. Biodiversity conservation through the lens of metacommunity ecology. Annals of the New York Academy of Sciences 1469, 86–104. https://doi.org/10.1111/nyas.14378

Chimienti, G., Bo, M., Taviani, M., Mastrototaro, F., 2019. 19 Occurrence and Biogeography of Mediterranean Cold-Water Corals, in: Orejas, C., Jiménez, C. (Eds.), Mediterranean Cold-Water Corals: Past, Present and Future. Springer International Publishing, Cham, pp. 213–243. https://doi.org/10.1007/978-3-319-91608-8_19

Clark, H.P., Smith, A.G., McKay Fletcher, D., Larsson, A.I., Jaspars, M., De Clippele, L.H., 2024. New interactive machine learning tool for marine image analysis. R. Soc. Open Sci. 11. https://doi.org/10.1098/rsos.231678

Corbera, G., Lo Iacono, C., Gràcia, E., Grinyó, J., Pierdomenico, M., Huvenne, V.A.I., Aguilar, R., Gili, J.M., 2019. Ecological characterisation of a Mediterranean cold-water coral reef: Cabliers Coral Mound Province (Alboran Sea, western Mediterranean). Progress in Oceanography 175, 245–262. https://doi.org/10.1016/j.pocean.2019.04.010

De Cáceres, M., Legendre, P., He, F., 2013. Dissimilarity measurements and the size structure of ecological communities. Methods in Ecology and Evolution 4, 1167–1177. https://doi.org/10.1111/2041-210X.12116

European Commission, Directorate-General for Environment, *The strict protection of animal species of Community interest under the Habitats Directive – Guidance document: a summary*, Publications Office of the European Union, 2021, https://data.europa.eu/doi/10.2779/3123

FAO (2009) International guidelines for the management of deep-sea fisheries in the High Seas. Rome.

Gili, J.-M., Coma, R., 1998. Benthic suspension feeders: their paramount role in littoral marine food webs. Trends in Ecology & Evolution 13, 316–321. https://doi.org/10.1016/S0169-5347(98)01365-2

Gonzalez, A., Germain, R.M., Srivastava, D.S., Filotas, E., Dee, L.E., Gravel, D., Thompson, P.L., Isbell, F., Wang, S., Kéfi, S., Montoya, J., Zelnik, Y.R., Loreau, M., 2020. Scaling-up biodiversity-ecosystem functioning research. Ecology Letters 23, 757–776. https://doi.org/10.1111/ele.13456

Guill, C., Hülsemann, J., Klauschies, T., 2021. Self-organised pattern formation increases local diversity in metacommunities. Ecology Letters 24, 2624–2634. https://doi.org/10.1111/ele.13880

Legendre, P., Legendre, L., 2012. Chapter 14 - Multiscale analysis: spatial eigenfunctions, in: Legendre, P., Legendre, L. (Eds.), Numerical Ecology, Developments in Environmental Modelling. Elsevier, pp. 859–906. https://doi.org/10.1016/B978-0-444-53868-0.50014-9

Maier, S.R., Brooke, S., De Clippele, L.H., de Froe, E., van der Kaaden, A.-S., Kutti, T., Mienis, F., van Oevelen, D., 2023. On the paradox of thriving cold-water coral reefs in the food-limited deep sea. Biological Reviews n/a. https://doi.org/10.1111/brv.12976

Smith, A.G., Han, E., Petersen, J., Olsen, N.A.F., Giese, C., Athmann, M., Dresbøll, D.B., Thorup-Kristensen, K., 2022. Root Painter: deep learning segmentation of biological images with corrective annotation. New Phytologist 236, 774–791. https://doi.org/10.1111/nph.18387

Van Der Kaaden, A., Maier, S.R., Siteur, K., De Clippele, L.H., Van De Koppel, J., Purkis, S.J., Rietkerk, M., Soetaert, K., Van Oevelen, D., 2023. Tiger reefs: Self-organized regular patterns in deep-sea cold-water coral reefs. Ecosphere 14, e4654. https://doi.org/10.1002/ecs2.4654