Machine learning applications to monitor marine Essential Biodiversity Variables of rocky shore communities

G. Bravo 1, **N. Moity** 2, **E. Londoño-Cruz** 3, **F. E. Muller-Karger** 4, **G. Bigatti** 1, 5 , **E. Montes** 4

1 Instituto de Biología de Organismos Marinos, Consejo Nacional de Investigaciones Científicas y Técnicas, Puerto Madryn, Argentina

2 Charles Darwin Research Station, Charles Darwin Foundation, Santa Cruz, Galapagos , Ecuador

3 Departamento de Biología, Universidad del Valle, Cali, Colombia

4 College of Marine Science, University of South Florida, St Petersburg, Florida, USA

5 Universidad Espíritu Santo, Guayaquil, Ecuador

Abstract

Monitoring marine ecosystems and biodiversity are necessary to understand ecological patterns and processes but also to detect natural or human induced changes such as those resulting from climate change or coastal pollution. A standard technique used in rocky shores is the estimation of cover of sessile organisms and macro-invertebrates. Photoquadrats are becoming standard practice for surveying biodiversity of intertidal and subtidal habitats. They allow to collect large volumes of reliable data efficiently and rapidly in addition to provide a permanent record of the sample. Despite known limitation in taxonomic resolution compared to visual quadrats, photoquadrats have demonstrated to perform well when estimating percent cover of functional groups. Nonetheless, photoquadrat analyses are time consuming and may lead to insufficient accuracy due to low sampling. Cutting-edge machine learning tools are now being used by marine ecologists to annotate species records from photoquadrat imagery. They allow the automatic identification of species, or functional groups, to examine community composition and biodiversity of rocky intertidal and subtidal habitats with high certainty. The use of these tools can significantly reduce the processing time of photo-quadrat imagery and optimize biodiversity survey programs. In this study we present results from visual versus photoquadrat assessments of rocky shores from Argentina, Galapagos Islands (Ecuador) and the Pacific Colombian coast using the CoralNet software. Photo-quadrat imagery was collected during visual surveys carried out at these sites following the South American Research Group on Coastal Ecosystems (SARCE) protocol implemented across the continent by the Marine Biodiversity Observation Network of the Americas (MBON Pole to Pole) program. We apply an ad hoc standardized list of benthic biota and substrata (i.e. CATAMI) as a common label set to enable the comparison between locations. Preliminary results show that CoralNet is able to identify key benthic species and substrate types with high levels of confidence (Pearson correlation coefficient (r) from computer vs visual annotations: Substrate Consolidated, r = 0.91, Molluscs Bivalves, r = 1, Macroalgae filamentous, r = 0.79, Macroalgae sheet-like r = 0.87). We conclude that the CoralNet software can be used to extract presence and percent cover of CATAMI categories. Change detection was tested with an unsupervised configuration of the CoralNet software successfully detecting changes in percent cover of bivalves at 3 sites in Puerto Madryn, Argentina, between Nov/2018 and Nov/2019. This method brings together two programs that are already working to facilitate data analyses over large latitudinal gradients. We suggest the suitability of this method to establish a protocol to rapidly describe rocky shore biodiversity and to detect changes in the biota in the time frame of the MBON Pole to Pole project along the American continent.

Objetives

- Test unsupervised configuration of the CoralNet software to identify CATAMI categories on intertidal photoquadrats from sampling sites of MBON.
- Provide a fast method to detect changes on intertidal habitats, using photoquadrats analysided by artificial intelligence (AI).

Methods

Here we present the firts results using photoquadrats from Argentina

The robot was trained with 151 photos (50 x 50 cm quadrat) with 100 annotations per photo performed by an experienced observator (Human annotator).

An independent set of photoquadrats (n=90) from the same sites and dates was analyised with CoralNet AI. Both results (Human annotator and Robot annotator) were compared with visual quadrats results. The quadrsts analyised visually were the same as the photoquadrats annotated by the CoralNet Robot.

Sampling Protocol for each site:

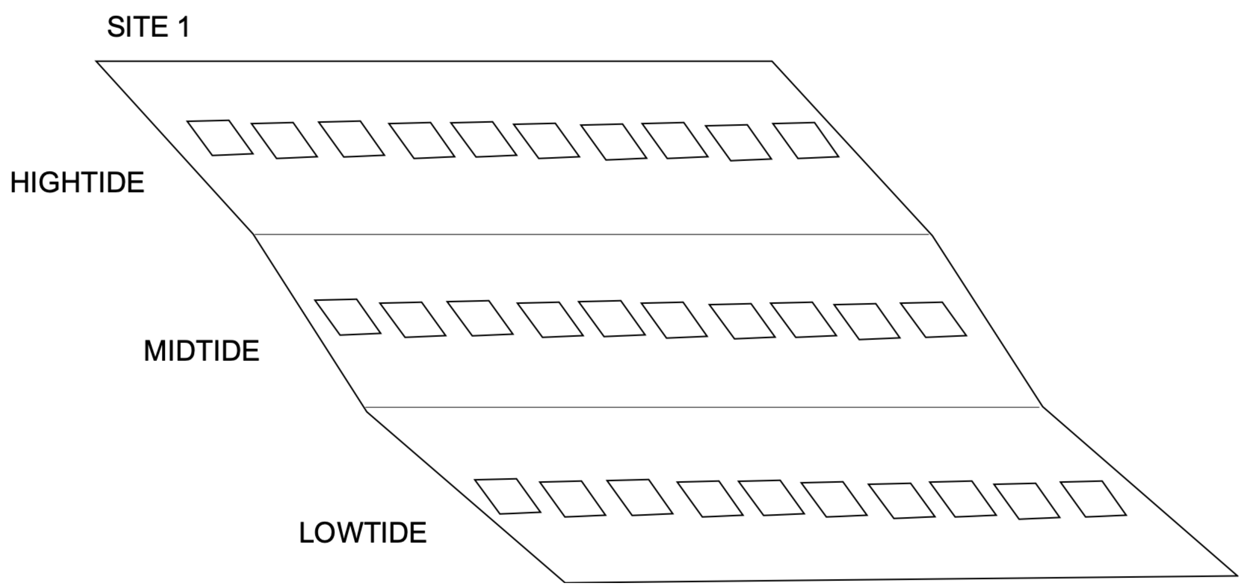

Figure 1: 10 phtoquadrats per strata

In Argentina ~50 extra photos were taken on each site, allowing to train the CoralNet robot with different photos from those analysed viasually. If there arent extra photos, the robot can be trained with the same photos but using alleatory points insted (working progress).

We decide to use CATAMI codes insted of species to unify among different countries.

The same number of CATAMI categories were used for visual and photoquadrats:

	CATAMI CODE	Frequency
3	MOB	9474
7	SC	8665
11	MAFG	4569
16	MASG	3945
8	MAA	3073
12	MAFR	1729
9	MAEN	615
2	CRB	523
5	WPOT	263
14	MALCB	165
13	MAG	46
15	MASB	18
1	CNTR	9
10	MAENRC	5
4	SP	1
6	SU	0

Name	Short.Code	Label.ID	Functional.Group
Ascidians	A	4762	Other Invertebrates
Bryozoan	BRY	1853	Other Invertebrates
Cnidaria: Colonial anemones	CNCA	4773	Other Invertebrates
Cnidaria: True anemones	CNTR	4774	Other Invertebrates
Crustacea: Barnacles	CRB	357	Other Invertebrates
Echinoderms: Sea stars	ESS	344	Other Invertebrates
Echinoderms: sea urchin	ESU	342	Other Invertebrates
Molluscs: Bivalves	MOB	355	Other Invertebrates
Molluscs: Chitons	MOCH	354	Other Invertebrates
Molluscs: Gastropods	MOG	353	Other Invertebrates
Sponge	SP	102	Other Invertebrates
Worms: Polychaetes: Tube worms	WPOT	361	Other Invertebrates
Substrate: Unconsolidated (soft)	SU	4115	Soft Substrate
Substrate: Consolidated (hard)	SC	4114	Hard Substrate
Verrucaria	LICHEN	4719	Other
Shadow	SHAD	222	Other
Unclear	Unc	118	Other
Macroalgae: Articulated calcareous	MAA	325	Algae
Macroalgae: Erect coarse branching	MAEC	317	Algae
Macroalgae: Erect fine branching	MAEF	313	Algae
Macroalgae: Encrusting	MAEN	321	Algae
Macroalgae: Encrusting: red	MAENR	322	Algae
Macroalgae: Filamentous / filiform	MAF	309	Algae
Macroalgae: Globose / saccate	MAG	305	Algae
Macroalgae: Laminate	MALA	301	Algae
Macroalgae: Large canopy-forming : brown	MALCB	299	Algae
Macroalgae: Sheet-like / membraneous	MAS	294	Algae

Results

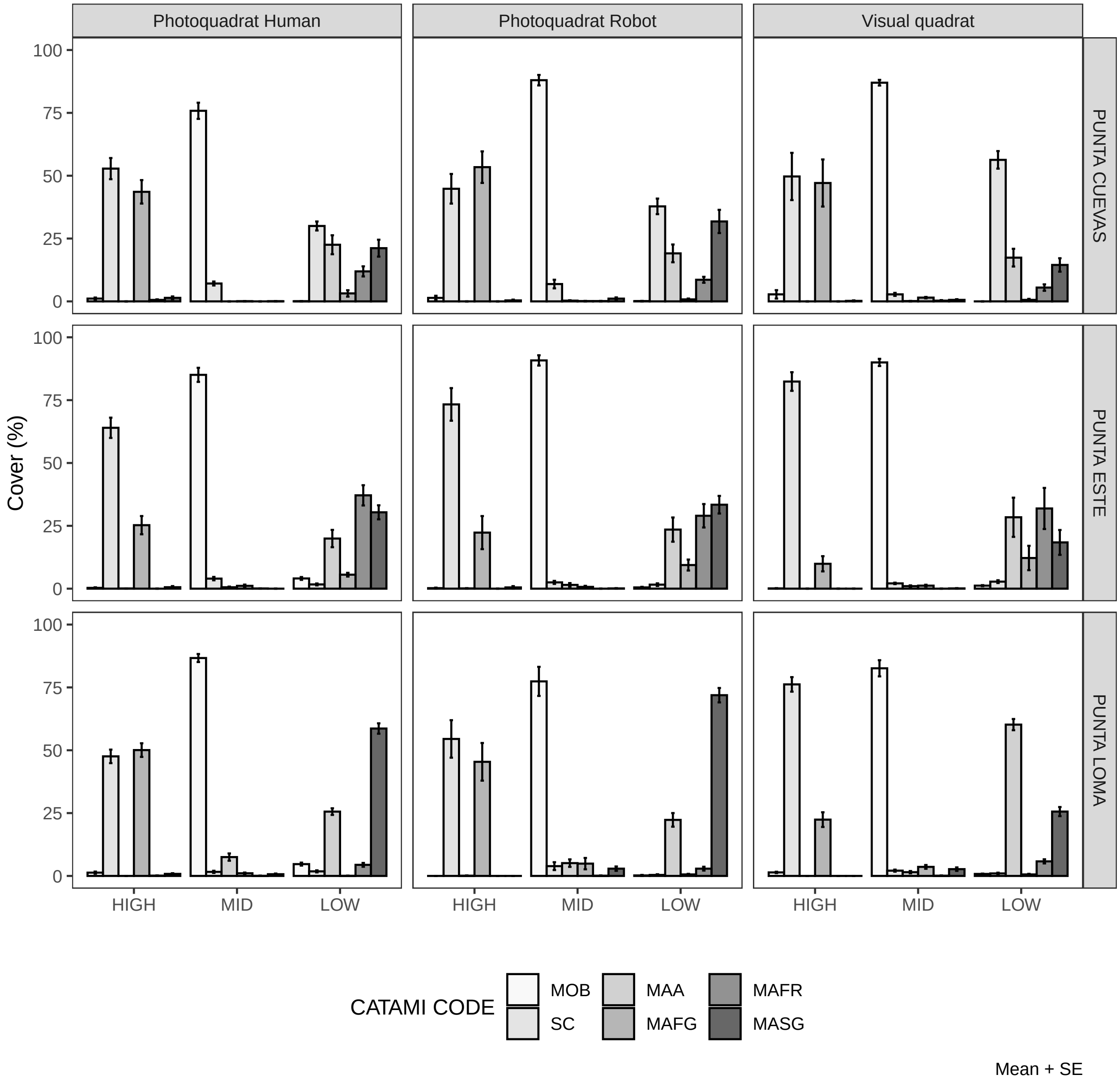

Figure 2: Overall view of differences between methods

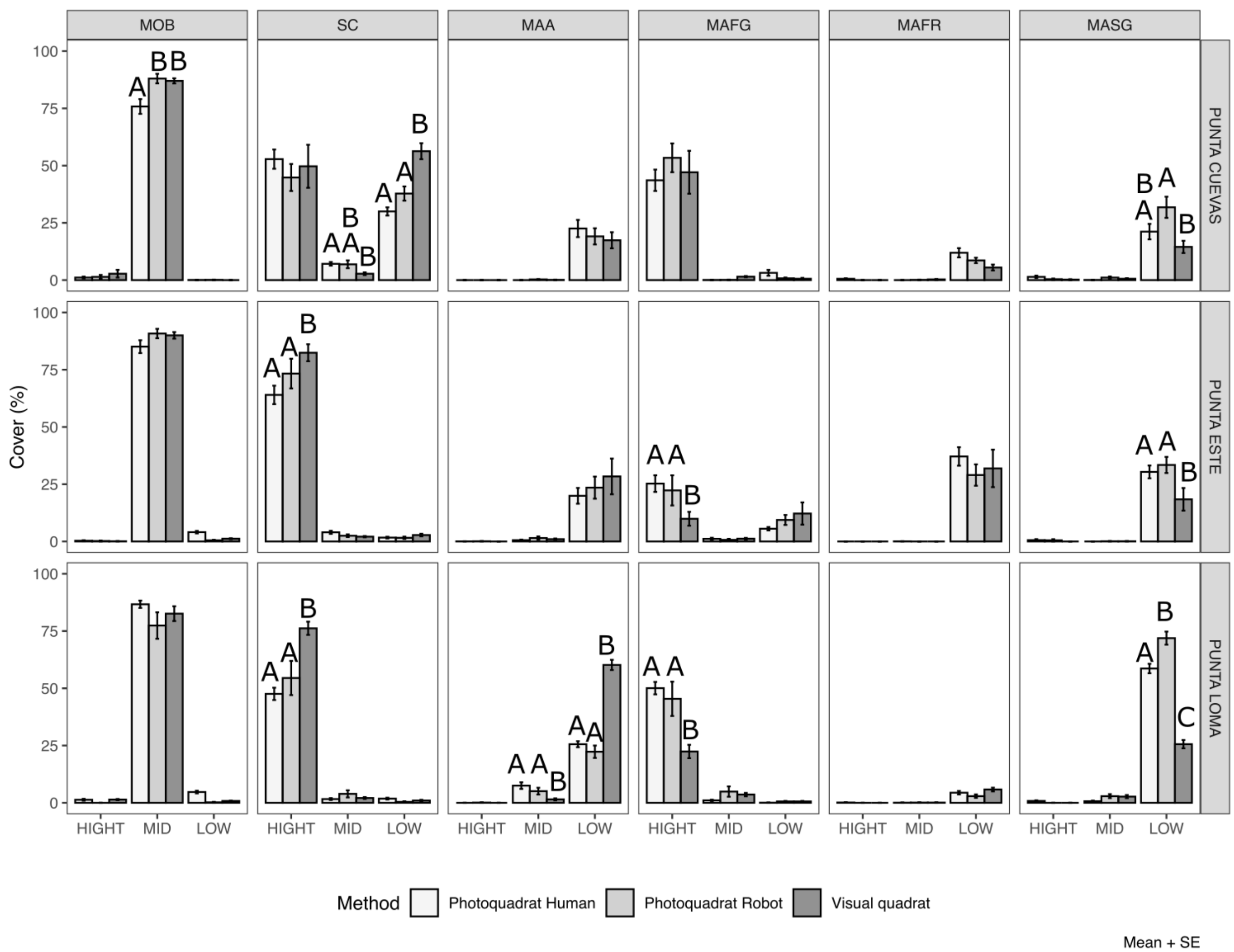

Figure 3: Diferences among method for percentage cover estimation of mayor CATAMI categories. Different letters indicate p<0.05 (Mann-Whitney test)

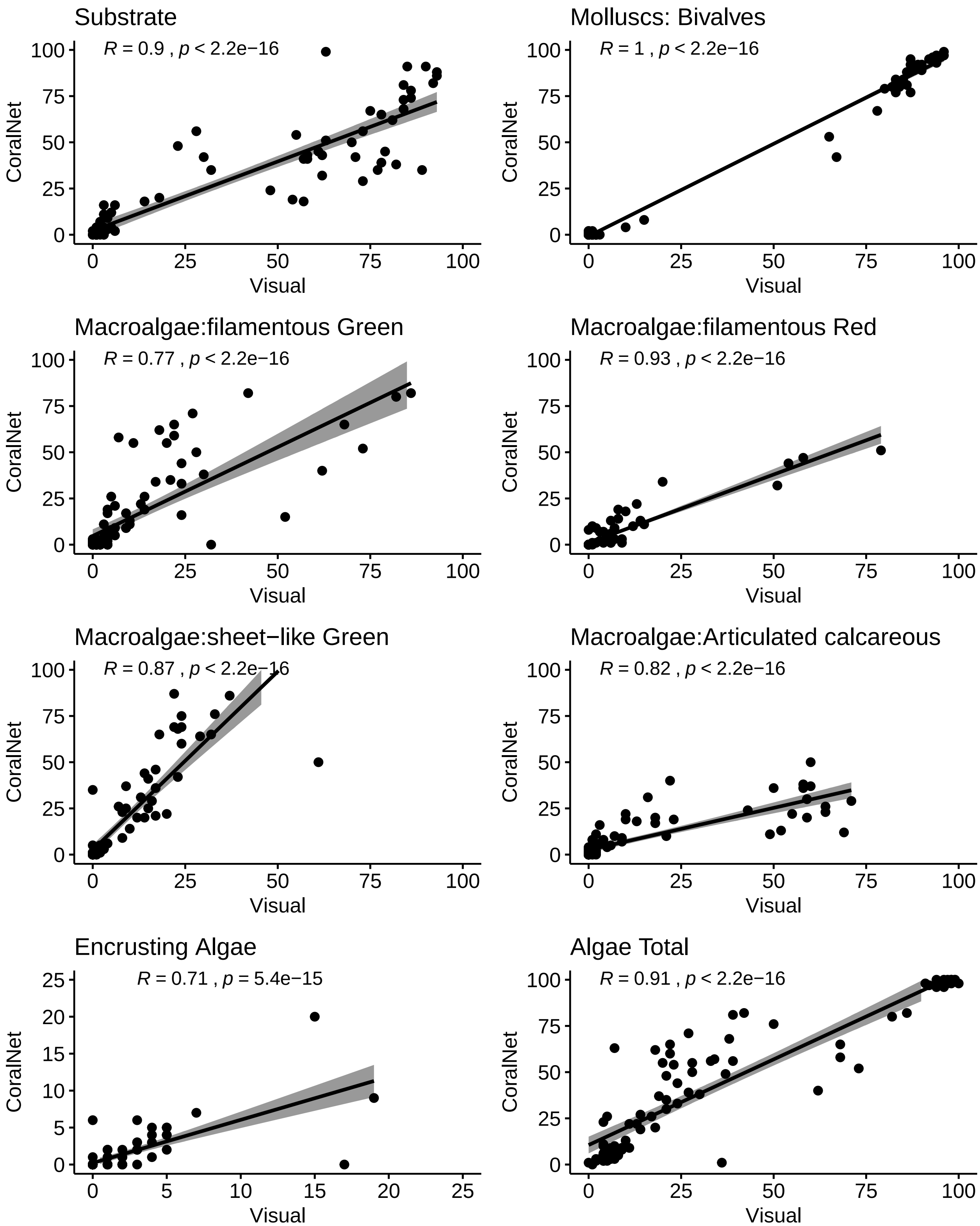

Figure 4: Correlations between CoralNet robot and visual estimations

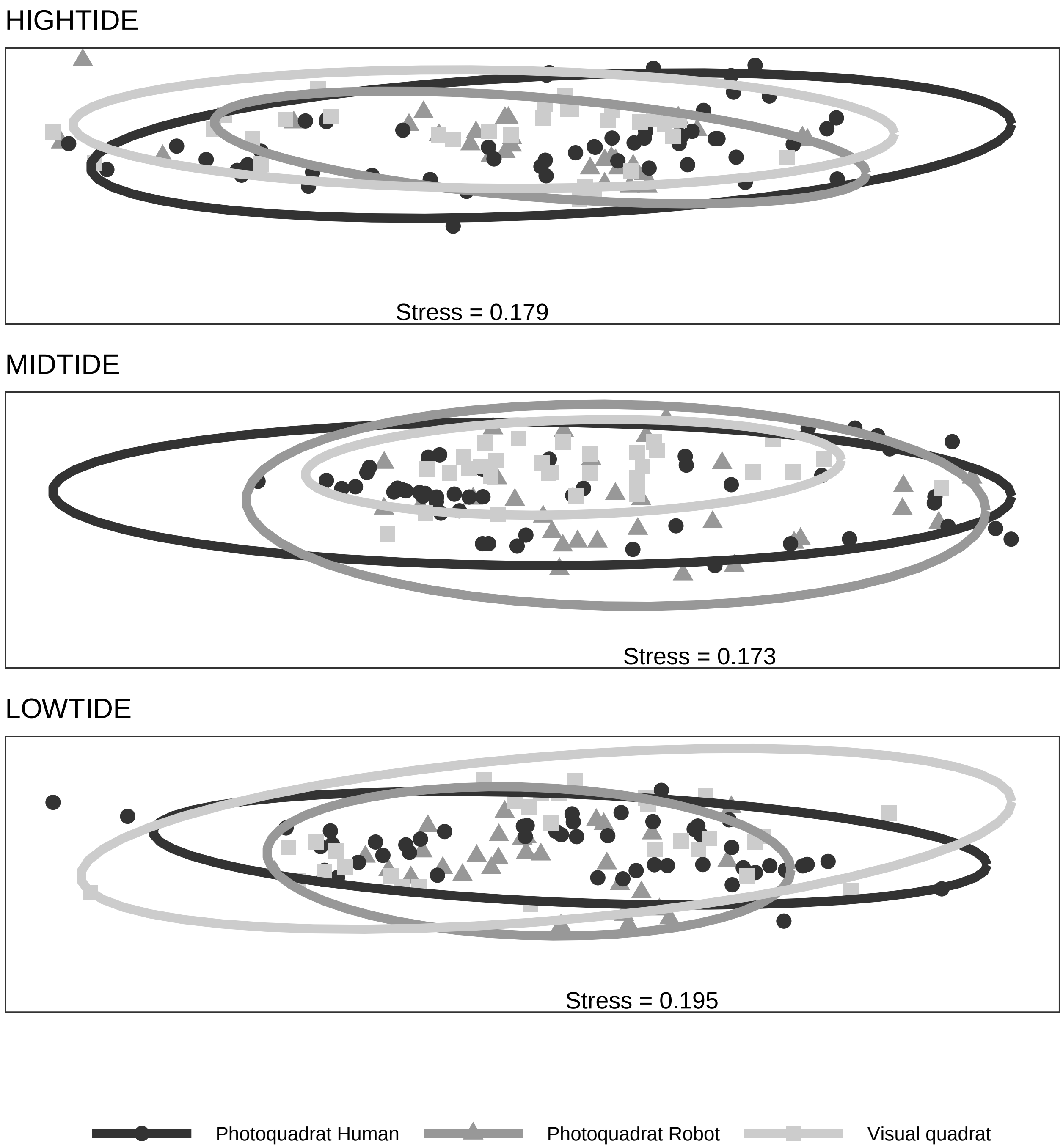
HIGHTIDE
Stress = 0.179
MIDTIDE
Stress = 0.173
LOWTIDE
Stress = 0.195
Photoquadrat Human
Photoquadrat Robot
Visual quadrat

Table 3: PERMANOVA Hightide

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	sig
Photoquadrat Robot vs Photoquadrat Human	1	0.0472122	0.9296186	0.0114868	0.343	1.000	
Photoquadrat Robot vs Visual quadrat	1	0.2763667	4.2708077	0.0685844	0.040	0.120	
Photoquadrat Human vs Visual quadrat	1	0.3986461	7.7778085	0.0886079	0.003	0.009	*

Table 4: PERMANOVA Midtide

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	sig
Photoquadrat Robot vs Photoquadrat Human	1	0.0473524	2.337965	0.0294684	0.092	0.276	
Photoquadrat Robot vs Visual quadrat	1	0.0237967	1.707691	0.0286008	0.196	0.588	
Photoquadrat Human vs Visual quadrat	1	0.0604702	3.948138	0.0487737	0.036	0.108	

Table 5: PERMANOVA Lowtide

pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	sig
Photoquadrat Robot vs Photoquadrat Human	1	0.1979604	1.852716	0.0232017	0.116	0.348	
Photoquadrat Robot vs Visual quadrat	1	1.0209450	8.239094	0.1243842	0.001	0.003	*
Photoquadrat Human vs Visual quadrat	1	0.8836800	6.973605	0.0820679	0.001	0.003	*

Code and data files

All the codes and data files are in github

https://github.com/gonzalobravoargentina/CoralNet_MBON