

Contents lists available at [ScienceDirect](http://ScienceDirect.com)

# Ecological Informatics

journal homepage: [www.elsevier.com/locate/ecolinf](http://www.elsevier.com/locate/ecolinf)

## Modelling benthic habitats and biotopes off the coast of Norway to support spatial management



Genoveva Gonzalez-Mirelis\*, Pål Buhl-Mortensen

Institute of Marine Research, Nordnesgaten 50, 5005 Bergen, Norway

### ARTICLE INFO

#### Article history:

Received 31 December 2014

Received in revised form 3 June 2015

Accepted 18 June 2015

Available online 25 June 2015

#### Keywords:

Benthic habitat

Biotope

Habitat modelling

Distribution modelling

Conditional Inference Forest

MAREANO

### ABSTRACT

Habitat conservation, and hence conservation of biodiversity hinges on knowledge of the spatial distribution of habitats, not least those that are particularly valuable or vulnerable. In offshore Norway, benthic habitats are systematically surveyed and described by the national programme MAREANO (Marine AREAL database for NORwegian waters). Benthic habitats and biotopes are defined in terms of the species composition of their epibenthic megafauna. Some habitats are of special conservation interest on account of their intrinsic value and/or vulnerability (e.g., long-lived species, rareness, to comply with international regulations such as OSPAR). In Norway, off Nordland and Troms, the following habitats of special interest can be found: *Umbellula encrinus* Stands, *Radicipes* sp. Meadows, Deep Sea Sponge Aggregations, Seapen and Burrowing Megafauna Communities, Hard Bottom Coral Gardens. In this paper, we used underwater video data collected within the MAREANO programme to define and describe benthic habitats and biotopes of special interest, and to map the geographic distribution thereof by means of habitat modelling.

We first evaluated the community structure of each habitat in the list using a SIMPROF test. We determined that the class Deep Sea Sponge Aggregations, as defined by OSPAR, had to be split into at least three classes. We then re-defined seven new types of ecological features, including habitats and biotopes that were sufficiently homogeneous. Then we modelled the spatial distributions of these habitats and biotopes using Conditional Inference Forests. Since the purpose of the distribution maps is to support spatial planning we classified the heat maps using density thresholds.

The accuracy of models ranged from fair to excellent. Hard Bottom Coral Gardens were the most rare habitat in terms of total area predicted (224 km<sup>2</sup>, 0.3% of the area modelled), closely followed by *Radicipes* Meadows (391 km<sup>2</sup>, 0.6%). Soft Bottom Demosponges (Geodid sponges and other taxa) represent the largest habitat, with a predicted area of 9288 km<sup>2</sup> (14%). Distribution maps of classes defined by habitat-forming species (Hard Bottom Coral Gardens) were more reliable than those defined by a host of species, or where no single species was a clear habitat provider (e.g. Seapen and Burrowing Megafauna Communities). We also put forward that a scale of patchiness larger than the scale of observation, and homogeneity of the community both play a role in model performance, and hence in map usefulness. These along with density threshold values based on observed data should all be taken into account in marine classifications and habitat definitions.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### 1. Introduction

Concepts like that of habitat have risen above others (e.g. species) in political agendas primarily because habitats can be used as surrogates for species distributions (Lindsay et al, 2008; Mumby et al., 2008) which are the ultimate target of conservation efforts, while being more readily described and documented. In addition, patterns in habitat distribution can be observed at the scale where management occurs. Hence, most governments now have a mandate to protect habitats

and/or related biological features (e.g. biotopes, ecosystems). Protecting habitats is however conditional on knowing their spatial distribution; therefore mapping the distribution of habitats paves the road to preserving them, which in turn leads to conservation of biological diversity.

When survey data are costly to obtain, as is the case in seabed environments, insight into the spatial distribution of habitats can be gained by use of spatially-explicit modelling, particularly distribution modelling at the community level (Ferrier and Guisan, 2006). This approach is similar to Species Distribution Modelling (Franklin, 2009) except a suite of species are modelled collectively, and it is also known as habitat modelling. First, biota–environment relationships are derived from a set of observational data. These relationships can easily be used to make predictions about the biological properties (e.g., the species

\* Corresponding author. Tel.: +47 55906504.

E-mail addresses: [genoveva.gonzalez-mirelis@imr.no](mailto:genoveva.gonzalez-mirelis@imr.no) (G. Gonzalez-Mirelis), [paal.buhl.mortensen@imr.no](mailto:paal.buhl.mortensen@imr.no) (P. Buhl-Mortensen).

composition) of sites where no direct observation of the biota has been made. When environmental data are available as map layers (i.e., full-coverage and sufficient resolution) a prediction can be made for every single pixel across a study area. In this way, ‘wall-to-wall’ maps of biota can be cost-effectively generated.

In the benthic realm the species composition of the epibenthic megafaunal community, comprising all organisms living on the sediment surface which are large enough to be visible to the naked eye, can construe habitats and biotopes. Megafaunal habitats, or simply benthic habitats are defined by two components: (1) an environmental setting characterized by the presence of one or more habitat-forming species and (2) the whole host of species that co-occur with it. This acceptance of the word “habitat” is absolute rather than relative. In other arenas habitat is understood as the type of environment where an organism lives, usually an animal, and is defined relative to that organism (hence the use of phrases such as “spotted-owl habitat”, or “Essential Fish Habitat”). Benthic ecologists, instead, used the term “associated fauna” to refer to the biota that utilize the resources and space procured by the habitat-forming species. The terrestrial concept closest in meaning would be that of vegetation types.

The term benthic habitat, however, has also been widely used in reference to areas that are simply homogeneous in terms of their geophysical (abiotic) environment, particularly in the hydroacoustic literature (e.g. Brown and Blondel, 2009, Jordan et al., 2005). This sense of the word implies that to assign an area to a type of habitat no data on the biota present in that area may be required; for example, “rocky reef”, or “coarse unconsolidated substrate”. When this is the case another concept is usually employed as a modifier, that of biotope, to designate units that can be characterized by a typical species assemblage in addition to the type of substrate (e.g. “*Mytilus edulis* beds on sublittoral sediment”, “*Phragmatopoma lapidosa* reefs on high energy sand”).

The biotope level is often the final (most-detailed) level available in marine classifications of ecosystems and communities (e.g. the Marine Habitat Classification for Britain and Ireland, Connor et al., 2004; the Coastal and Marine Ecological Classification Standard in the US, Federal Geographic Data Committee, 2012). Biotope is usually defined as the sum of habitat, which in this case is taken to be the abiotic component, and the species assemblage as the biotic component (Olenin and Ducrotoy, 2006). Note that if the habitat is best described by something of a biological, rather than geophysical nature (i.e., a habitat-forming species) the terms biotope and habitat become undistinguishable. In general, the term “habitat” can assume a wider range of meanings, including referring simply to a naturally occurring, relatively homogeneous type of nature.

Norway has a programme in place to systematically collect ecological data from the seabed and define, describe, and map the distribution of benthic habitats and biotopes beyond the continental platform: MAREANO (Marine AREAL database for NORwegian waters). Among other sampling tools the MAREANO programme operates an underwater video platform which yields geo-referenced, species occurrence data covering nearly the whole range of species of epibenthic megafauna. In this paper we use MAREANO video data to (1) test methods to define and describe benthic habitats and biotopes of management interest, and (2) map the geographic distribution of benthic habitats and biotopes by means of distribution modelling. The intended use of all the data products generated is to support spatial management of Norwegian offshore areas, with a view to conserving biological diversity rather than understanding patterns of distribution.

## 2. Material and methods

### 2.1. Biological data

Under the MAREANO programme Norway has conducted over 1300 video transects using the Campod and Chimaera platforms. On average the spacing between the video transects is 10 km (Buhl-Mortensen

et al., 2015, van Son et al., in preparation), although more complex areas can be sampled more densely. These transects are generally 700 m in length, occasionally longer. During each transect the video platform, carrying a forward-looking, high definition, colour video camera is towed behind the survey vessel at a speed of 0.7 knots, 1.5 m above the seabed. Laser pointers allow calculating the width of the field of view of every video frame. Navigation data is collected via a hydroacoustic positioning system and a transponder mounted on the video platform.

In the laboratory, video footage is translated into quantitative, species occurrence data. Every encountered organism is identified to the most-detailed possible taxonomical resolution (generally species level), and individually geo-referenced to approximately 5 m-accuracy. The abundance of each occurrence of megafauna is also quantified for the frame in view either as a count of individuals or percent cover, taxon by taxon. Following video analysis, data can be pooled into equal-length segments depending on the analysis scale of interest. Final abundance for all fauna of interest is reported as density, where area is calculated from the estimated values of field width and the length of the segment. Densities are standardized to numbers of individuals (or colonies, where relevant) per 100 m<sup>2</sup>.

In this paper we used data from video transects conducted off Nordland and Troms (Fig. 1) between 2006 and 2009. The area selection was dictated by management needs and is approximately 63,000 km<sup>2</sup>. The total number of video transects was 388. The species occurrence data was pooled at a 200-m scale. Each transect yielded 3–5 point-locality samples, and the total number of samples (*n*) was 1709.

Unfortunately, not all video data for the chosen area had been analyzed at the time of this study and a gap had to be filled in with spatial predictions larger than the general spacing between the video transects (see Fig. 1). The gap area is known to be similar in its biological composition to the adjacent areas and is not thought to contain any completely unsampled communities or landscapes. While representativeness is not believed to be compromised, predictions in that area will be affected by a larger degree of uncertainty.

### 2.2. Habitat selection and assessment of species structure

We used three commonly-used criteria to prioritize marine benthic habitats: being dominated by long-lived species, being unique in Norway, or simply, being included in the OSPAR list of Threatened and/or Declining Habitats (OSPAR, 2008). Habitat types used and the list of taxa whose presence prescribes the presence of the habitat (termed here “qualifying taxa”) are described in Table 1 and in the Supplementary material. They are all considered highly vulnerable to physical or mechanical damage by e.g. demersal fishing gear, and some host significant biological diversity.

After removing from the dataset all samples which did not contain any of the qualifying taxa we compared the species compositions between all pairs of samples within each class. We applied a SIMPROF (Similarity Profile) test (Clarke et al, 2008) to each subset. This test results in an assessment of whether there is any internal structure in a group of samples. When the test is not significant the group of samples is said to represent only one type of community (or is homogeneous in terms of its species composition) and further division is not warranted.

### 2.3. Modelling habitats

For each habitat that was deemed homogeneous we created a spatially-explicit dataset of the total abundance of all of the qualifying species pooled together, which was used as the response variable (total density of organisms/colonies of qualifying species). Samples where none of the qualifying species were found for each habitat were used as absence data. We reserved 677 samples for model evaluation.

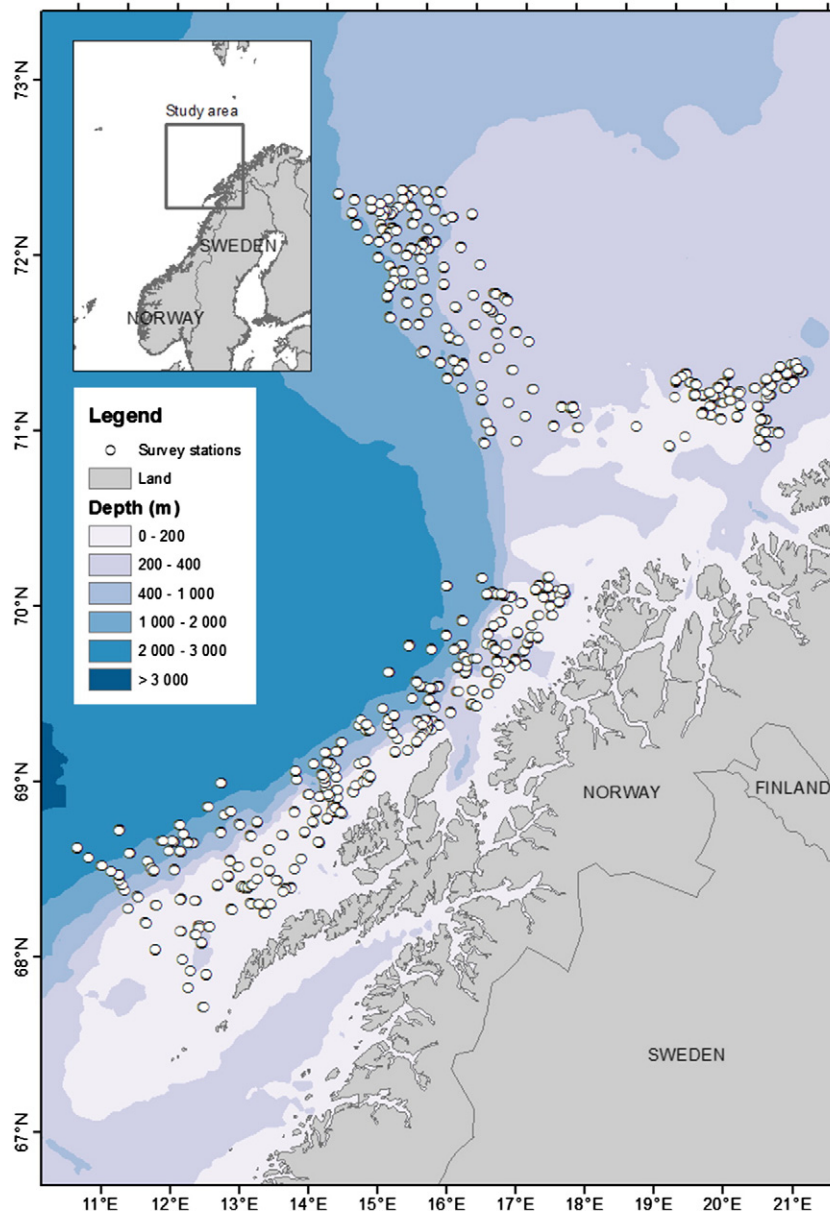


Fig. 1. Map showing the general location of the area and the layout of the 388 sampled sites (video-surveyed stations).

**Table 1**  
Habitat and biotope classes considered of special concern in offshore, Northern Norway.

Class	Qualifying taxa	Criterion
Umbellula Stands	<i>Umbellula encrinus</i>	Long-lived species
Radicipes Meadows	<i>Radicipes cf. gracilis</i>	Red-listed in Norway
Deep Sea Sponge Aggregations	Fam. Axinellidae, Geodiidae, Mycaliidae, Darwinellidae, Ancorinidae; class Hexactinellida	OSPAR
Seapen and Burrowing Megafauna Communities	<i>Funiculina quadrangularis</i> , <i>Kophobelemnon stelliferum</i> , <i>Pennatula</i> sp. (and other pennatulaceans), <i>Virgularia</i> spp.	OSPAR
Coral Gardens (particularly, Hard Bottom Coral Gardens)	<i>Paragorgia arborea</i> , <i>Paramuricea placomus</i> , <i>Primnoa resedaeformis</i>	OSPAR

These were selected randomly, with the additional condition that they be as geographically isolated as possible to avoid inflating model performance measures on account of spatial autocorrelation.

The predictor data was derived from geophysical data collected through hydroacoustic remote sensing techniques (multibeam echosounder), from which environmental proxies can be derived and stored digitally as Geographic Information System (GIS) layers. We obtained depth, slope, topographic complexity (surface area), and landscape type from bathymetry data. Layers on sedimentary environment, and dominant grain size were generated from backscatter data, aided by expert interpretation and ground-truthed by video data (for details see Bøe et al., 2010).

We used Conditional Inference Forests (CIF, Hothorn et al., 2006a, 2006b) to model habitat distribution. This method has been applied to modelling of vegetation types (Czúcz et al., 2011) and benthic biotopes (Gonzalez-Mirelis & Lindegarth, 2012). As a machine-learning method, it is data-driven rather than model-driven, and it is based on recursive partitioning. Recursive partitioning-based methods have become popular owing to their ability to solve classification

**Table 2**

Results from SIMPROF tests. A  $p$ -value  $< 0.05$  (represented by the symbol “\*\*”) means that samples represent more than one community.

Habitat	SIMPROF test $p$ -value ( $\alpha = 0.05$ )	Final habitat type
<i>Umbellula</i> Stands	0.07	<i>Umbellula</i> Stands
<i>Radicipes</i> Meadows	0.29	<i>Radicipes</i> Meadows
Deep Sea Sponge Aggregations	0.0001 (*)	Hard bottom + Demospongiae Soft bottom + Demospongiae Other deep sea sponges, including Hexactinellida (glass sponges)
Seapen and Burrowing Megafauna Communities	0.12	Seapen and Burrowing Megafauna Communities
Hard Bottom Coral Gardens	0.35	Hard Bottoms Coral Gardens

problems where data are multidimensional, explanatory variables are correlated, and relationships are non-linear (Strobl et al., 2008), as is the case in ecology.

For each habitat/biotope we: fit a CIF to the training data ( $n = 1032$ ); used this model to predict the probability of presence (of any of the species in the group) for every observation in the evaluation dataset; and checked the observed presences and absences in the evaluation dataset against the probabilities returned by the model to compute a measure of model performance (see below). We approximately repeated this procedure to generate spatial predictions, except that at this time we used the whole dataset ( $n = 1709$ ), and we set the model to predict values in the same scale as the response in the data (density).

Accuracy of the models was measured by the AUC (area under curve) of the ROC (receiver operating characteristic, Fielding and Bell, 1997). We have used the following system for classifying the accuracy of each model: 0.9–1 = excellent, 0.8–0.9 = good, 0.7–0.8 = fair, 0.6–0.7 = poor, and 0.5–0.6 = fail.

We implemented all analyses in R (R Core Team, 2012) with packages ‘vegan’ (Oksanen et al., 2013) and ‘party’ (Hothorn et al., 2006a, 2006b, Strobl et al., 2007, 2008).

#### 2.4. Mapping distributions for spatial management

Plotting the predictions from the model in geographic space results in a “heat map”. Management authorities, however, are often more interested in crisp boundaries, which in turn allow implementation of Marine Spatial Planning. Therefore, we had to convert the heat maps to binary maps. To achieve this a decision on an appropriate threshold value needs to be made.

We attempted to generate threshold values by splitting up all samples (within each habitat type) into density groups; fit a separate model for each subset (i.e., first using only low-density samples, then

intermediate-density samples, etc.); and plot model performance against density, where we expected an increase in performance at a given density value. The assumption behind this approach was that where the density of the species of interest is low the species should be present by chance and patterns should be harder to detect, just as if one were trying to model the distribution of a species outside of its range. Unfortunately, for most habitats the number of samples was not enough and this could only be implemented for the Seapens case. All other thresholds were picked arbitrarily, only in reference to the range of observed values and the total area predicted.

We also wanted to minimize overlap between the different habitats at any given location. Two conditions were then applied for a pixel  $n$  to be classified as habitat  $H$ : that the habitat of maximum predicted density for pixel  $n$  is  $H$ , and that density is larger than selected threshold for  $H$ . Lastly we ran a boundary clean algorithm, to remove single (or very small groups of) pixels.

### 3. Results

Deep Sea Sponge Aggregations were the only habitat that was found to be heterogeneous, and it was subdivided before being modelled. Subdivision was according to the type of dominant substrate (hard versus soft); a third sub-class of deep sea sponges was defined in terms of species composition (Table 2). These three new classes all had a homogeneous community structure (SIMPROF test not significant).

Soft bottom, deep sea sponges (typically Geodid sponges) and *Radicipes* sp. were the most densely-occurring qualifying species observed in this area (Table 3), with upwards of 200 individuals (or colonies) per 100 m<sup>2</sup>. Notwithstanding, sponges contribute much more biomass than pigtail corals.

Model accuracies ranged from fair to excellent. The results from the models for the final seven classes are summarized in Table 3, and shown in Figs. 2–8. Hard Bottom Coral Gardens were the most rare habitat in terms of total area predicted (224 km<sup>2</sup>, 0.3% of the area modelled), followed by *Radicipes* Meadows (391 km<sup>2</sup>, 0.6%). Hard Bottom Coral Gardens also had the smallest mean patch size (2 km<sup>2</sup>). At the other extreme, Soft Bottom Demosponges (Geodid sponges and other taxa) represent the largest habitat, with a predicted area of 9288 km<sup>2</sup> (14%). This habitat also had the largest mean patch size (52 km<sup>2</sup>).

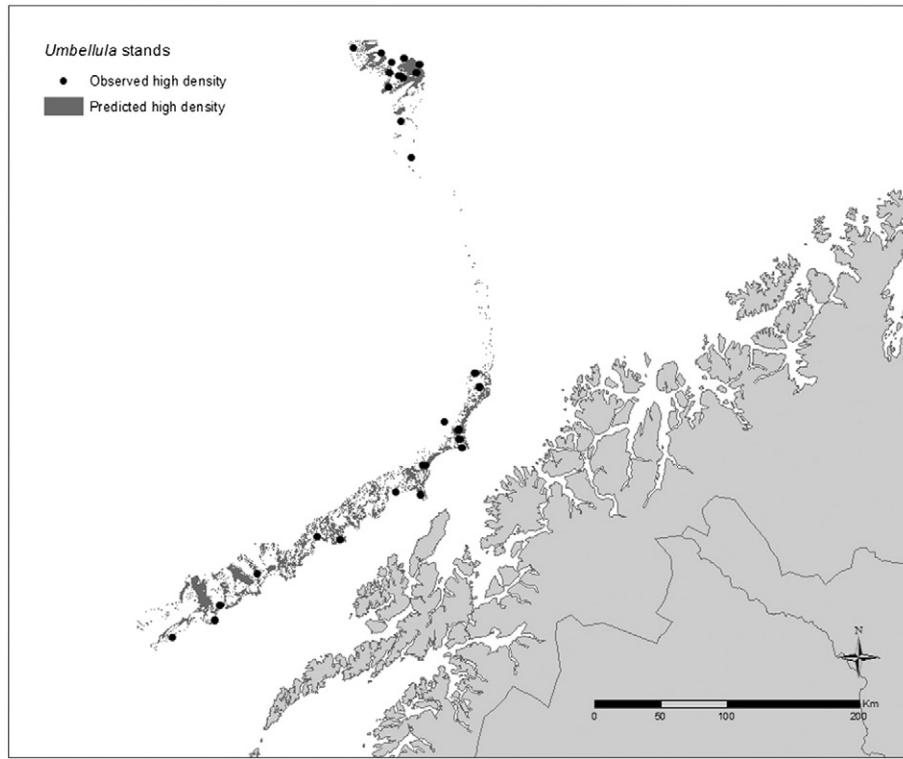
### 4. Discussion

Given sufficient data and the right modelling technique any type of ecological unit (from species to ecosystems) can be modelled in terms of its spatial distribution. Model interpretation aside, the resulting maps can be more, or less suited to support management decisions. One fundamental aspect is how reliable was the model? Here we have

**Table 3**

Modelling results, including model performance as area under curve (AUC), prevalence of the qualifying species in the dataset, maximum observed density, density threshold used to classify the heat maps, and variables that ranked first and second as explanatory variables.

Habitat name	AUC	Prevalence (%)	Maximum observed density (n/100 m <sup>2</sup> )	Threshold used (n/100 m <sup>2</sup> )	Top-two explanatory variables
<i>Umbellula</i> Stands	0.85	0.3	6.4	0.06	Surface area Depth
<i>Radicipes</i> Meadow	0.99	0.1	239.8	4.4	Landscape type sedimentary environment
Hard Bottom Demosponges	0.81	44.4	127.4	5	Depth Landscape type
Soft Bottom Demosponges	0.80	37.7	258.3	6	Landscape type Grain size
Other deep sea sponges, including glass sponges	0.77	9.0	30.6	1	Sedimentary environment Landscape type
Seapens and Burrowing Megafauna	0.73	14.0	52.0	2	Landscape type Grain size
Hard Bottom Coral Gardens	0.85	4.0	44.4	0.26	Landscape type Surface area

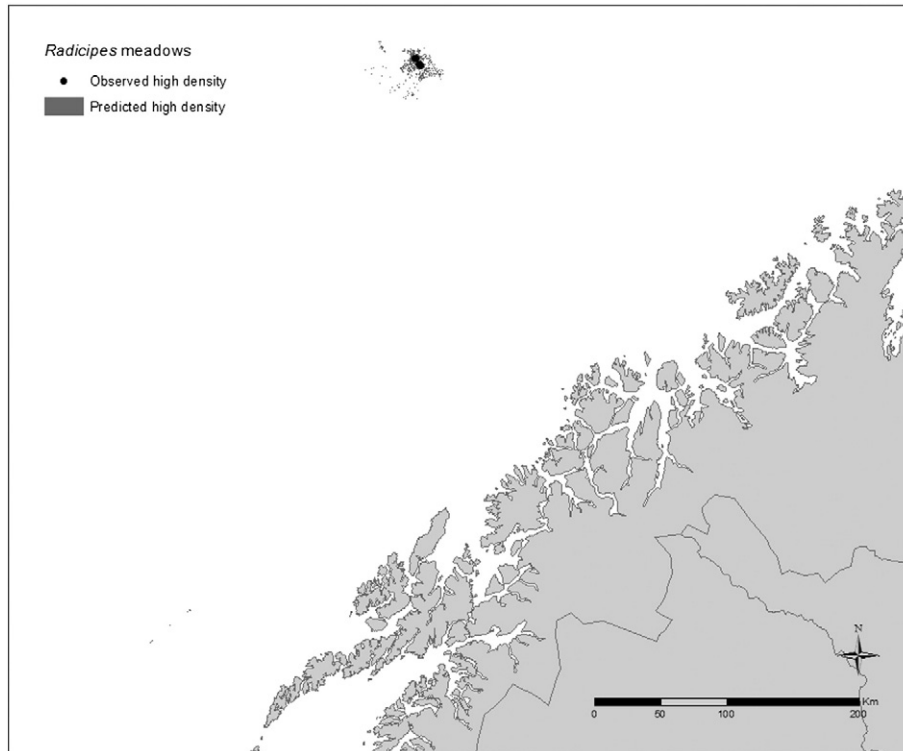


**Fig. 2.** Distribution of *Umbellula* Stands; model performance is considered “good”. “High density” is defined as above the threshold used for this biotope (0.06 colonies/100 m<sup>2</sup>), whether observed or predicted.

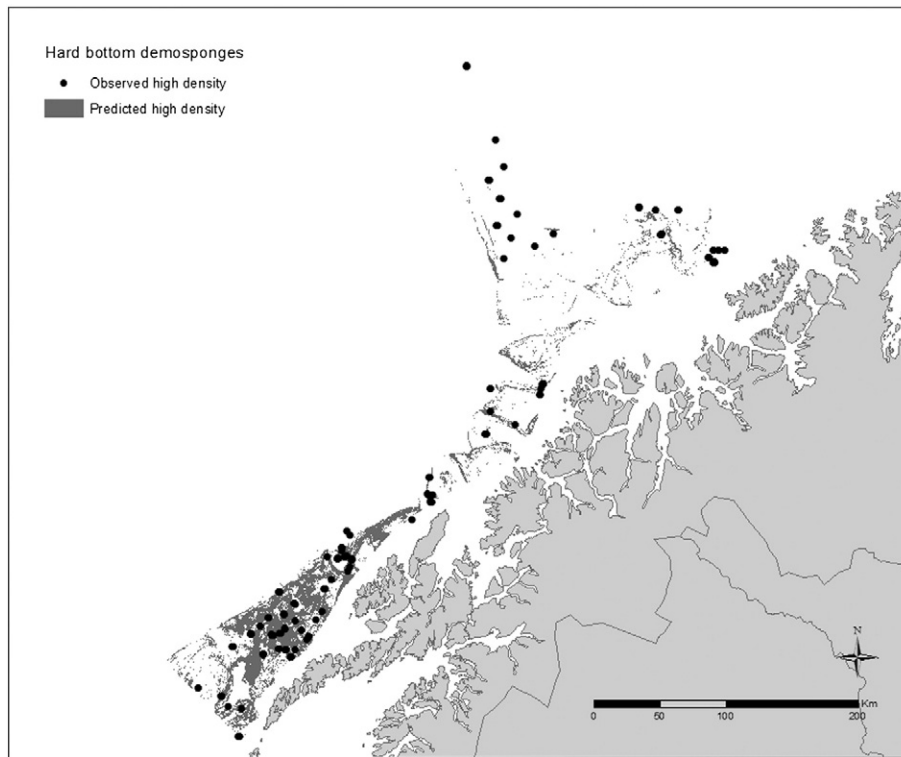
measured model performance by its AUC, which provides an indication of how well the model discriminates presence from absence.

The best performing models were those for *Radicipes* Meadows, and *Umbellula* Stands with AUC values of 0.99, and 0.85. These were

conventional Species Distribution Models and their high performance is a result (among other factors) of models dealing with a single species and thus a uniform response (Zimmermann and Kienast, 1999). The high accuracy of the *Radicipes* model can also be partly due to the low



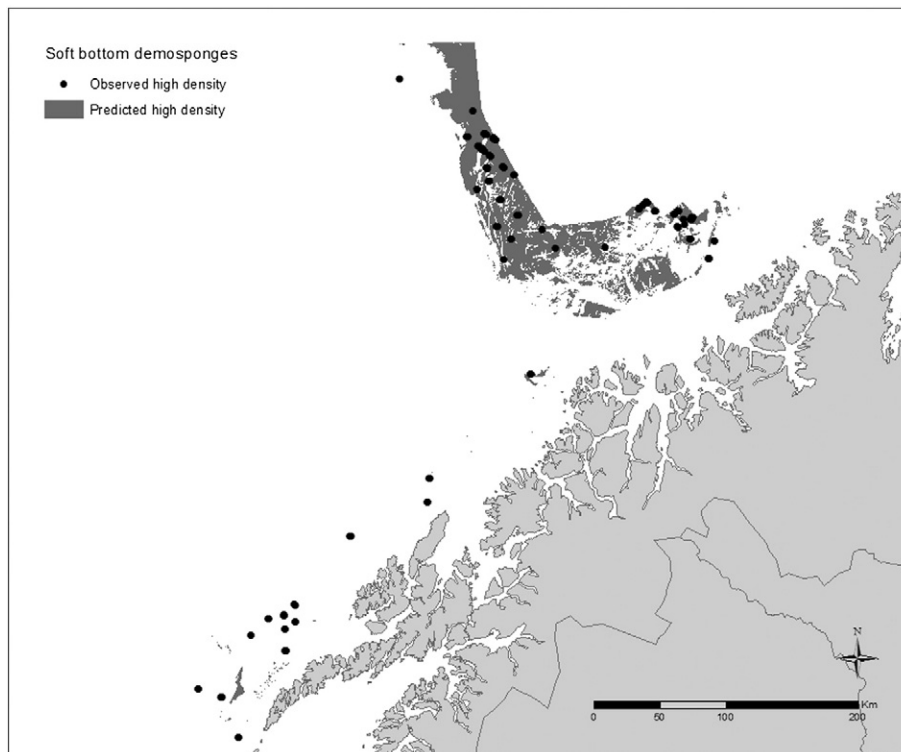
**Fig. 3.** Distribution of *Radicipes* Meadows; model performance is considered “excellent”. “High density” is defined as above the threshold used for this biotope (4.4 colonies/100 m<sup>2</sup>), whether observed or predicted.



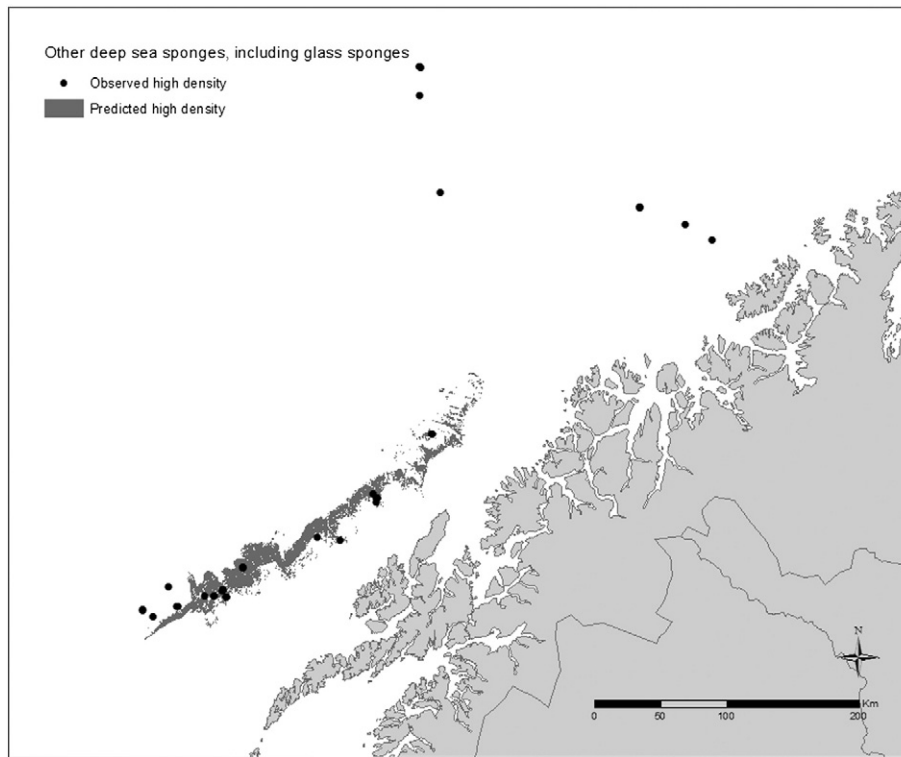
**Fig. 4.** Distribution of hard-bottom, demosponge habitat; model performance is considered “good”. “High density” is defined as above the threshold used for this habitat (5 individuals/100 m<sup>2</sup>), whether observed or predicted.

prevalence of this taxon (0.1%, Table 3), which makes it easy to predict its absence. However, it is in fact known that this biotope is not present outside the Bjørnøya slide area, as the data extends well beyond that locality. The limited distribution of this species and the fact that where it

occurs it forms extensive, dense meadows makes this species easy to model and the model particularly useful to define the boundaries of the distribution range of this red-listed biotope (Lindgaard and Henriksen, 2011).



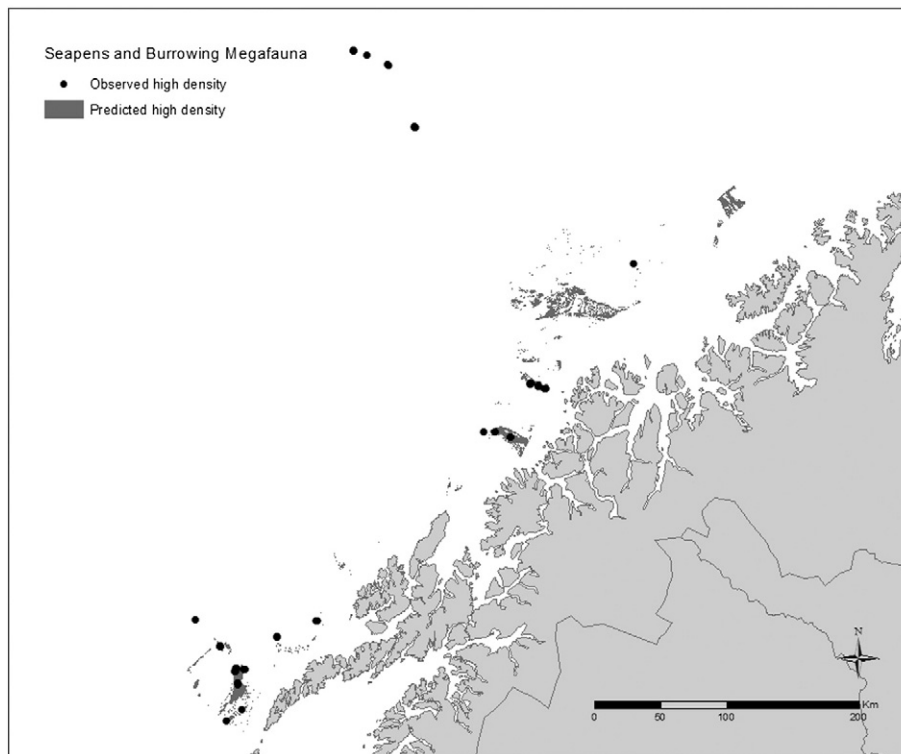
**Fig. 5.** Distribution of soft-bottom, demosponge habitat with shedded spicule carpet. This habitat is also known as *Ostur*. Model performance is considered “good”. “High density” is defined as above the threshold used for this habitat (6 individuals/100 m<sup>2</sup>), whether observed or predicted.



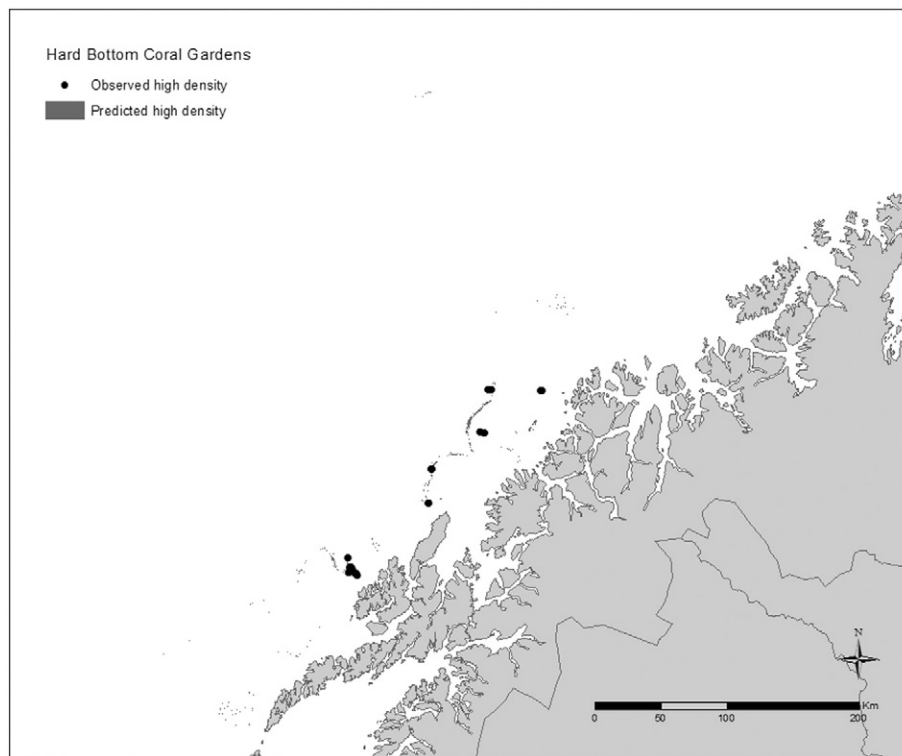
**Fig. 6.** Distribution of deep sea sponges including glass sponges; model performance is considered “fair”. “High density” is defined as above the threshold used for this habitat (1 individual/100 m<sup>2</sup>), whether observed or predicted.

We put forward that classes characterized by habitat-forming species are easier to model. Indeed, the next best model was Hard Bottom Coral Gardens (AUC = 0.85). Only hard bottom gorgonians can form habitats in the sense we have used here: that of providing

habitat for a range of larger organisms at the landscape scale (e.g. [Buhl-Mortensen and Mortensen, 2005](#), [Buhl-Mortensen et al., 2010](#)), although certain types of Deep Sea Sponge Aggregations may be said to provide similar functionality. Habitat-forming organisms can cover



**Fig. 7.** Distribution of Seapen and Burrowing Megafauna; model performance is considered “fair”. “High density” is defined as above the threshold used for this biotope (2 colonies/100 m<sup>2</sup>), whether observed or predicted.



**Fig. 8.** Distribution of Hard Bottom Coral Gardens; model performance is considered “good”. “High density” is defined as above the threshold used for this habitat (0.26 colonies/100 m<sup>2</sup>), whether observed or predicted.

the whole surface of an area (and fill up a large volume over the seabed) at this scale (hundreds of metres) leaving only small, randomly-distributed gaps, much like vegetation does. A situation where the scale of analysis is smaller than the size of the patch of the feature of interest is known to be the optimal scenario to detect patterns from observational data (Fortin and Dale, 2005). The predictions from this model can aid both in understanding the factors shaping the patterns of distribution, and in selecting target areas for further investigation and ultimately for protection of hard bottom gorgonian corals and Coral Gardens.

Without a clear habitat-forming species, generally with large gaps between occurrences (although extreme densities can also be observed), and prescribed by a suite of different species it is perhaps not surprising that the model for Seapens and Burrowing Megafauna had the lowest performance (AUC = 0.73). This model simply confirmed the presence of these taxa in marine valleys filled with fine sediment (notice that the most important variables were landscape type and grain size, Table 2). OSPAR-signatory countries, which have a responsibility to preserve this biotope, would need to make efforts to: better define the biotope, and model it at the appropriate, possibly regional scale.

A non-significant SIMPROF test effectively means here that only by picking samples with the qualifying taxa present the same community is being sampled; from this point of view, all of our qualifying taxa can be called “indicator species”. By the same token, all these habitats and biotopes are good surrogates for the qualifying taxa. If we further assume that the level of significance of the SIMPROF test is indicative of the degree of homogeneity of the habitat, the Hard Bottom Coral Gardens also have the most homogenous community structure ( $p$ -value = 0.35, Table 1) Table 3. Bullimore et al (2013) found that Coral Gardens off the UK encompass a wide range of different communities, concluding that the term is ambiguous. Here, we have modelled one specific type of Coral Gardens, and the exercise proved useful. Along the same lines as Bullimore et al. (2013) we also found that Deep Sea Sponge Aggregations are not adequately defined by OSPAR, and that they contain at least three distinct habitat types, including the

community known as *ostur* characterized by the presence of Geodid sponges on soft substrate with a sponge spicule cover. Once the class was divided, the three resulting classes turned out to have non-overlapping spatial distributions (Figs. 4–6), which in turn responded to different sets of environmental variables (Table 3). *Ostur* is predicted (and known) to occur in massive, dense aggregations at the Egga shelf and Tromsøflaket area (Fig. 5).

Direct or indirect coupling between predictors and responses is also needed for increased model performance and these links can vary with spatial resolution (e.g. Svensson et al., 2013). While it is most common to use terrain variables to predict habitat (e.g. Rattray et al., 2009, Ross and Howell, 2013), in this paper we have used surficial geology variables (grain size and sedimentary environment), and at least one of them was among the top-two predictors in four of the six classes modelled, highlighting the importance of using substrate data as predictors.

Another factor playing an important role in the applicability of the final map is what threshold was used to delineate the boundary of the habitat? Other authors have said that the density of the qualifying species needs to be at least 10 times the background density (Rogers and Gianni, 2009), but the problem remains as to how to measure the background density? OSPAR has made attempts to make the density of colonies or individuals a criterion in the definition of habitats (e.g. for Coral Gardens, Christiansen, 2010). This approach has not succeeded owing to difficulties in adequately quantifying the variation of density across space (Bullimore et al., 2013; ICES, 2007) but should such a figure be available it would be straightforward to use it to classify the spatial predictions from a model.

In conclusion, for mapping (and ultimately conservation) purposes it is fundamental that classes, whether habitats or biotopes are clear natural units with homogeneous species composition and indicator species that are representative of their distributions, and to adapt the scale of observation to the scale of the patches they form. Using the relevant variables should prove most useful in model interpretation but it



significantly improves models. Threshold values based on observed variation of density across space urgently need to be incorporated into habitat and biotope definitions so that they can be used to classify maps objectively.

### Acknowledgements

This work was entirely funded by the MAREANO programme (Norway) which is supported by the Norwegian Ministry for the Environment and the Ministry of Trade, Industry and Fisheries. The analysis of video footage was largely carried out by Gjertrud Jensen and Anne Sveistrup. Margaret Dolan and Valérie Bellec (Geological Survey of Norway) generated the surficial geology layers from backscatter data.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.ecoinf.2015.06.005>.

### References

- Brown, C.J., Blondel, P., 2009. Developments in the application of multibeam sonar backscatter for seafloor habitat mapping. *Appl. Acoust.* 70 (10), 1242–1247.
- Buhl-Mortensen, L., Mortensen, P., 2005. Distribution and diversity of species associated with deep-sea gorgonian corals off Atlantic Canada. In: Freiwald, A., Roberts, J.M. (Eds.), *Cold-Water Corals and Ecosystems*. Springer-Verlag, Berlin, pp. 849–879.
- Buhl-Mortensen, L., Buhl-Mortensen, P., Dolan, M., Gonzalez-Mirelis, G., 2015. Habitat mapping as a tool for conservation and sustainable use of marine resources: some perspectives from the MAREANO programme, Norway. *J. Sea Res.* 100, 46–61.
- Buhl-Mortensen, L., Vanreusel, A., Gooday, A.J., Levin, L.A., Priede, I.G., Buhl-Mortensen, P., Gheerardyn, H., King, N.J., Raes, M., 2010. Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. *Mar. Ecol.* 31, 21–50.
- Bøe, R., Dolan, M., Thorsnes, T., Lepland, A., Olsen, H., Totland, O., Elvenes, S., 2010. Standard for geological seabed mapping offshore. *Geol. Surv. Norway* 15–pp (Report 2010.033).
- Bullimore, R.D., Foster, N.L., Howell, K.L., 2013. Coral-characterized benthic assemblages of the deep Northeast Atlantic: defining “Coral Gardens” to support future habitat mapping efforts. *ICES J. Mar. Sci.* 70, 511–522.
- Christiansen, S., 2010. Background Document for Coral Gardens. OSPAR commission 978-1-907390-27-2 (Publication Number: 486/2010, 39 pp.).
- Clarke, K.R., Somerfield, P.J., Gorley, R.N., 2008. Testing of null hypotheses in exploratory community analyses: similarity profiles and biota-environment linkage. *J. Exp. Mar. Biol. Ecol.* 366, 56–69.
- Connor, D.W., Allen, J., Golding, N., Howell, K., Lieberknecht, L., Northen, K.O., Reker, J., 2004. The marine habitat classification for Britain and Ireland, Version 04.05. Technical report. JNCC, Peterborough. ISBN: 1 861 07561 8 (internet version, <http://www.jncc.gov.uk/MarineHabitatClassification/>).
- Czúcz, B., Gálhidy, L., Mátyás, C., 2011. Present and forecasted xeric climatic limits of beech and sessile oak distribution at low altitudes in Central Europe. *Ann. For. Sci.* 68, 99–108.
- Ferrier, S., Guisan, A., 2006. Spatial modelling of biodiversity at the community level. *J. Appl. Ecol.* 43, 393–404.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24, 38–49.
- Fortin, M.-J., Dale, M.R.T., 2005. *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press, New York 365 pp.
- Franklin, J., 2009. *Mapping Species Distribution: Spatial Inference and Prediction*. Cambridge University Press, New York 336 pp.
- Gonzalez-Mirelis, G., Lindegarth, M., 2012. Predicting the distribution of out-of-reach biotopes with decision trees in a Swedish marine protected area. *Ecol. Appl.* 22, 2248–2264.
- Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., van der Laan, M., 2006a. Survival ensembles. *Biostatistics* 7 (3), 355–373.
- Hothorn, T., Hornik, K., Zeileis, A., 2006b. Unbiased recursive partitioning: a conditional inference framework. *J. Comput. Graph. Stat.* 15, 651–674.
- ICES, 2007. Report of the Working Group on Deep-Water Ecology (WGDEC). 26–28 February 2007. 2007. ICES Advisory Committee on Ecosystems (ICES Document CM 2007/ACE: 01).
- Jordan, A., Lawler, M., Halley, V., Barrett, N., 2005. Seabed habitat mapping in the Kent Group of islands and its role in marine protected area planning. *Aquat. Conserv. Mar. Freshwat. Ecosyst.* 15 (1), 51–70.
- Lindgaard, A., Henriksen, S. (Eds.), 2011. *The 2011 Norwegian Red List for Ecosystems and Habitat Types*. Norwegian Biodiversity Information Centre, Trondheim.
- Lindsay, M.J., Patterson, H.M., Swearer, S.E., 2008. Habitat as a surrogate measure of reef fish diversity in the zoning of the Lord Howe Island Marine Park, Australia. *Mar. Ecol. Prog. Ser.* 353, 265–273.
- Mummy, P.J., Broad, K., Brumbaugh, D.R., Dahlgren, Craig, Harborne, P., Hastings, A.R., Holmes, A., Kappel, K.E., Micheli, C.V., F. & Sanchirico, J. N., 2008. Coral reef habitats as surrogates of species, ecological functions, and ecosystem services. *Conserv. Biol.* 22 (4), 941–951.
- Oksanen, J.F., Blanchet, G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Wagner, H., 2013. *Vegan: community ecology package*. R package version 2.0-6 (<http://CRAN.R-project.org/package=vegan>).
- Olenin, S., Ducrottoy, J.-P., 2006. The concept of biotope in marine ecology and coastal management. *Mar. Pollut. Bull.* 53, 20–29.
- OSPAR, 2008. OSPAR list of threatened and/or declining species and habitats. OSPAR Agreement 2008-06pp. 1–4.
- Core Team, R., 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria 3-900051-07-0 (URL <http://www.R-project.org/>).
- Rattray, A., Ierodiaconou, D., Laurenson, L., Burq, S., Reston, M., 2009. Hydro-acoustic remote sensing of benthic biological communities on the shallow South East Australian continental shelf. *Estuar. Coast. Shelf Sci.* 84, 237–245.
- Rogers, A., Gianni, M., 2009. Implementation of UN resolution 61/105 in the management of deep-sea fisheries on the high seas. North Atlantic status and recommendations. Technical report, International Programme on the State of the Ocean, Provisional Report.
- Ross, R.E., Howell, K.L., 2013. Use of predictive habitat modelling to assess the distribution and extent of the current protection of 'listed' deep-sea habitats. *Divers. Distrib.* 19, 433–445. <http://dx.doi.org/10.1111/ddi.12010>.
- Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. *Bioinformatics* 9, 307.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinforma.* 8 (25) URL <http://www.biomedcentral.com/1471-2105/8/25>.
- Svensson, J.R., Jonsson, L., Lindegarth, M., 2013. Excessive spatial resolution decreases performance of quantitative models, contrary to expectations from error analyses. *Mar. Ecol. Prog. Ser.* 485, 57–73 06/2013.
- van Son, T.C., Dolan, M., Gonzalez-Mirelis, G., Thorsnes, T., Bjarnadóttir, L.R., Buhl-Mortensen, P., 2015. Environmental Variability Index (EVI) - a MAREANO methods study for guidance of sampling effort. NGU Report 2015 027, 55 pp.
- Zimmermann, N., Kienast, F., 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *J. Veg. Sci.* 10, 469–482.