

36 ABSTRACT

37 In the deep sea, biological data are often sparse; hence models capturing relationships
38 between observed fauna and environmental variables (acquired via acoustic mapping
39 techniques) are often used to produce full coverage species assemblage maps. Many
40 statistical modelling techniques are being developed, but there remains a need to determine
41 the most appropriate mapping techniques. Predictive habitat modelling approaches
42 (redundancy analysis, maximum entropy and random forest) were applied to a heterogeneous
43 section of seabed on Rockall Bank, NE Atlantic, for which landscape indices describing the
44 spatial arrangement of habitat patches were calculated. The predictive maps were based on
45 remotely operated vehicle (ROV) imagery transects, high-resolution autonomous underwater
46 vehicle (AUV) sidescan backscatter maps and ship-based multibeam bathymetry. Area under
47 the curve (AUC) and accuracy indicated similar performances for the three models tested, but
48 performance varied by species assemblage, with the transitional species assemblage showing
49 the weakest predictive performances. Spatial predictions of habitat suitability differed
50 between statistical approaches, but niche similarity metrics showed redundancy analysis and
51 random forest predictions to be most similar. As one statistical technique could not be found
52 to outperform the others when all assemblages were considered, ensemble mapping
53 techniques, where the outputs of many models are combined, were applied. They showed
54 higher accuracy than any single model. Different statistical approaches for predictive habitat
55 modelling possess varied strengths and weaknesses and by examining the outputs of a range
56 of modelling techniques and their differences, more robust predictions, with better described
57 variation and areas of uncertainties, can be achieved. As improvements to prediction outputs
58 can be achieved without additional costly data collection, ensemble mapping approaches have
59 clear value for spatial management.

60 **KEYWORDS:** Cold-water corals, Deep sea, Ensemble approaches, Habitat mapping,
61 Megabenthos

62 **1. INTRODUCTION**

63 As the anthropogenic footprint extends deeper into our oceans, reliable descriptions of
64 the seafloor and the species present are required to devise appropriate management and
65 conservation measures. With very limited areas of seafloor mapped at comparable resolution
66 to terrestrial environments (Sandwell et al., 2006), quantitative spatial information regarding
67 distributions of marine biotic and abiotic components is needed to build benthic habitat maps
68 (Kostylev et al., 2001). Recent advances in acoustic techniques for seafloor mapping (Brown
69 et al., 2011) have made it possible to create detailed geomorphological maps more rapidly.
70 However, the biological information needed to supplement complete coverage topographic
71 and geological maps has remained limited owing to the time-consuming process of specimen
72 collection and taxonomic identification (Przeslawski et al., 2011).

73 Full coverage biological sampling is often not an option, and hierarchical approaches
74 involving nested survey designs are often employed. They involve a combination of broader-
75 scale geological map creation based on acoustic data, and detailed ground-truthing biological
76 studies covering smaller spatial extents, often taking the form of imagery transects (Elvenes
77 et al., 2014; Robert et al., 2015). These broader-scale geological maps can be used to define
78 habitat patches allowing the relationships between the spatial arrangement of these patches
79 within the surrounding landscape and their effect on species spatial patterns (Turner and
80 Gardner, 1991) to be examined, modelled and used to make biological predictions across the
81 larger extent covered by the acoustic surveys. The spatial arrangement of habitat patches can
82 be described using a variety of class and landscape metrics, the former used to describe
83 properties of patches from a single habitat type while the latter are used to characterise all
84 patches present within a landscape (McGarigal et al., 2012). Although such metrics have
85 been shown to help explain species spatial patterns (Teixidó et al., 2002), they have so far
86 rarely been employed for predictive mapping.

87 In recent years, there have been an increasing number of studies employing a variety
88 of techniques to produce predictive full coverage megabenthic invertebrate habitat maps:
89 maximum entropy (Rengstorf et al., 2012; Ross and Howell, 2012), many types of decision or
90 classification trees (Compton et al., 2013; Gonzalez-Mirelis and Lindegarth, 2012), a variety
91 of multivariate analyses or ordination methods (Buhl-Mortensen et al., 2012; Shumchenia
92 and King, 2010), general additive models, neural networks (Palialexis et al., 2011) and many
93 more. Some of these techniques, such as maximum entropy, are based on records of presence
94 only (with background points), as obtaining reliable absence data can be particularly difficult
95 (Pearce and Boyce, 2006). However, when absence data is available, presence-absence
96 models (such as general linear/additive models or classification trees) can provide more
97 information regarding unsuitable habitats (Brotons et al., 2004; Pearson et al., 2006) and
98 avoid the difficulties associated with selecting appropriate background points (Phillips et al.,
99 2009).

100 In this study, we used benthic imagery data (photographs and extracted video frames),
101 in addition to acoustic maps to produce predictive maps for megabenthic invertebrate species
102 assemblages, specifically demonstrating the usefulness of class and landscape indices to
103 improve prediction results. The prediction accuracy of three different modelling approaches
104 was investigated for highly heterogeneous sections of seabed mapped at very high resolutions
105 (~0.5 m pixel size): redundancy analysis (RDA) (ter Braak, 1994), maximum entropy
106 (MaxEnt) (Phillips and Dudík, 2008) and random forest (RF) (Breiman, 2001). As these
107 techniques represent three very different modelling approaches ('assemble and predict
108 together', RDA; 'assemble first, predict later' using a presence only model, MaxEnt; as well
109 as a presence-absence model, RF) (Ferrier and Guisan, 2006), the aim was to determine
110 whether, and which, a single approach may be most appropriate when considered across
111 multiple species assemblages. Based on the results obtained, we also examined whether

112 ensemble maps, which take into account predictions and uncertainties from more than one
113 model (Araújo and New, 2007; Marmion et al., 2009b), could further improve predictions.

114 **2. MATERIALS AND METHODS**

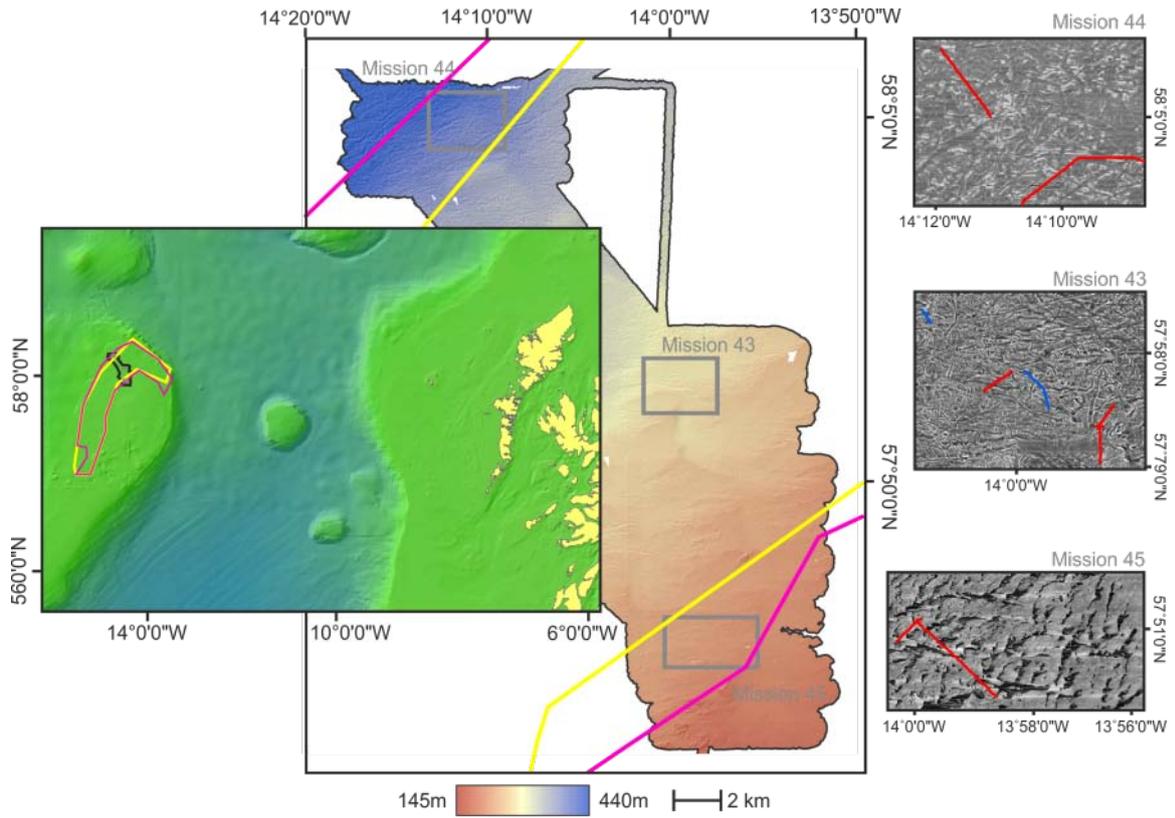
115 **2.1 Survey Design**

116 As part of the 'UK Marine Environmental Mapping Programme' (MAREMAP;
117 <http://www.maremap.ac.uk/index.html>) and the 'COMplex Deep-sea Environments: Mapping
118 habitat heterogeneity As Proxy for biodiversity' project (CODEMAP;
119 <http://www.codemap.eu/>), a section of the western flank of Rockall Bank (200-400 m),
120 Northeast Atlantic, was mapped during the RRS *James Cook* 060 cruise carried out in May -
121 June 2011 (Figure 1). Over 380 km² of ship-based multibeam bathymetry (pixel size of
122 10x10 m), three 12-13 km² Autosub6000 autonomous underwater vehicle (AUV, missions
123 M43, M44 and M45) sidescan sonar surveys (pixel size of 0.5x0.5 m) and five *Lynx* remotely
124 operated vehicle (ROV) photographic imagery (2592x1944 pixels) transects (1,222 images
125 along ~8 km using a Kongsberg OE14-208 digital stills camera) were collected. During the
126 *JC-073* cruise carried out in June 2012 as part of the UK Ocean Acidification programme's
127 'Changing Oceans Expedition', two additional *Holland I* ROV high-resolution (1920x1080
128 pixels) video transects (Insite Mini Zeus camera with direct HDSDI fibre output) were
129 carried out within the more heterogeneous area (M43) surveyed by the AUV. Frames were
130 extracted at a rate representing the distance separating the previously collected digital stills (~
131 1 frame per 5 m, 514 images).

132 The high resolution AUV surveys were positioned in areas of proposed boundary
133 changes to a conservation zone (Figure 1). In 2007, a Fisheries Closure was established by
134 the North East Atlantic Fisheries Commission (NEAFC Recommendation IX-2008, EC
135 Regulation No 40/2008) based on reports of cold-water coral occurrence. In 2011, a nearly

136 overlapping, but slightly extended area was put forth as candidate Special Area of
137 Conservation (cSAC) with the main aim of protecting stony and biogenic reefs (JNCC,
138 2010), a habitat listed under Annex I of the Habitat Directive (92/43/EEC). As such, AUV
139 mapping was conducted in areas outside of the Fisheries Closure, but still inside the cSAC
140 (M44 and M45) as well as inside both protected areas (M43), to identify the status of the
141 seabed habitats. ROV imagery transects were positioned to sample a variety of sediment
142 types within each of the three areas, including areas of high backscatter likely to harbour
143 cold-water corals. To reduce the influence of spatial autocorrelation, images were
144 systematically subsampled into 8 groups in which neighbouring pictures were located at a
145 distance of 40m (Figure 2).

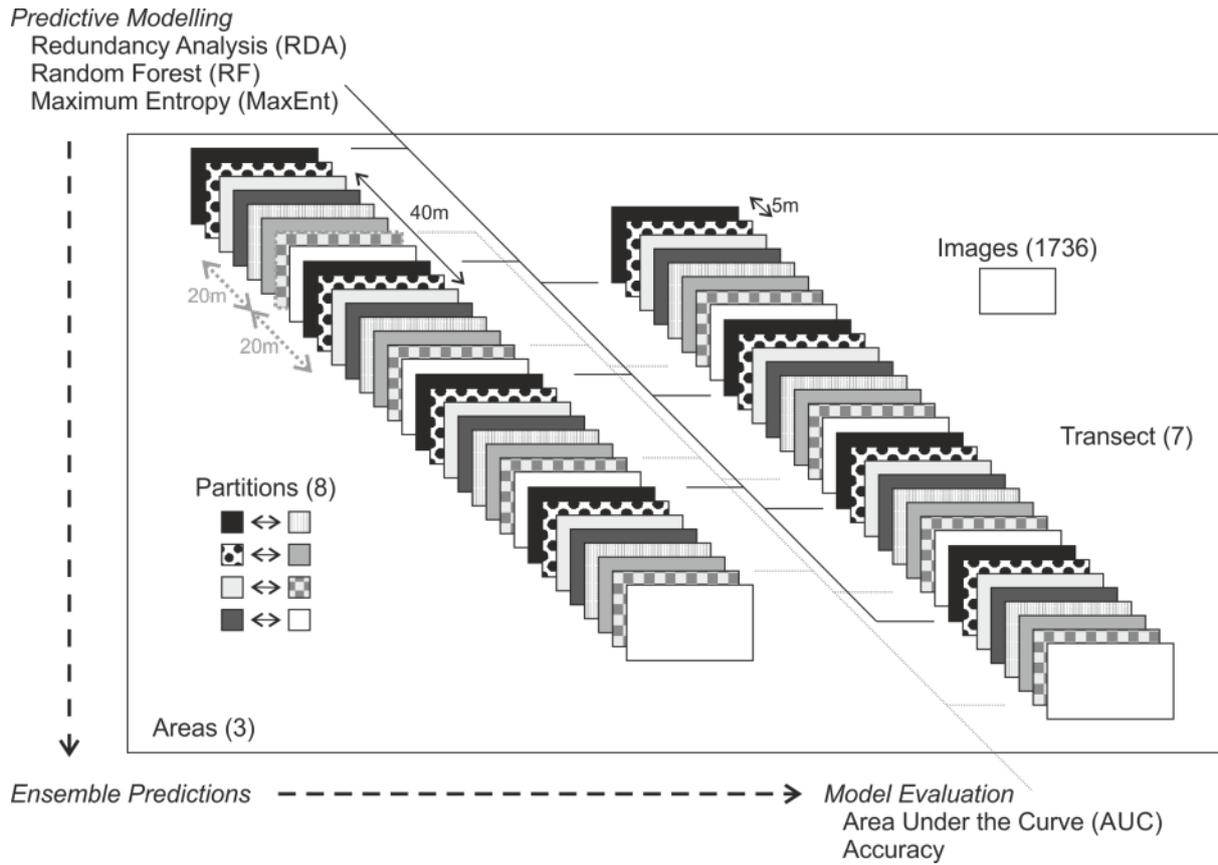
146 All individual organisms larger than 1 cm were counted and identified, using
147 morphospecies when species-level identification could not be achieved. Identification was
148 achieved by consulting image catalogues (Guillaumont et al., 2014; Howell and Davies,
149 2010; Jones and Gates, 2010; KeyToNature programme, 2015; SERPENT project, 2009;
150 WoRDSS, 2016), species lists compiled from conservation work in the area (Howell et al.,
151 2009; JNCC, 2010) and taxonomic resources (Hayward and Ryland, 1995 ; Mortensen,
152 1927). Sponges were only described to morphological categories as outlined in Bell and
153 Barnes (2001). Parallel lasers (with 10 cm separation) were mounted on the ROVs to provide
154 a scale on all recorded images. Positioning was achieved using the ROVs' ultra-short
155 baseline (USBL) navigation systems. Only common species, which occurred in at least 10
156 images, were retained for the analysis, which was carried out with the images as sampling
157 units.



158

159 **Fig. 1.** Map of the surveys carried out on Rockall Bank, Northeast Atlantic. Ship-based
 160 bathymetry (black outline) displayed with superimposed outlines of the sidescan sonar data
 161 (grey boxes) collected during three autonomous underwater vehicle missions. The remotely
 162 operated vehicle imagery transects carried out during the *JC-060* cruise are shown in red and
 163 the two from *JC-073* in blue, superimposed on the autonomous underwater vehicle sidescan
 164 sonar maps (high backscatter shown in white). The boundaries of a 2007 fisheries closure
 165 area (yellow) and a candidate for ‘Special Area of Conservation’ (pink) are also displayed.
 166 Depth contours for Northeast Atlantic background from GEBCO bathymetry (General
 167 Bathymetric Chart of the Oceans (IOC IHO and BODC, 2003).

168



169
 170

171 **Fig. 2.** Schematic of the steps taken. The data were separated into eight partitions where
 172 images were located every 40m. Three statistical approaches were applied separately to each
 173 partition and the results were evaluated using the partition whose images were halfway
 174 (20m). For each partition, the three statistical approaches were combined to form ensemble
 175 models. These steps were carried out for four species assemblages.

176

177

178 Environmental descriptors were derived from the sidescan backscatter maps
179 (EdgeTech FS2200, 410 kHz). These maps had been classified into sediment interpretation
180 maps (0.5x0.5 m pixel size) representing six seabed facies (soft and mixed sediments, hard
181 substratum, exposed bedrock as well as coral stand and rubble) using an unsupervised
182 classification (Robert et al., 2014). From the sediment interpretation maps, class and
183 landscape indices were derived to describe the shape, size, diversity and spatial arrangement
184 (connectivity) of habitat patches (selected for their explanatory power using redundancy
185 analysis and forward selection based on Robert et al. (2014), listed in Table 1). Bathymetry
186 and CTD derived environmental variables were examined, but as they did not significantly
187 improve the models, they were not included and are not discussed further.

188 Landscape and class metrics were calculated for each pixel of the sediment
189 interpretation maps using moving windows (sizes of 60x60 m and 150x150 m, see Robert et
190 al. (2014) for description of size choice). Owing to the large number of computations
191 involved, the high performance computer cluster IRIDIS 3 (University of Southampton) was
192 used to run an R script (R Development Core Team, 2011) written for parallel computation.
193 The R package 'SDMTools' was used to compute the metrics and the package 'Snowfall'
194 was used to run the computations in parallel. On smaller datasets, these computations could
195 easily be accomplished on a regular desktop computer (see Appendix A for R code).

196 **Table 1.** List of class and landscape variables used to construct predictive maps. Values in
 197 italics indicate the size of the moving window used to calculate the metrics. For formulas
 198 and descriptions see McGarigal (2012).
 199

200	Class Metrics	
201	<i>60m</i>	<i>150m</i>
202		Sand
203	Proportion of Like Adjacencies	Max. Shape Index
204	Landscape Shape Index	Min. Shape Index
205		Proportion of Like Adjacencies
206		Mean Patch Area
207		
208		Mixed Sediments
209	Landscape Shape Index	Mean Shape Index
210	Mean Patch Core Area	
211		
212		Hard Substrate
213	Mean Shape Index	Mean Patch Core Area
214		
215		Bedrock
216	Effective Mesh Size	Min. Patch Area
217	Patch Density	
218	Mean Shape Index	
219		Live Coral
220	Total Area	Mean Shape Index
221	Min. Patch Core Area	Largest Patch Index
222		
223		Coral Rubble
224	Patch Density	Min. Patch Core Area
225		
226	Landscape Metrics	
227	<i>60m</i>	<i>150m</i>
228	β Max. Shape Index	Number of Patches
229	β Largest Patch Index	Mean Core Area Index
230		Mean Shape Index
231		
232		

233 2.2 Predictive Modelling

234 Four species assemblages (A1- *Parastichopus tremulus*, A2- *Munida sarsi* and
235 associated species, A3- *Reteporella* sp. and various sponge spp., and A4- *Lophelia pertusa*
236 and associated species) were identified using K-mean classification, ANOSIM and ‘species
237 indicator values’ (described in Robert et al. (2014)). Although low numbers of individuals
238 were generally found, the holothurian *Parastichopus tremulus* was most commonly observed
239 in soft sediments. Bryozoan species (mostly *Reteporella* sp.) and various sponge
240 morphotypes were characteristic of hard substratum, while the abundant squat lobster *Munida*
241 *sarsi* dominated mixed sediment areas. Species associated with the cold-water coral *Lophelia*
242 *pertusa* included sabellid worms, an unsampled actinarian sp. and many asteroid spp.
243 Analyses were carried out using the R libraries ‘vegan’, ‘randomForest’, ‘dismo’, ‘raster’ and
244 ‘caret’.

245 2.2.1 Redundancy Analysis

246 A multivariate approach was first considered, and redundancy analysis (RDA) was
247 used to create the first set of full coverage fine scale biological maps. Similarly to Oldeland
248 et al. (2010), we used the estimated coefficients of the linear combination of environmental
249 predictors to position each pixel along the canonical axes. To assign each pixel to a species
250 assemblage, a nearest neighbour classification (k=10) was carried out. The probability of
251 belonging to each of the species assemblages was estimated using the proportion of nearest
252 neighbours.

253 2.2.2 MAXENT

254 Maximum entropy (MaxEnt) predicts an index of relative habitat suitability using
255 presence data compared to randomly selected background points (Phillips and Dudík, 2008)
256 by minimizing the distance between the probability density of species occurrence and the
257 probability density of the covariates as they occur in space (relative entropy) (Elith et al.,
258 2011). The software MaxEnt (version 3.3.3, freely available online

259 <http://www.cs.princeton.edu/~schapire/maxent/>) was employed with sampling bias grids to
260 select the background points and help account for the transect design. Weighted surfaces
261 based on sampling density were built (with more weight given to areas closer to sampled
262 locations) using a Gaussian kernel estimation (with SD of 500 m) (Clements et al., 2012).
263 Habitat suitability predictions were made separately for each of the four species assemblages.
264 Presence/absence predictions were obtained by setting the threshold level to optimize
265 sensitivity and specificity.

266 2.2.3 Random Forest

267 Random Forest (RF) is a technique that allows for the building of multiple trees for a
268 dataset, hence the term forest (Breiman, 2001). Each tree is built based on a sub-sample of
269 the data and at each node the data are split based on the best predictor variable, selected out
270 of a smaller number of randomly selected variables. A probability estimate can be obtained
271 based on the number of votes given to each class for a given pixel. Forests were built using a
272 varying number of trees and environmental variables, but a forest containing 1,000 trees and
273 considering 15 environmental predictors per node was selected.

274 2.3 Model Evaluation

275 To minimize spatial autocorrelation between the training and testing datasets,
276 systematic data splitting was carried out. For each of the 8 data partitions, the dataset whose
277 images were located at a distance of 20m (for example models based on partition 3 were
278 assessed using images in partition 7, Figure 2) were used to calculate the area under the curve
279 (AUC) of the receiver operating characteristics (ROC) (Fielding and Bell, 1997; Manel et al.,
280 2001). This distance was chosen based on spatial analysis of the data which indicated that
281 spatial autocorrelation became negligible at distances of ~20 m as a result of the high
282 heterogeneity introduced by iceberg ploughmarks in the area (Robert et al., 2014). AUC was
283 calculated for each partition, species assemblage and statistical approach. The test AUC

284 values reported for MaxEnt were based on absences identified within the acquired imagery as
285 opposed to background data. Prediction accuracy (the proportion of correctly assigned
286 presences and absences over total sample size) was also calculated. Full coverage maps were
287 produced for each of the eight partitions. To assess similarities between the predictive maps
288 obtained from each of the statistical approaches, the Hellinger-based niche similarity metric
289 described in Warren, Glor & Turelli (2008) was computed on maps averaged across
290 partitions. This measure can vary from 0 (no overlap) to 1 (identical niches).

291 **2.4 Ensemble Predictions**

292 Considering that different models are likely to produce different predictive outputs,
293 but with each containing separate information and areas of uncertainties, the idea of ensemble
294 predictions is to summarise a range of potential outcomes to produce more robust predictions
295 (Araújo and New, 2007). Using the same partitioning of training and test datasets as
296 previously described, for each partition, AUC values for the ensembles were calculated by
297 averaging probability maps from all three models for each species assemblage. Accuracy of
298 the ensemble predictions was calculated by first assigning, for each statistical approach and
299 partition, the species assemblage with the highest predicted probability of occurrence.
300 Subsequently, for each partition, majority voting was carried out based on the species
301 assemblage predicted by each statistical technique. To obtain a visual depiction of prediction
302 confidence, the number of models in agreement at each pixel was also calculated.

303 **3. RESULTS**

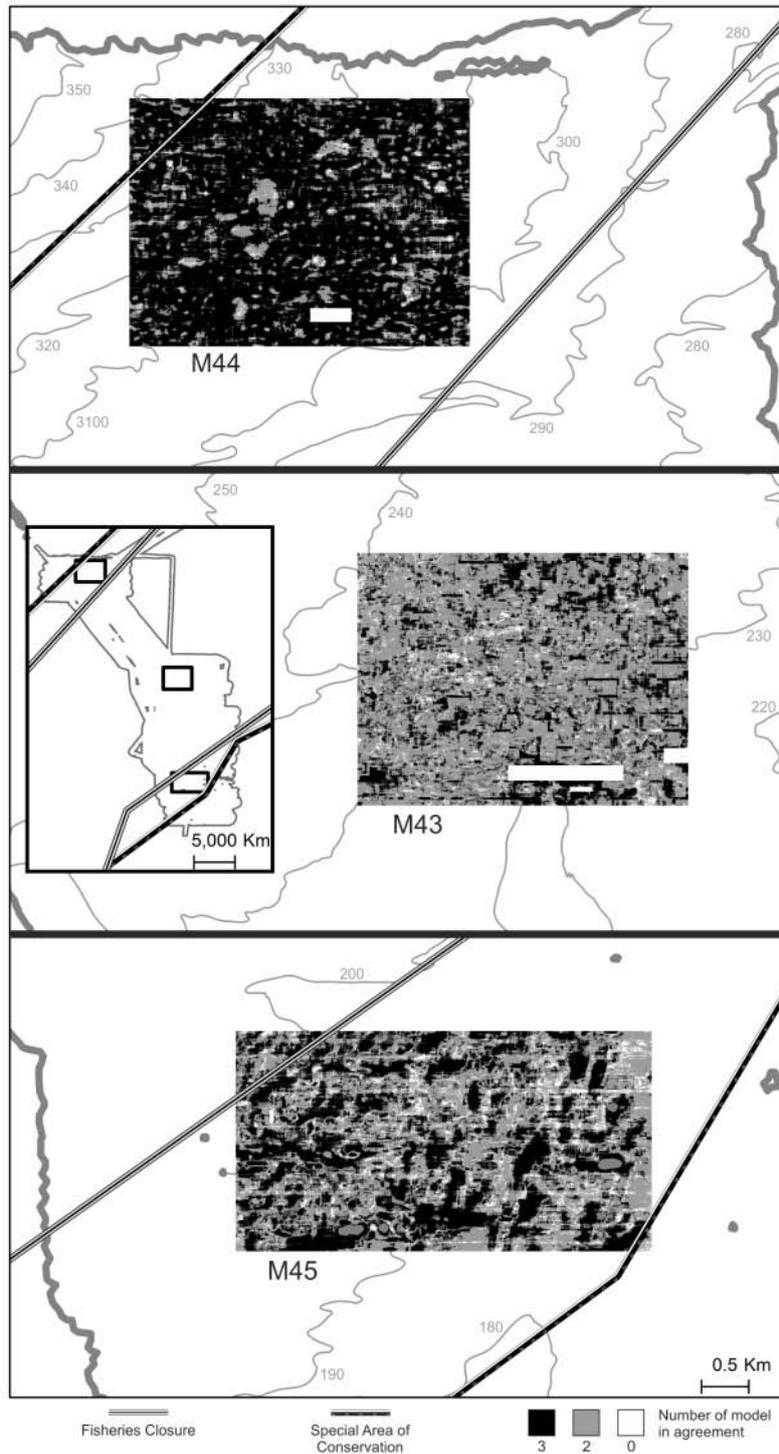
304 For the combined *JC-060* and *JC-073* datasets, a total of 11,268 individual organisms
305 were observed from 38 morphospecies (present in at least 10 images). For each of the
306 species assemblages considered, AUC values showed all models to perform better than could
307 be expected by chance (Table2). Based on the eight partitions MaxEnt showed average AUC
308 values ranging from 0.73 (SD=0.05 Assemblage A4; SD=0.02 Assemblage A3) to 0.63

309 (SD=0.05 Assemblage A2), values of 0.81 (SD=0.02 Assemblage A1) to 0.63 (SD=0.04
310 Assemblage A2) for RDA and 0.83 (SD=0.02 Assemblage A1) to 0.68 (SD= 0.05
311 Assemblage A2) for the RF classifier. All models had the most difficulties predicting
312 Assemblage A2. Assemblage A1 could be captured by RDA and RF, but predictions using
313 MaxEnt were lower (Mean=0.67, SD=0.02). Overall, RF had the highest AUC values across
314 species assemblages.

315 The three models showed differences in the maps of habitat suitability for the various
316 species assemblages, but measures of environmental niche indicated similarities between
317 model predictions (Table 3). Generally, RDA and RF showed the most similarities across all
318 species assemblages, and all models tended to select a similar set of environmental
319 descriptors as most important. For RF, the number of patches (150 m), the mean patch size
320 for soft sediments (150 m) and the total area occupied by coral (60 m) were the most
321 important variables. For individual species assemblages, all three environmental descriptors
322 were also the most important for Assemblage A1, while for Assemblage A3, it was the
323 former and for Assemblage A4, the latter. Proportion of like adjacencies for soft sediment
324 (150 m) was also important for Assemblage A2. For MaxEnt, total area covered by coral
325 (60m) and effective mesh size of bedrock (60 m) were also important for Assemblage A4,
326 while number of patches (150 m) was selected for both Assemblages A2 and A3. For
327 Assemblage A1, it was maximum shape index and proportion of like adjacencies for soft
328 sediment (150m). Similarly for RDA, the number of patches (150 m) and the proportion of
329 like adjacencies for soft sediment (150 m) were most important, but the mean shape index for
330 corals (150 m) and the effective mesh size for bedrock (60 m) were also valuable.

331 **Table 2.** Area under the curve (AUC) values observed based on eight partitions for four different species assemblages using three statistical
 332 approaches (Maximum Entropy (MaxEnt), Redundancy Analysis (RDA) and Random Forest (RF)) and one ensemble mapping approach.
 333 Numbers in bold show highest values obtained for each assemblage and partition. Assemblage A1 - *Parastichopus tremulus* and associated
 334 species, Assemblage A2 - *Munida sarsi* and associated species, Assemblage A3 - *Reteporella* sp. and various sponge spp. and Assemblage A4 -
 335 *Lophelia pertusa* and associated species.

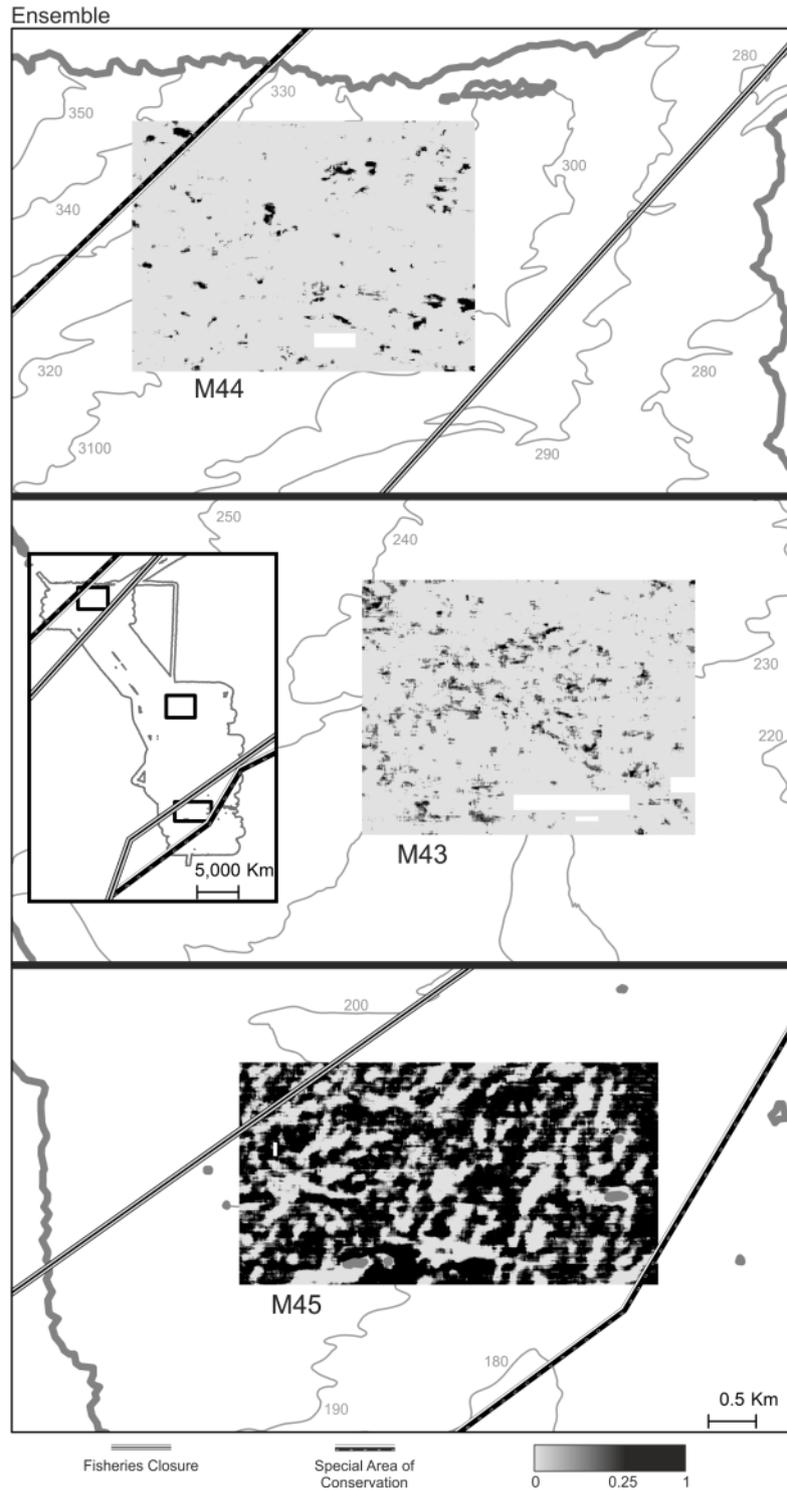
	Assemblage A1			Assemblage A2			Assemblage A3				Assemblage A4						
	RDA	RF	MaxEnt	EnsembleRDA	RF	MaxEnt	EnsembleRDA	RF	MaxEnt	EnsembleRDA	RF	MaxEnt	Ensemble	RF	MaxEnt	Ensemble	
339	1	0.84	0.83	0.63	0.83	0.69	0.69	0.68	0.74	0.70	0.81	0.71	0.75	0.76	0.82	0.81	0.83
340	2	0.80	0.83	0.66	0.82	0.63	0.76	0.65	0.72	0.74	0.81	0.74	0.78	0.59	0.67	0.71	0.69
341	3	0.82	0.86	0.66	0.85	0.55	0.67	0.63	0.64	0.75	0.81	0.71	0.76	0.78	0.81	0.76	0.80
342	4	0.79	0.84	0.66	0.82	0.64	0.66	0.61	0.67	0.69	0.73	0.72	0.72	0.56	0.67	0.65	0.67
343	5	0.83	0.79	0.72	0.84	0.67	0.69	0.67	0.72	0.77	0.80	0.75	0.80	0.76	0.78	0.73	0.78
344	6	0.81	0.84	0.69	0.84	0.60	0.73	0.65	0.69	0.84	0.85	0.74	0.83	0.81	0.76	0.70	0.77
345	7	0.83	0.83	0.65	0.84	0.63	0.63	0.63	0.66	0.79	0.80	0.75	0.79	0.72	0.78	0.74	0.77
346	8	0.80	0.83	0.67	0.83	0.64	0.62	0.55	0.59	0.68	0.75	0.74	0.74	0.68	0.74	0.74	0.79
347	Mean	0.81	0.83	0.67	0.83	0.63	0.68	0.63	0.68	0.74	0.79	0.73	0.77	0.71	0.75	0.73	0.76
348	SD	0.02	0.02	0.03	0.01	0.04	0.05	0.04	0.05	0.05	0.04	0.02	0.04	0.09	0.06	0.05	0.06



368

369 **Fig. 3.** Maps showing the prediction agreement based on the three models considered for
 370 Assemblage A4 (*Lophelia pertusa* and associated species) for three survey areas around two
 371 conservation zone boundaries. The maps were based on the habitat suitability averaged
 372 across all eight partitions. The inset on the left shows the relative position of the three survey
 373 areas M43, M44 and M45 with respect to the boundaries of the conservation zones and the
 374 outline of the shipboard-multibeam survey. The white rectangles represent areas for which
 375 the acoustic data was corrupted and were not included in the prediction models.

376



377

378 **Fig. 4.** Maps showing the relative habitat suitability resulting from an ensemble modelling
 379 approach for Assemblage A4 (*Lophelia pertusa* and associated species) for three survey areas
 380 around two conservation zone boundaries (higher suitability in black). The maps show the
 381 habitat suitability averaged across all eight partitions. The inset on the left shows the relative
 382 position of the three survey areas M43, M44 and M45 with respect to the boundaries of the
 383 conservation zones and the outline of the shipboard-multibeam survey. Maps for the other
 384 three groups are presented in Appendix B. The white rectangles represent areas for which the
 385 acoustic data was corrupted and were not included in the prediction models.

386 The areas of variability also differed between models (Figure 3) and ensemble
387 predictions (Figure 4 and Appendix B) made by combining all three models exhibited a
388 slightly higher accuracy across species assemblages than could be obtained based on any
389 single model (Table 4). Ensemble predictions showed the highest (or equal) accuracies in
390 every partition for Assemblage A3, and all but one partition for Assemblage A1. For the
391 other two species assemblages, ensemble predictions still showed the highest (or equal)
392 accuracy in half of the partitions. Overall, models completely disagree in less than 10% of
393 the area surveyed while all three agreed in 25.8% of M43, 79.0% of M44 and 42.1% of M45
394 (Figure 3).

395 **4. DISCUSSION**

396 By taking advantage of species-environment relationships, abiotic proxies can provide
397 direct applications for the management of natural resources by establishing representations of
398 biotic components via high resolution acoustic survey techniques. The spatial arrangement of
399 habitat patches was successfully included to predict the spatial patterns of four species
400 assemblages across a highly heterogeneous area of seabed. No single approach consistently
401 surpassed the others across species assemblages and although differences occurred between
402 spatial predictions of habitat suitability from the different statistical approaches, ensemble
403 models appeared as a meaningful improvement.

404 **4.1 Model Predictions**

405 Of the three models (RDA, RF and MaxEnt) compared in this study, similar AUC values
406 were obtained, but performance varied by species assemblage. As species turnover generally
407 occurs over a gradient, the predictions showed a similar pattern, and overlap between habitat
408 suitability predictions occurred, particularly between Assemblage A3 and A4. This is to be
409 expected as cold-water corals need hard substratum for attachment .

410 **Table 4.** Accuracy values obtained based on eight partitions for four different species assemblages using three statistical approaches (Maximum
 411 Entropy (MaxEnt), Redundancy Analysis (RDA) and Random Forest (RF)) and an ensemble mapping approach. Accuracy defined as the
 412 proportion of correctly assigned presences and absences over sample size. Number in bold show highest values obtained for each assemblage
 413 and partition. Assemblage A1 - *Parastichopus tremulus* and associated species, Assemblage A2 - *Munida sarsi* and associated species,
 414 Assemblage A3 - *Reteporella* sp. and various sponge spp. and Assemblage A4 - *Lophelia pertusa* and associated species.

415	Assemblage A1				Assemblage A2				Assemblage A3				Assemblage A4				
416	Partition	RDA	RF	MaxEnt	Ensemble	RDA	RF	MaxEnt	Ensemble	RDA	RF	MaxEnt	Ensemble	RDA	RF	MaxEnt	Ensemble
417	1	0.76	0.70	0.76	0.81	0.74	0.73	0.71	0.73	0.76	0.88	0.84	0.88	0.82	0.87	0.83	0.83
418	2	0.72	0.73	0.76	0.78	0.68	0.77	0.68	0.75	0.80	0.83	0.82	0.86	0.82	0.85	0.82	0.87
419	3	0.70	0.76	0.76	0.78	0.78	0.78	0.70	0.76	0.79	0.85	0.79	0.88	0.82	0.84	0.82	0.83
420	4	0.69	0.74	0.72	0.76	0.78	0.75	0.61	0.83	0.79	0.80	0.81	0.92	0.81	0.77	0.82	0.84
421	5	0.72	0.70	0.76	0.71	0.71	0.71	0.76	0.71	0.83	0.87	0.79	0.88	0.82	0.85	0.85	0.81
422	6	0.70	0.73	0.76	0.76	0.76	0.73	0.70	0.73	0.81	0.83	0.82	0.90	0.85	0.82	0.84	0.78
423	7	0.75	0.71	0.76	0.77	0.76	0.76	0.71	0.76	0.83	0.87	0.80	0.87	0.79	0.81	0.81	0.82
424	8	0.71	0.73	0.73	0.76	0.69	0.72	0.72	0.74	0.85	0.85	0.82	0.86	0.76	0.82	0.80	0.82
425	Mean	0.72	0.72	0.75	0.77	0.74	0.74	0.70	0.75	0.81	0.85	0.81	0.88	0.81	0.83	0.82	0.83
426	SD	0.02	0.02	0.02	0.03	0.04	0.02	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.03	0.02	0.02

427 (Wilson, 1979) and in turn provide hard substratum to a number of species. Assemblage A2
428 appeared as a transition between the more defined hard substratum and soft sediment
429 associated fauna, and as such prediction performance for this assemblage generally tended to
430 be lower. Across models, areas of highest disagreement tended to occur at the edge of patches
431 and highlighted the difficulty associated in delineating hard boundaries for otherwise
432 continuous gradients of species assemblages. Albeit at a larger scale, higher levels of
433 discrepancies between modelling techniques have been shown to occur at the edge of a
434 species distribution (Grenouillet et al., 2011). Assemblage A2 tended to be found in
435 particularly complex areas where a high number of patches, of both hard and soft sediments,
436 appeared. On the other hand, Assemblage A1 was found in areas characterised by few large
437 patches in proximity to other large soft sediment patches. Assemblage A3 or even A4 were
438 generally found in regions of harder substratum, particularly if bedrock was present.

439 As these three statistical approaches are based on very different modelling strategies,
440 differences in their predictions are to be expected. Presence-absence models generally
441 provide more information about less suitable habitats (if adequate absences are available). As
442 this information is not available to presence only models, overestimation of suitable habitats
443 can occur (Brotons et al. 2004; Pearson et al. 2006). Results can also depend on species
444 characteristics, with generalist species being more difficult to predict accurately, and absence
445 data being more valuable for such species (Brotons et al., 2004; Marmion et al., 2009a). This
446 might be another reason why lower prediction performances were obtained for Assemblage
447 A2. Overall MaxEnt tended to show a lower niche similarity than RDA and RF, which may
448 be due to its different data requirement. In the case of RDA, classification into assemblages
449 was only conducted after predictions of individual species, and as such could be more
450 affected by difficulties associated with predicting rarer species. However, since species are
451 predicted instead of assemblages, it might also be possible to define potentially new

452 assemblages as occurring in areas outside of the originally sampled locations (Ferrier and
453 Guisan, 2006). RF predictions for Assemblage A4 equalled those of the ensemble model.
454 Other studies have found RF to often equal ensemble approaches (Grenouillet et al., 2011;
455 Marmion et al., 2009b), potentially because it already includes a consensus step and might be
456 less affected by species geographical attributes, such as prevalence, range and spatial
457 autocorrelation (Marmion et al., 2009a). On the other hand, Meynard and Quinn (2007)
458 found that although GAM tended to outperform classification trees under many simulated
459 scenarios, the latter were particularly effective at predicting species displaying threshold
460 (on/off) response curves to environmental variables. In the case of Assemblages A3 and A4,
461 a threshold response to the presence of hard substratum could be expected while
462 Assemblages A1 and A2 may be more likely to exhibit more continuous response curves.
463 Additional factors also complicate predictive habitat modelling, particularly in deeper waters.
464 Precise spatial positioning can be problematic for underwater vehicles, particularly AUVs,
465 owing to difficulties associated with determining the initial position following the descent
466 and correcting for drift (McPhail, 2009). However, as our study site was located at 200-400
467 m water depth, limited drift would be expected. Accuracy of the ROV's Sonardyne USBL
468 navigation is also expected to be better than 1% of the depth. With the landscape and class
469 metrics calculated at scales of 60 m and 150 m, a small shift in position would have had
470 limited effects on the values of the explanatory variables. The spatial extent for which
471 predictions can be valid is also of importance, as predictions made for areas outside the range
472 of environmental conditions captured by the survey design are problematic (Elith and
473 Leathwick, 2009). The use of transects limited the area surveyed to single narrow lines
474 leaving most of the regions covered acoustically without any biological sampling. Transects
475 are designed to maximise seafloor survey areal coverage for a given bottom time, but also
476 have the disadvantage of causing issues of spatial autocorrelation which need to be taken into

477 account in order to adequately capture predictive ability (Hirzel and Guisan, 2002; Legendre
478 et al., 2002). In our study, this effect was mitigated through a subsampling scheme which
479 increased distances between sample images used for model building.

480 **4.2 Ensemble Mapping for Conservation**

481 Comparison of the statistical approaches showed differences in predictions, but a single
482 approach did not consistently outperform the others when multiple species assemblages were
483 considered. Instead, our results suggest that taking into account the output of many different
484 models may provide a valuable alternative. Ensembles can be created using an array of
485 approaches (Marmion et al., 2009b), but even the relatively simple approach taken in this
486 study was effective at optimizing different model strengths and increasing accuracy. All
487 three statistical approaches were included in the ensemble mapping of all four species
488 assemblages, but in other cases, the consideration of thresholds for the exclusion of lower
489 performing models may also be valuable. In any case, diversity in the type of approaches
490 selected is needed to increase the likelihood of obtaining better performing ensemble models
491 (Du et al., 2012). Identifying regions of prediction disagreement across models also provides
492 an easy to understand depiction of spatial uncertainties.

493 On the other hand, under certain circumstances, ensemble mapping may be of less
494 value. One example may be when the main aim is to derive habitat indicators such as extent
495 to monitor the achievement of "Good Environmental Status" as suggested by the Marine
496 Strategy Framework Directive (MSFD) (2008/56/EC) (CEFAS 2012; OSPAR 2012). In this
497 case, having the most thorough description of an habitat, including across model
498 uncertainties, may not be as important as having a consistent approach with minimum
499 deviation over time from which to monitor change (Strong, 2015). Employing multiple
500 models can increase the variability as some models may perform less adequately for certain
501 assemblages and make it more difficult to assess the degree of change across surveys.

502 However, this should still not preclude the examination of the data using multiple statistical
503 approaches, as one approach may be more sensitive to a given environmental variable and be
504 able to detect change earlier. Once prediction similarly has been ascertained, the final
505 measure of extent could still rely on one specific technique for consistency.

506 Cold-water corals can have a strong impact on local diversity and much effort is being
507 made to improve their conservation (Roberts and Hirshfield, 2004), but owing to limited data
508 spatial planning often must rely only on spatial predictions of habitat suitability. Even so,
509 these maps provide greater insights into their spatial distribution patterns, which helps in
510 understanding their ecology and supports adequate management better than single point
511 observation obtained from limited imagery transects or physical samples. As illustrated by
512 the case of Rockall Bank, different statistical approaches may provide different predictive
513 maps of coral suitability. Predictions of assemblage A4 (mostly composed of the cold-water
514 coral *L. pertusa* and associated filter-feeding species), the least common assemblage, were
515 particularly sensitive to changes in modelling approach. For example, if only random forest
516 had been considered, it would have been tempting to conclude that area M44 was as suitable
517 a conservation area as M45. However, M44 was only found to contain coral rubble in ROV
518 video surveys, likely resulting from past trawling activities. Ensemble models better
519 represented the spatial patterns observed in the video survey as they highlight areas where
520 predictions were consistent across at least two models.

521 Even though it is the broader-scale patterns in species distributions that may be of
522 interest for management purposes, it is the fine-scale habitat characterisation of the
523 environment, through high-resolution sidescan sonar mapping, that allowed the heterogeneity
524 of the region to be accurately captured and the driving processes identified. The ship-board
525 bathymetry survey carried out during *JC-60* covered less than 10% of the 4,365 km²
526 conservation zone and took approximately 2.3 days. Although of much higher resolution

527 than other datasets available for the remainder of this area, compared to the even higher
528 resolutions obtained with the AUV, the ship-board dataset was of limited use in explaining
529 species distribution patterns for the extent covered in this survey (Robert et al. 2014). It is
530 clear that AUV mapping shows great promise for marine management; however there
531 remains a distinct trade-off between the resolution achieved and the extent that can be
532 covered. With current AUV technologies, Autosub6000 can be sent out from a ship to
533 autonomously map an area for ~30 hrs, covering a distance of ~150 km (the size of the
534 resulting area mapped will vary based on the acquired resolution) (Wynn et al., 2012). In
535 order to map the entirety of the conservation zone to the resolution acquired in this study,
536 >200 days would be required. This is well outside the scope of most scientific cruises or
537 conservation projects, but AUVs have been successfully employed to target certain features
538 in other conservations zones such as Haig Fras and the Darwin Mounds (Wynn et al., 2012).
539 The Marine Autonomous and Robotic Systems (MARS) facility is also currently working on
540 developing long-distance AUVs which could be deployed from shore to reach the closer
541 offshore conservation areas with the aims of eventually covering greater extents at high
542 resolutions and instituting repeat long-term monitoring of specific areas without the need for
543 expensive ship-based surveys.

544 **4.3 Conclusion**

545 Predictive habitat maps are of great use for marine management as they represent the
546 best available information to support decision making, but, as they are typically based on a
547 very limited amount of data, they should only serve as general guides until more data become
548 available. The presentation of uncertainty maps should help emphasize this point and can be
549 employed to help select target areas for which further biological sampling will be particularly
550 valuable. Uncritical reliance on a particular statistical method, without comparison with
551 others, may lead to decisions being biased by the chosen method since predictions made from

552 different modelling strategies have been shown to give differing outputs, but whose
553 combination into ensemble models can lead to increased accuracy. Comparison between
554 statistical methods showing one method to outperform the others may not always be
555 extendable to other habitats, species or assemblages, and similarly our results cannot be
556 perfunctorily generalized to all habitats. However, in cases where one statistical approach
557 cannot be identified as performing significantly better, ensemble approaches may provide an
558 elegant alternative. Although this approach can be more involved than other techniques, the
559 additional work requires no further costly sampling or access to specialized equipment and
560 potential increases in prediction performances are clearly of value for spatial planning.

561

562 **ACKNOWLEDGEMENTS**

563 We would like to thank the captain, crew, technicians and scientific party of cruises
564 *JC-060* and *JC-073*. We would also like to acknowledge the following funding sources:
565 MAREMAP (NERC), HERMIONE (EU FP7 project, Contract number 226354), JNCC, the
566 Lenfest Ocean Programme (PEW Foundation), CODEMAP(ERC Starting Grant no. 258482)
567 and the UK Ocean Acidification programme (NERC grant NE/H017305/1 to J.M.R.). K.R. is
568 supported by funding from CODEMAP and a Postgraduate Scholarship (PGSD3-408364-
569 2011) from NSERC-CRSNG. We would also like to acknowledge the use of the IRIDIS
570 HPC Facility, and associated support services at the University of Southampton. Finally, a
571 special thanks to Prof.Paul Tyler for helpful discussions and support throughout the PhD.

572 **LITERATURE CITED**

- 573 Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. *Trends in*
574 *Ecology & Evolution* 22 (1), 42-47.
- 575 Bell, J.J., Barnes, D.K.A., 2001. Sponge morphological diversity: a qualitative predictor of
576 species diversity? *Aquatic Conservation: Marine and Freshwater Ecosystems* 11 (2), 109-121.
- 577 Breiman, L., 2001. Random forests. *Machine Learning* 45 (1), 5-32.
- 578 Brotons, L., Thuiller, W., Araújo, M.B., Hirzel, A.H., 2004. Presence-absence versus
579 presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27 (4),
580 437-448.
- 581 Brown, C.J., Smith, S.J., Lawton, P., Anderson, J.T., 2011. Benthic habitat mapping: A
582 review of progress towards improved understanding of the spatial ecology of the seafloor
583 using acoustic techniques. *Estuarine, Coastal and Shelf Science* 92 (3), 502-520.
- 584 Buhl-Mortensen, L., Mortensen, P.B., Dolan, M., Dannheim, J., Bellec, V., Holte, B., 2012.
585 Habitat complexity and bottom fauna composition at different scales on the continental shelf
586 and slope of northern Norway. *Hydrobiologia* 685 (1), 191-219.
- 587 Centre for Environment Fisheries and Aquaculture Science (CEFAS), 2012. Proposed UK
588 targets for achieving GES and cost-benefit analysis for the MSFD.
- 589 Clements, G.R., Rayan, D.M., Aziz, S.A., Kawanishi, K., Traeholt, C., Magintan, D., Yazi,
590 M.F.A., Tingley, R., 2012. Predicting the distribution of the Asian tapir in Peninsular
591 Malaysia using maximum entropy modeling. *Integrative Zoology* 7 (4), 400-406.
- 592 Compton, T.J., Bowden, D.A., Pitcher, C.R., Hewitt, J.E., Ellis, N., 2013. Biophysical
593 patterns in benthic assemblage composition across contrasting continental margins off New
594 Zealand. *Journal of Biogeography* 40 (1), 75-89.
- 595 Du, P., Xia, J., Zhang, W., Tan, K., Liu, Y., Liu, S., 2012. Multiple classifier system for
596 remote sensing image classification: A review. *Sensors* 12, 4764-4792.
- 597 Elith, J., Leathwick, J.R., 2009. Species distribution models: Ecological explanation and
598 prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40
599 (1), 677-697.
- 600 Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical
601 explanation of MaxEnt for ecologists. *Diversity and Distributions* 17 (1), 43-57.
- 602 Elvenes, S., Dolan, M.F.J., Buhl-Mortensen, P., Bellec, V.K., 2014. An evaluation of
603 compiled single-beam bathymetry data as a basis for regional sediment and biotope mapping.
604 *ICES Journal of Marine Science* 71, 867-881.
- 605 Ferrier, S., Guisan, A., 2006. Spatial modelling of biodiversity at the community level.
606 *Journal of Applied Ecology* 43 (3), 393-404.
- 607 Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors
608 in conservation presence/absence models. *Environmental Conservation* 24 (1), 38-49.

- 609 Gonzalez-Mirelis, G., Lindegarth, M., 2012. Predicting the distribution of out-of-reach
610 biotopes with decision trees in a Swedish marine protected area. *Ecological Applications* 22
611 (8), 2248-2264.
- 612 Grenouillet, G., Buisson, L., Casajus, N., Lek, S., 2011. Ensemble modelling of species
613 distribution: the effects of geographical and environmental ranges. *Ecography* 34 (1), 9-17.
- 614 Guillaumont, B., van den Beld, I., Davies, J., Menot, L., 2014. Catalogue des coraux d'eau
615 froide et éponges du Golfe de Gascogne. IFREMER.
- 616 Hayward, P.J., Ryland, J.S., 1995 *The Marine Fauna of the British Isles and North-West*
617 *Europe*. Oxford University Press.
- 618 Hirzel, A., Guisan, A., 2002. Which is the optimal sampling strategy for habitat suitability
619 modelling. *Ecological Modelling* 157 (2-3), 331-341.
- 620 Howell, K.L., Davies, J.S., 2010. Deep-sea species image catalogue. [www.marlin.ac.uk/deep-](http://www.marlin.ac.uk/deep-sea-species-image-catalogue/)
621 [sea-species-image-catalogue/](http://www.marlin.ac.uk/deep-sea-species-image-catalogue/).
- 622 Howell, K.L., Davies, J.S., Jacobs, C., Narayanaswamy, B.E., 2009. Broadscale survey of the
623 habitats of Rockall Bank and mapping of Annex I 'reef' habitat. Joint Nature Conservation
624 Committee, Peterborough, UK.
- 625 IOC IHO and BODC, 2003. GEBCO digital atlas. Centenary edition of the GEBCO digital
626 atlas, published on CD-ROM on behalf of the Intergovernmental Oceanographic Commission
627 and the International Hydrographic Organization as part of the General Bathymetric Chart of
628 the Oceans, British Oceanographic Data Centre, Liverpool, U.K.
- 629 JNCC, 2010. Conservation objectives and advice on operations. Offshore special area of
630 conservation: North West Rockall Bank. Joint Nature Conservation Committee,
631 Peterborough, UK.
- 632 Jones, D.O.B., Gates, A.R., 2010. *Deep-sea life of Scotland and Norway*. Ophiura Publishing.
- 633 KeyToNature programme, 2015. Marine species identification portal. [http://species-](http://species-identification.org/index.php)
634 [identification.org/index.php](http://species-identification.org/index.php).
- 635 Kostylev, V.E., Todd, B.J., Fader, G.B.J., Courtney, R.C., Cameron, G.D.M., Pickrill, R.A.,
636 2001. Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial
637 geology and sea floor photographs. *Marine Ecology Progress Series* 219, 121-137.
- 638 Legendre, P., Dale, M.R.T., Fortin, M.-J., Gurevitch, J., Hohn, M., Myers, D., 2002. The
639 consequences of spatial structure for the design and analysis of ecological field surveys.
640 *Ecography* 25 (5), 601-615.
- 641 Manel, S., Williams, H.C., Ormerod, S.J., 2001. Evaluating presence-absence models in
642 ecology: The need to account for prevalence. *Journal of Applied Ecology* 38 (5), 921-931.
- 643 Marmion, M., Luoto, M., Heikkinen, R.K., Thuiller, W., 2009a. The performance of state-of-
644 the-art modelling techniques depends on geographical distribution of species. *Ecological*
645 *Modelling* 220 (24), 3512-3520.

- 646 Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W., 2009b. Evaluation
647 of consensus methods in predictive species distribution modelling. *Diversity and*
648 *Distributions* 15 (1), 59-69.
- 649 McGarigal, K., Cushman, S.A., Ene, E., 2012. FRAGSTATS (4.x): Spatial pattern analysis
650 program for categorical and continuous maps. University of Massachusetts, Amherst, USA.
- 651 McPhail, S., 2009. Autosub6000: A deep diving long range AUV. *Journal of Bionic*
652 *Engineering* 6 (1), 55-62.
- 653 Meynard, C.N., Quinn, J.F., 2007. Predicting species distributions: a critical comparison of
654 the most common statistical models using artificial species. *Journal of Biogeography* 34,
655 1455-1469.
- 656 Mortensen, T.H., 1927. , 1927. *Handbook of the Echinoderms of the British Isles*. Humphrey
657 Milford, Oxford University Press.
- 658 Oldeland, J., Dorigo, W., Lieckfeld, L., Lucieer, A., Jürgens, N., 2010. Combining vegetation
659 indices, constrained ordination and fuzzy classification for mapping semi-natural vegetation
660 units from hyperspectral imagery. *Remote Sensing of Environment* 114 (6), 1155-1166.
- 661 OSPAR Intersessional Correspondence Group on the Coordination of Biodiversity
662 Assessment and Monitoring, 2012. MSFD advice manual and background document on
663 biodiversity. In: OSPAR Biodiversity Committee (Ed.).
- 664 Palialexis, A., Georgakarakos, S., Karakassis, I., Lika, K., Valavanis, V.D., 2011. Prediction
665 of marine species distribution from presence-absence acoustic data: Comparing the fitting
666 efficiency and the predictive capacity of conventional and novel distribution models.
667 *Hydrobiologia* 670 (1), 241-266.
- 668 Pearce, J.L., Boyce, M.S., 2006. Modelling distribution and abundance with presence-only
669 data. *Journal of Applied Ecology* 43 (3), 405-412.
- 670 Pearson, R.G., Thuiller, W., Araújo, M.B., Martinez-Meyer, E., Brotons, L., McClean, C.,
671 Miles, L., Segurado, P., Dawson, T.P., Lees, D.C., 2006. Model-based uncertainty in species
672 range prediction. *Journal of Biogeography* 33 (10), 1704-1711.
- 673 Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: New
674 extensions and a comprehensive evaluation. *Ecography* 31 (2), 161-175.
- 675 Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J., Ferrier, S.,
676 2009. Sample selection bias and presence-only distribution models: Implications for
677 background and pseudo-absence data. *Ecological Applications* 19 (1), 181-197.
- 678 Przeslawski, R., Currie, D.R., Sorokin, S.J., Ward, T.M., Althaus, F., Williams, A., 2011.
679 Utility of a spatial habitat classification system as a surrogate of marine benthic community
680 structure for the Australian margin. *ICES Journal of Marine Science* 68 (9), 1954-1962.
- 681 R Development Core Team, 2011. R: A language and environment for statistical computing
682 R Foundation for Statistical Computing.

- 683 Rengstorf, A.M., Grehan, A., Yesson, C., Brown, C., 2012. Towards high-resolution habitat
684 suitability modeling of vulnerable marine ecosystems in the deep-sea: Resolving terrain
685 attribute dependencies. *Marine Geodesy* 35 (4), 343-361.
- 686 Robert, K., Jones, D.O.B., Huvenne, V.A.I., 2014. Megafaunal distribution and biodiversity
687 in a heterogeneous landscape: The iceberg scoured Rockall Bank, NE Atlantic. *Marine*
688 *Ecology Progress Series* 501, 67-88.
- 689 Robert, K., Jones, D.O.B., Tyler, P.A., Van Rooij, D., Huvenne, V.A.I., 2015. Finding the
690 hotspots within a biodiversity hotspot: fine-scale biological predictions within a submarine
691 canyon using high-resolution acoustic mapping techniques. *Marine Ecology*, DOI:
692 10.1111/maec.12228.
- 693 Roberts, S., Hirshfield, M., 2004. Deep-sea corals: Out of sight, but no longer out of mind.
694 *Frontiers in Ecology and the Environment* 2, 123-130.
- 695 Ross, R.E., Howell, K.L., 2012. Use of predictive habitat modelling to assess the distribution
696 and extent of the current protection of 'listed' deep-sea habitats. *Diversity and Distributions*
697 19, 433-445.
- 698 Sandwell, D.T., Smith, W.H.F., Gille, S., Kappel, E., Jayne, S., Soofi, K., Coakley, B., Géli,
699 L., 2006. Bathymetry from space: Rationale and requirements for a new, high-resolution
700 altimetric mission. *Comptes Rendus Geoscience* 338 (14-15), 1049-1062.
- 701 SERPENT project, 2009. Scientific and Environmental ROV Partnership using Existing
702 Industrial Technology (SERPENT): Media database archive. Jones, D.O.B., Gates, A.R.,
703 Curry, R.A., Thomson, M., Pile, A., Benfield, M. (Eds.) <http://archive.serpentproject.com>.
- 704 Shumchenia, E.J., King, J.W., 2010. Comparison of methods for integrating biological and
705 physical data for marine habitat mapping and classification. *Continental Shelf Research* 30
706 (16), 1717-1729.
- 707 Strong, J.A., 2015. Habitat area as an indicator of Good Environmental Status under the
708 Marine Strategy Framework Directive: the identification of suitable habitat mapping methods
709 with recommendations on best-practice for the reduction of uncertainty. In: Defra Project
710 Code: ME5318 (Ed.).
- 711 Teixidó, N., Garrabou, J., Arntz, W.E., 2002. Spatial pattern quantification of Antarctic
712 benthic communities using landscape indices. *Marine Ecology Progress Series* 242, 1-14.
- 713 ter Braak, C.J.F., 1994. Canonical community ordination. Part I: Basic theory and linear
714 methods *Ecoscience* 1, 127-140.
- 715 Turner, M.G., Gardner, R.H., 1991. Quantitative methods in landscape ecology: The analysis
716 and interpretation of landscape heterogeneity. Springer, USA.
- 717 Warren, D.L., Glor, R.E., Turelli, M., 2008. Environmental niche equivalency versus
718 conservatism: Quantitative approaches to niche evolution. *Evolution* 62 (11), 2868-2883.
- 719 Wilson, J.B., 1979. 'Patch' development of the deep-water coral *Lophelia pertusa* on Rockall
720 Bank. *Journal of the Marine Biological Association of the United Kingdom* 59 (1), 165-177.

- 721 WoRDSS, 2016. World Register of Deep-Sea species. Glover, A.G., Higgs, N., Horton, T.
722 (Eds.) <http://www.marinespecies.org/deepsea/photogallery.php>.
- 723 Wynn, R.B., Bett, B.J., Evans, A.J., Griffiths, G., Huvenne, V.A.I., Jones, A.R., Palmer,
724 M.R., Dove, D., Howe, J.A., Boyd, T.J., MAREMAP partners, 2012. Investigating the
725 feasibility of utilizing AUV and Glider technology for mapping and monitoring of the UK
726 MPA network. In: MB0118., D.p. (Ed.). National Oceanography Centre, Southampton, p.
727 244.
- 728