When the species is also a habitat: comparing the predictively modelled distributions of *Lophelia pertusa* and the reef habitat it forms.

Kerry L. Howell^a, Rebecca Holt^a, Inés Pulido Endrino^a, Heather Stewart^b

a. Marine Biology and Ecology Research Centre, Marine Institute at the University of Plymouth, Drake Circus, Plymouth, PL3 5EE. UK. Email: kerry.howell@plymouth.ac.uk

b. British Geological Survey, Murchison House, West Mains Road, Edinburgh, EH9 3LA. UK. Email: hast@bgs.ac.uk

Corresponding authors details: Email: kerry.howell@plymouth.ac.uk Tel: +44 (0)1752 584544

Abstract

Internationally there is political momentum to establish networks of marine protected areas for the conservation of threatened species and habitats. Practical implementation of such networks requires an understanding of the distribution of these species and habitats. Predictive modelling provides a method by which continuous distribution maps can be produced from limited sample data. This method is particularly useful in the deep sea where a number of biological communities have been identified as vulnerable ‘habitats’, including *Lophelia pertusa* reefs. Recent modelling efforts have focused on predicting the distribution of this species. However the species is widely distributed where as reef habitat is not. This study uses Maxent predictive modelling to investigate whether the distribution of the species acts as a

suitable proxy for the reef habitat. Models of both species and habitat distribution across Hatton Bank and George Bligh Bank are constructed using multibeam bathymetry, interpreted substrate and geomorphology layers, and derived layers of bathymetric position index (BPI), rugosity, slope and aspect. Species and reef presence records were obtained from video observations. For both models performance is fair to excellent assessed using AUC and additional threshold dependant metrics. 7.17% of the study area is predicted as highly suitable for the species presence while only 0.56% is suitable for reef presence, using the sensitivity-specificity sum maximization approach to determine the appropriate threshold. Substrate is the most important variable in the both models followed by geomorphology in the RD model and fine scale BPI in the SD model. The difference in the distributions of reef and species suggest that mapping efforts should focus on the habitat rather than the species at fine (100m) scales.

Keywords: deep-sea; *Lophelia pertusa*; habitat mapping; biotopes; marine protected area; maxent;

1. Introduction

The call for better spatial management of our marine environment is growing globally. Specifically, there is momentum for the establishment of networks of marine protected areas (MPAs) driven by global, European and (within the UK) national

initiatives. One of the criteria by which MPAs are selected includes the protection of habitats and species that have been identified as rare, sensitive, functionally important, threatened and / or declining.

Within the NE Atlantic region the 1992 Convention for the Protection of the Marine Environment of the north-east Atlantic (OSPAR Convention) gives the OSPAR Commission a duty to develop means, consistent with international law, for establishing protective, conservation, restorative or precautionary measures related to specific species or habitats. A target date of 2010 has been set by OSPAR contracting parties to achieve “an ecologically coherent network of well managed Marine Protected Areas” that serve (at least in part) to protect those habitats and species listed under Annex V of the Convention, on the OSPAR List of Threatened and/or Declining Species and Habitats.

At a European level the EU Habitats and Species Directive (92/43/EEC) requires the establishment of protected areas (Special Areas of Conservation – SACs) for habitats and species listed under Annex I and V respectively of the Directive, in areas of sea under the jurisdiction of member states (i.e. out to the 200 nm limit). In addition, the 2006 United Nations General Assembly Resolution 61/105 called “upon States to take action immediately, individually and through regional fisheries management organizations and arrangements, and consistent with the precautionary approach and ecosystem approaches, to sustainably manage fish stocks and protect vulnerable

marine ecosystems [VMEs], including seamounts, hydrothermal vents and cold water corals, from destructive fishing practices". This resolution has also ultimately resulted in the establishment of MPAs for the protection of specific species and habitats.

In order to establish MPAs for the protection of listed habitats and species there is a clear need to have a firm understanding of the distribution of those species and habitats (i.e. maps). The difficulties and expense of collecting species and habitat distribution data has led to the approach of using surrogates (Howell, 2010) and / or predictive species modelling techniques to provide maps of the distribution of vulnerable species (Bryan & Metaxas 2007; Holmes et al., 2007; Embling et al., 2010). This approach is particularly useful for the deep-sea and high seas. Here, the vast area involved, sparse and highly localised data available, and distance from land, compound the problems encountered in shallow water settings. Within the deep-sea (high seas) ecosystem there are few actual species that are listed as of conservation concern under the legislation detailed above. Only commercially important fish species including orange roughy (*Hoplostethus atlanticus*), Portuguese dogfish (*Centroscymnus coelolepis*), and Leafscale gulper shark (*Centrophorus squamosus*), known to have undergone significant declines (ICES, 2008; ICES, 2010) are included. However a number of deep-sea habitats are listed. These habitats are predominantly biogenic in origin or are in fact biological assemblages, and include *Lophelia pertusa* reefs, coral gardens, sponge aggregations and sea-pen and burrowing megafauna communities, as

well as other geogenic habitats such as carbonate mounds, seamounts, and oceanic ridges with hydrothermal vents/fields.

Recently, efforts have been made to model the distribution of the cold water coral *L. pertusa* from global to local scales, in order to identify areas of conservation importance (Davies et al., 2008; Dolan et al., 2008; Guinan et al., 2009; Tittensor et al., 2009). *L. pertusa* is a widely distributed species and occurs as isolated colonies on boulders, cobbles, sand ripples, and even flat sea bed where some form of hard substrate is available for attachment (Wilson, 1979; Mortensen and Buhl-Mortensen, 2004a, b; Hovland et al., 2005). Its conservation importance stems from its reef forming capacity. *L. pertusa* can form large reefs and giant carbonate mounds up to 300m high and several km in diameter (Roberts et al., 2006). Reef structures are highly biodiverse, possibly rivalling tropical coral reefs (Roberts et al., 2006). They may also have an important role as essential fish habitat but this is not yet clear (Husebø et al., 2002; Auster, 2005; Costello et al., 2005). *L. pertusa* only forms reefs under specific environmental conditions that are not yet fully understood but are controlled by the interplay between local hydrography and sedimentary dynamics (Thiem et al., 2006).

Given that the species is widely spread while the reef habitat has specific environmental requirements likely to result in a more confined distribution, to what extent does the distribution of *L. pertusa* species act as a proxy for the reef habitat? Mapping efforts that focus on the species rather than the habitat may produce maps

of limited use to marine environmental managers if the distribution of the species is so broad as to indicate reef habitat is widely spread. Using the species distribution as a proxy for the habitat could provide a false impression of the extent of reef habitat and effect assessments of rarity and threat from human activities. The aim of this study is to use predictive modelling to investigate the difference in the distribution of the species and habitat and assess the implications of any difference to marine environmental management. Multibeam bathymetry, interpreted substrate and geomorphology layers, and derived layers of bathymetric position index (BPI), rugosity, slope and aspect are used as environmental input layers together with presence of *L. pertusa* reef and species. Here we conform to the definition of *L. pertusa* reef following Roberts et al (2006) as biogenic structures formed by *L. pertusa* frameworks that alter sediment deposition, provide complex structural habitat and are subject to the processes of growth and (bio)erosion. However, presence data are derived from records of living *L. pertusa* reef only and not dead framework structures. The study focuses on Hatton and George Bligh Banks in the N.E. Atlantic.

2. Methods

2.1 Site description

Hatton Bank and George Bligh Bank are part of the Rockall Plateau (Hitchen, 2004), which is a large piece of continental crust that separated from the European continent during the early Cretaceous (Hauterivian-Cenomanian) when the North Atlantic was in the early stages of formation (Musgrove and Mitchener, 1996) (Fig. 1). Hatton Bank

forms an elongate arc that stretches over 400 km and descends >2500 m below sea level into the Iceland Basin to the west and 1100 m below sea level into the Hatton Basin, (sometimes referred to as the Hatton-Rockall Basin), to the east. At its summit Hatton Bank lies less than 500 m below sea level. South of 59°N Hatton Bank is orientated approximately southwest-northeast, further north the orientation is more east-west. George Bligh Bank is broadly conical in shape and situated at the north-eastern end of the Rockall Plateau (and Hatton Basin). It rises to a summit at 450 m below sea level, and has a diameter of roughly 75 km.

2.2 Data collection

Collection of biological (video) data and low resolution multibeam data from both Hatton Bank and George Bligh Bank were undertaken over a one month period (August-September) in 2005 using the commercial research vessel *S/V Kommandor Jack*. Further collection of biological data and collection of higher resolution multibeam data were undertaken over a two month period (August – October) in 2006 using the commercial research vessel *M/V Franklin*. Video sampling stations were selected during operations using multibeam bathymetry and backscatter data. Video tows were selected to cover a range of geomorphology, substrate types and water depths (Fig. 1).

Video data were collected using the Seatronics drop frame camera system. The system comprised an integrated DTS 6000 digital video telemetry system, which provided a

151 real time video link to the surface, and a digital stills camera (5 mega pixel, Kongsberg
152 OE14-208). In the 2005 surveys, the video stream from the viewing screen of the
153 digital stills camera provided video data, in 2006 separate video (Kongsberg 14-366)
154 and stills cameras were used. Cameras were mounted at an oblique angle (video: 24°;
155 stills: 22° from the horizontal) to the sea bed to aid in species and habitat
156 identification. Sensors monitored depth and altitude, and an Ultra Short Base Line
157 (USBL) beacon provided accurate (to approximately 1m) position data for the camera
158 frame.

159

160 The system was deployed from the starboard side of the vessel. Video tows were
161 between 250 and 1200 m long. For the majority of tows, vessel speed was
162 approximately 0.5 knots (min 0.3 and max 0.7 knots), with most tows lasting between
163 0.5 and 1.5 h. The drop frame was towed in the water column between one and three
164 metres (dependant on substrate type, topography and currents) above the sea bed. At
165 the beginning of each tow, starting from when the sea floor became visible, a 2-3
166 minute period was allowed before sampling, to enable the camera to stabilise before
167 commencing the transect.

168

169 Videos were reviewed and the occurrence of *L. pertusa* colonies and *L. pertusa* reef
170 habitat noted and linked to the navigational data from the USBL on the camera system,
171 such that the location of each colony / habitat occurrence was recorded. Presence
172 data for both species and habitat were then plotted in ArcGIS 9.3.

173

174 **2.3 Multibeam**

175 Multibeam data were acquired in 2005 on *S/V Kommander Jack* using an EM120 (12
176 kHz; 191 beams) multibeam echosounder system which achieved good quality
177 bathymetric data with marginal quality backscatter data. During the 2005 survey
178 complete data coverage was achieved through 2500m line spacing on Hatton Bank and
179 1500m line spacing on the other survey areas. In 2006 multibeam data were acquired
180 on *M/V Franklin* using an EM1002 (95 kHz; 111 beams) multibeam echosounder
181 system which achieved excellent quality bathymetric and moderate quality backscatter
182 data. During the 2006 survey complete data coverage was achieved through 650m line
183 spacing on Hatton Bank and 500m line spacing on the other survey areas. It should be
184 noted that weather and sea conditions adversely impacted on the backscatter data
185 quality during both the 2005 and 2006 surveys. Positioning was accomplished using
186 real-time Differential GPS (DGPS) systems. The C-Nav system was used in 2005 and the
187 ARON 2000 system in 2006. All data acquisition systems took their time stamp from
188 the primary DGPS which had a theoretical accuracy of better than 0.5m. Sound velocity
189 measurements were performed at regular intervals to account for hydrology effects
190 during both surveys. Multibeam data were processed onboard ship and ashore by
191 OSAE Ltd in 2005 and Marin Mätteknik AB in 2006. Data were gridded at resolutions
192 appropriate to the quality of the data (2005: 200m grids; 2006: 25m grids). Minimum
193 waters depths encountered over Hatton Bank were -483m, Lyonesse Seamount -507m
194 and George Bligh Bank -435m. The maximum water depth of -1679m was encountered

on the eastern flank of George Bligh Bank as it descends into the Rockall Trough. As Maxent requires all environmental data to have the same geographic bounds and cell size the 2006 multibeam dataset was regridded in ArcGIS 9.3 to a 200m grid and merged with the 2005 multibeam dataset to produce a single bathymetry data layer, which was used to produce subsequent derived layers (see section 2.4).

2.4 Derived layers

Additional environmental layers included in the model were layers derived from the multibeam bathymetry and were generated in ArcGIS using the spatial analyst and benthic terrain modeller extensions. Derived layers included slope, rugosity (indicates the ratio of surface area to planar area), aspect (identifies the direction of the steepest slope) and bathymetric position index at broad and fine scale.

Bathymetric Position Index (BPI) is a measure of where a referenced location is relative to the locations surrounding it. Derived from an input bathymetric data set, a neighbourhood analysis function produces an output raster in which the output cell value at each location is a function of the input cell value and the values of the cells in a specified "neighborhood" surrounding that location. As bathymetric position is an inherently scale-dependent phenomenon (Weiss, 2001) both fine scale and broad scale BPI data sets are usually created, whereby the fine scale BPI layer is generated using a smaller analysis neighbourhood than the broad scale BPI layer. In the present study

the default settings used in the Benthic Terrain Modeller extension to ArcGIS were applied to calculate coarse scale and fine scale BPI layers. These are, for broad scale BPI: inner radius = 1, outer radius = 5, and fine scale: inner radius = 1, outer radius = 3. Positive cell values within a BPI data set denote features and regions that are higher than the surrounding area. Conversely negative cell values within a BPI data set denote features and regions that are lower than the surrounding area. BPI values near zero are either flat areas (where the slope is near zero) or areas of constant slope (where the slope of the point is significantly greater than zero) (Weiss, 2001).

224

2.5 Sea-bed Substrate and Geomorphological Interpretations

All data were used to produce ArcGIS layers of sea-bed substrate and geomorphological features. For each digital stills image acquired a sea-bed sediment classification was assigned. These point classifications were used to ground-truth the multibeam echosounder and backscatter data allowing a complete sea-bed substrate interpretation to be created utilising all available data layers including derived layers. It should be noted that the backscatter quality was not suitable for accurate habitat differentiation as poor weather conditions and sea-state introduced noise which masked the more subtle geological variations of the sea floor. Following the sea-bed substrate classification, a geomorphological interpretation was created using standard geological terms and definitions. Geomorphology from the study area is described by 10 classes, with substrate described by 9 classes (Table 1).

237

238 The following 8 environmental data layers were prepared in ArcGIS 9.3 for use in
239 Maxent: continuous variables - bathymetry, slope, aspect, fine scale BPI, broad scale
240 BPI, rugosity, categorical variables - substrate and geomorphology.

241

242 **2.6 Maximum Entropy Modelling**

243 Maximum entropy modelling was introduced as a general approach to presence only
244 modelling of species distributions by Phillips et al. (2004; 2006). It has subsequently
245 been shown to perform very well against other presence only models (Elith et al.,
246 2006) with specific comparisons made between Maxent and environmental niche
247 factor analysis (ENFA) applied to predictions of the global distribution of stony corals
248 (Tittensor et al., 2009). Maxent estimates a target probability distribution by finding
249 the probability distribution of maximum entropy subject to a set of constraints that
250 represent our incomplete information about the target distribution (Phillips et al.,
251 2006). Put simply and in the context of the present study Maxent allows the user to
252 predict the distribution of a species/ habitat in terms of probability of occurrence, by
253 finding the distribution that agrees with everything known about the distribution of
254 the species / habitat (given the environmental data that has been provided to the
255 model), without making any assumptions about what is not known.

256

257 Single models were constructed for *L. pertusa* species distribution (SD) and *L. pertusa*
258 reef distribution (RD) using Maxent version 3.3.2, available for free download on

<http://www.cs.princeton.edu/~schapire/maxent/>. Maxent was run with default settings: convergence threshold 10^{-5} and maximum number of iterations of 500, regularisation set to 1, that have been shown to achieve good performance (Phillips and Dudík, 2008) even with small size datasets. However, for the RD model following visual inspection of the response curves and subsequent trials with increased regularisation, the regularisation parameter was set to 2 to reduce over fitting of the model. Maxent results are given in a logistic system where values near 0 mean low probability of presence and values near 1 mean high probability of presence.

2.7 Model evaluation

The models generated were evaluated in two ways.

Firstly, threshold independent ROC (receiver operating characteristic) curves were used to measure how successful the prediction was using the area under the curve (AUC) (Fielding & Bell, 1997). As a result of the sampling method used (video transects), the presence data for both SD and more obviously for RD were spatially autocorrelated within transects. To attempt to account for this in the model evaluation process cross validation of the both models was performed manually rather than using the Maxent replicates setting. For the SD model 1479 presence records (reduced to 102 cells with presence records) were obtained from 43 transects. These data were partitioned such that approximately 25% of the transects (10 or 11) constituting ~25-30% of the presence records were omitted from model building and used as a test dataset. This process was repeated 10 times and average AUC and standard deviation

of AUC across all 10 models was calculated. For the RD model the nature of the reef presence data was such that although there were 9 cells with reef presence, in truth this amounted to observation of 6 complete reefs (as one reef occupied more than one cell). Therefore cross validation of the RD model was performed by splitting the presence data into groups corresponding to the 6 reefs observed and using these data to manually run the Maxent model 6 times, leaving out one ‘complete reef’ presence each time. Average AUC and standard deviation of AUC across all 6 models was calculated. However the small total number of presence samples available to the model suggests that cross validation may be inappropriate given that test data sets may consist of a single presence point.

Secondly, the model assessment indices: percent correctly classified (PCC), specificity and sensitivity (Fielding & Bell, 1997), were calculated using the Presence-Absence Model Evaluation library (Freeman, 2007) in R (R Development Core Team, 2010). These indices require that a threshold be used to convert the continuous Maxent probability of occurrence prediction to a binary prediction delimiting presence or absence. Determining the appropriate threshold for Maxent models is an interesting and ongoing area of research (Liu et al., 2005). In this study three possible thresholds were assessed for their use in producing a reliable binary output map (Table 2). The three selected for testing were from the group of metrics identified as ‘good’ by Liu et al. (2005). The effectiveness of each threshold was evaluated in R (R Development Core Team, 2010) using the Presence-Absence Model Evaluation library (Freeman, 2007)

and model assessment indices listed previously. For SD and RD models the model build datasets were used together with absence data obtained from video analysis to assess the appropriateness of the thresholds. In addition thresholds were also assessed for each of the training and test model datasets and average performance of each threshold calculated. The most appropriate threshold was defined as that which resulted in constantly delivering the highest sensitivity score, since the precautionary principle suggests that false positives are less of a concern than false absences. For this study the sensitivity-specificity sum maximization approach where the sum of sensitivity and specificity is maximized Cantor et al. (1999), was selected (Table 2).

2.8 Assessment of variable importance within the models

Jackknife tests were undertaken to assess variable importance during model development by comparing the model gain (a measure of goodness of fit closely related to deviance) associated with models constructed with each variable omitted in turn, models constructed using individual variables only, and the full final model. Relative changes in gain between the full model and models constructed without one variable and with only one variable allow an assessment of relative importance of each variable to the final model build.

2.9 Quantification of species and habitat distribution

Binary maps for both SD and RD predicted distribution, produced using the sensitivity-specificity sum maximization approach threshold obtained for the full models (SD=0.2, RD=0.25), were used to quantify the difference in area suitable for *L. pertusa* species and *L. pertusa* reef habitat in ArcGIS 9.3.

3. Results

3.1 Model evaluation

For all partitions of the occurrence data, for both the *Lophelia pertusa* species distribution (SD) and *Lophelia pertusa* reef distribution (RD) models, the AUC values achieved by the training-test data were better than random (Table 3) (AUC>0.5). Full model AUC and mean training and test AUC for both SD and RD models could be rated a fair (0.7-0.8), good (0.8-0.9) or excellent (0.9-1). However, the consistently higher AUC value for the RD training-test models and the higher AUCs for the full model (Table 3) indicates that the whole RD model can better discriminate between suitable and unsuitable distribution areas for *L. pertusa* reef than for *L. pertusa* species.

Threshold dependent model assessment indices for the sensitivity-specificity sum maximization approach threshold also indicated that, when measured using PCC and sensitivity, the SD models performance was generally fair, while the RD models performance was generally excellent. Specificity scores for the SD models ranged

between awful (<0.6), poor (0.6-0.7) and good (0.8-0.9) reflecting the decision to select a threshold to maximise sensitivity scores.

3.2 Importance of environmental variables within each model

Jackknife tests of variable importance revealed the models generated for both SD and RD relied heavily on the substrate variable both in terms of having the most useful information by itself and having the most information that was not present in the other variables (Fig. 2). Analysis of response curves created through construction of Maxent models using single variables illustrated the dependence of predicted suitability on substrate, with the bedrock, gravel and sandy gravel categories as most important in the SD model and bedrock as most important in the RD model (Fig. 3).

Within the RD model geomorphology was the second most important variable in terms of having the most useful information by itself, but was of limited importance to the SD model (Fig. 2). The geomorphological categories flank, pinnacle/mound, ridge, and escarpment were of most use. However, for both SD and RD models omission of the geomorphological variable resulted in the second largest drop in gain after depth and substrate respectively suggesting that the geomorphological layer contained information that was not present in the other variables (Fig. 2).

For both SD and RD models fine scale BPI and broad scale BPI were the next most important variables in terms of providing the highest gain when used in isolation to construct models (Fig. 2). As the information in one BPI layer was essentially also contained in the other BPI layer, albeit at different resolution, in order to fully assess the importance of BPI to the SD and RD models, both models were rerun including only one BPI variable. However, in both the original models and the rerun models there was a negligible change in gain when BPI was omitted in jackknife analyses. This suggests that BPI does not contain any information that was not present in other variables. Analysis of response curves illustrated subtle differences between models in the dependence of predicted suitability on both fine scale and broad scale BPI. Within the SD model both high negative values and high positive values were of greatest importance (Fig. 3). However within the RD model only high positive values were of greatest importance.

Within the SD model rugosity and slope were the next most important variables (Fig. 2). However, omission of the slope or rugosity variable in jackknife tests resulted in a negligible drop in training gain. This suggests neither variable contained information that was not present in other variables. Analysis of response curves suggested the highest probability of occurrence of SD was achieved at rugosity of >1.01 and slopes of >20° (Fig. 3).

Jackknife tests of variable importance suggested that for both models bathymetry was of limited importance alone, however omission of the bathymetry variable from the models resulted in the third largest drop in training gain, suggesting the bathymetry variable contained information not present in other variables (Fig. 2). Heuristic estimates of relative contributions of the environmental variables to the Maxent model suggests that for SD and RD models bathymetry contributed 14.7% and 0.7% respectively. Analysis of response curves suggest the depths that resulted in the highest probability of SD presence were 500-900m, however for RD the probability of presence was predicted to increase with depth (Fig. 3), most likely a reflection of the limited depth range sampled.

Aspect was of least importance in both SD and RD models (Fig. 2) although interestingly probability of species occurrence was lowest at a bearing of 300°.

3.3 Potential distribution of *Lophelia pertusa* species and reef

Binary maps (1-presence, 0-absence) of both species and habitat distribution show *L. pertusa* species is distributed over a broader area (7.17% of the map area) than *L. pertusa* reef habitat (0.56% of the map area) (Fig. 4).

4. Discussion

4.1 Species vs habitat distribution

407 On Hatton Bank and George Bligh Bank, the models identify a broader area of high
408 suitability for the species than for reef (7.17% vs 0.56% of the total area respectively
409 using the sensitivity-specificity sum maximization threshold). Visual analysis of the
410 spatial distribution of areas predicted as highly suitable for species and reef suggests
411 that reef distribution is a highly restricted subset of species distribution.

412

413 Within the SD model the species was associated with bedrock, gravel and sandy gravel
414 substrate categories, with no one geomorphological class of particular importance.
415 These findings support what is currently known of the ecology of this species. *L.*
416 *pertusa* has a cosmopolitan distribution (Zibrowius, 1980) and occurs as individual
417 colonies under a relatively broad range of conditions, from depths of 40-3400m,
418 temperatures of between 4-13°C, salinities of 32-38‰ and across different oceans
419 including the NE Atlantic, Barents Sea, the Mediterranean, and the Gulf of Mexico
420 (Freiwald et al, 2004). *L. pertusa* is found on hard and mixed bottoms (Dons 1944;
421 Frederiksen et al., 1992) and in areas of fine sand where some form of hard substrate
422 is present for initial attachment. Wilson (1979) suggested that suitable substrata for
423 colony growth may be small e.g. mollusc shells, cobbles and boulders. It is not
424 surprising then, to find that the species is likely to be found over a relatively wide area

425

426 In comparison to the species, *L. pertusa* reef habitat is not widely distributed. The RD
427 model indicates that reef habitat is only likely to be present over small areas on both
428 Hatton Bank and George Bligh Bank. Within the wider NE Atlantic a limited number of

large reef structures (mound regions) have been identified (see Wheeler et al, 2007 for a review). *L. pertusa* only forms reefs under a specific set of environmental conditions. The largest reefs occur in depths between 500-1200m (Frederiksen et al., 1992; Wheeler et al., 2007) and may be associated with topographic features such as ridges (Sula Ridge), escarpments (Pelagia Mounds) and channels (Hovland Mounds) (Wheeler et al., 2007). Within the RD model reef habitat was clearly associated with bedrock substrate, and ridge, escarpment, flank and pinnacle/mound features.

While fundamental questions remain concerning the physical factors that are important in the development of reefs, recent research has highlighted the significance of hydrodynamic conditions in reef formation. Reef habitat forms in areas of enhanced turbidity, within a narrow density envelope, with high current velocities that prevent local sedimentation but provide enhanced encounter-rates with food particles (Thiem et al., 2006; Mienis et al., 2007; Dullo et al., 2008). These conditions must be stable over long periods of time to allow reef development (Thiem et al., 2006). Inclusion of hydrographic data in the model would undoubtedly improve the model fit and predictive power, however fine scale oceanographic data are not widely available. Geomorphology acts as a surrogate for fine scale current speed. The relationship between reef habitat and geomorphological features such as ridges and escarpments identified by the model most likely reflects both the substrate and hydrodynamic requirements of reef habitat development.

4.2 Importance of environmental parameters to predictive modelling of *L. pertusa* species and habitat distribution.

Within both the SD and RD models, substrate, geomorphology and BPI were the most important variables in terms of their importance to predicting species and habitat distribution, followed by rugosity and slope for SD and depth and slope for RD. These findings support those of Guinan et al. (2009a) who also found that at a slightly finer but comparable spatial scale (30m multibeam grids) the most important variables in predicting *L. pertusa* species distribution using GARP modelling, were rugosity and slope. At finer resolution (0.5m multibeam grids) Dolan et al. (2008) found BPI, structural complexity (fractal dimension) and orientation were important variables in using ENFA modelling. At coarser resolution (550m grids) Guinan et al. (2009a) found in addition to rugosity and slope, that aspect was weakly important. Neither Dolan et al. (2008) nor Guinan et al. (2009) had produced interpreted layers for substrate and geomorphology from multibeam bathymetry and backscatter and thus did not include these variables in their models. Following these studies Guinan et al. (2009b) concluded that coral abundance increases with increasing BPI, rugosity and slope.

Within the SD and RD models high positive BPI values were associated with high probability of occurrence suggesting both species and habitat are associated with raised features. High probability of species presence was also associated with high negative BPI values, suggesting the species may be associated with depressions as well as raised features. Probability of occurrence also increased with increasing rugosity

and slope to an asymptote (~1.16 and 30° respectively) suggesting values above this may not increase the probability of presence.

4.3 Implications for marine environmental management

The difference in area identified by the models as suitable for the species compared to the habitat has important implications for environmental management and the design of marine protected area networks (MPAs). As reef distribution is a restricted subset of species distribution, calculations of habitat extent based on the species distribution (e.g. within a given countries EEZ or an area covered by a particular convention such as OSPAR in the NE Atlantic), will be gross overestimates, and will thus mask the relative rarity of the habitat. In addition, in the complex task of identifying suitable boundaries for MPAs for the purpose of conserving reef habitat, boundaries drawn on the basis of species distribution may fail to include the target habitat. It is therefore desirable, where possible, to focus on the distribution of the habitat over the species at least at fine scales. However, it would be unwise to consider *only* reef habitat distribution in conservation planning, since connectivity between reef areas is likely to be maintained by the wider species distribution. Isolated colonies on cobbles may well provide a mechanism for gene-flow between larger reefs. This is particularly important as gene flow occurring among subpopulations is moderate at best with high levels of inbreeding and self-recruitment (Le Goff-Vitry et al., 2004). Research is needed into issues of connectivity with respect to MPA planning.

495 The use of predictive species modelling in the deep-sea is a relatively new field.
496 Recently models constructed at global and regional scales used broad-scale
497 oceanographic data at cell sizes of 130km, 1 degree and 0.25° to predict the
498 distribution of *L. pertusa* species (Davies et al., 2008; Tittensor et al., 2009.). These
499 models identified the levels of nitrate silicate and phosphate, aragonite saturation,
500 dissolved oxygen, and percent oxygen saturation as important in predicting the
501 distribution of *L. pertusa*. The spatial resolution of the environmental data used in
502 these models (in many cases data derived from model predictions) are inadequate to
503 capture the fine-scale current regimes likely to determine the specific sites at which
504 reefs are present (Guinan et al., 2007). Therefore, while on global and regional scales a
505 focus on modelling (and mapping) the distribution of the species is useful and
506 appropriate for assisting in targeting and coordinating conservation efforts (Davies et
507 al., 2008; Tittensor et al., 2009), the discrepancy between the areas of predicted
508 presence of the species and the habitat in this study suggest that predictive modelling
509 of habitat distribution at fine scales is more useful in terms of identifying areas of reef
510 occurrence.

511

512 The importance of substrate, geomorphology, BPI, rugosity and slope to habitat
513 distribution reflect the hydrodynamic conditions required for reef formation. This is
514 important in terms of future mapping efforts. For *L. pertusa* reef, variables derived
515 from multibeam bathymetry and interpreted backscatter data collectively act as
516 suitable surrogates for those environmental factors, which are critical in determining

517 reef distribution, but for which we generally lack fine-scale data. This suggests that it
518 may be possible in future to undertake multibeam survey of appropriate resolution for
519 large areas of the deep-sea and from that produce reasonable maps of reef
520 distribution with limited ground truthing required. It also suggests that the final model
521 produced here could be used to predict the distribution of reef habitat in other areas.
522 However, the restricted depth and temperature range of the study area limits the final
523 model to use in areas of similar environmental conditions.

524

525 What constitutes an appropriate resolution (multibeam grid size) for a given accuracy
526 of predictive map requires further investigation if managers are to make informed
527 decisions to balance predictive accuracy with survey cost. In addition the inclusion of
528 interpreted substrate and geomorphology layers and their resulting importance in
529 both SD and RD models suggests that these variables are particularly useful in
530 providing good predictions. However, these interpretations take considerable time and
531 skill to produce. In practical terms there may be a trade off between the time (and
532 expense) taken to interpret such datasets and the gain in the accuracy of predicted
533 distributions. Further research is needed into assessing such tradeoffs in the
534 application of these methods.

535

536 **4.4 Use of Maxent in predictive modelling of biological community distribution**
537 **(biotope mapping).**

The use of predictive modelling in marine community mapping is in its infancy (Kelly et al., 2001; Meleder et al., 2010). However, this technique has considerable benefits to offer to conservation efforts in the deep-sea where areas are vast, biological data are sparse and new survey is expensive. In shallow water areas remote sensing tools such as airborne and satellite imagery and aerial photography may be used to map the distribution of some habitats (Holmes et al., 2007). However, these tools rapidly reach their limits for sub-tidal surveys because of the absorption of visible radiation by water. At greater depths mapping is achieved using acoustic devices such as multibeam and sidescan sonar which are then ground truthed using video or other physical sampling methods (Brown et al., 2002; Huvenne et al., 2005; Brown and Blondel, 2009; Buhl-Mortensen et al., 2009). Although methods of mapping benthic assemblages vary, in general expert judgement is used to predict where assemblages will occur based on where they have been observed. This effectively amounts to predictive modelling using the mind. There is therefore a potential role for predictive modelling in biological assemblage (or biotope as defined by Dahl (1908)) mapping (Eastwood et al., 2006; Wilson et al., 2007). This study has demonstrated the potential use of the freely downloadable software Maxent to model the distribution of *L. pertusa* reef in benthic mapping efforts. This approach could be broadened and applied to other listed biogenic habitats / biological communities such as coral gardens, sponge aggregation etc, as well as any other defined benthic assemblages (biotope).

560 **Acknowledgements**

561 The authors would like to acknowledge with thanks the scientists, officers and crew of
562 SV Kommandor Jack and MV Franklin, the staff at Geotek and Marin Mätteknik AB, the
563 wider project partners J. Davies, C. Marshall, S. Mowles, L. Robinson, C. Jacobs, N.
564 Golding, N. Coltman, and B. Narayanaswamy; also D. Bilton for helpful comments on
565 the manuscript, C. Embling for useful discussions on thresholding and modelling in
566 general, and two anonymous reviewers for their helpful comments. The collection of
567 data used within was funded by the Department for Business, Enterprise and
568 Regulatory Reform through Strategic Environmental Assessment 7 (formerly the
569 Department for Trade and Industry) and the Department for Environment, Food and
570 Rural Affairs through the their advisors the Joint Nature Conservation Committee and
571 the offshore Special Areas for Conservation programme. Data analysis was partly
572 funded by a mini grant awarded to K.L.H. from the Census of Seamounts, a joint
573 Research Councils of the UK fellowship awarded to K.L.H, and University of Plymouth,
574 and the ERASMUS student exchange programme. H.S. publishes with the permission
575 of the Executive Director of the British Geological Survey (NERC).

576

577 **References**

578 Anderson, R.P., Gómez-Laverde, M., Peterson, A.T., 2002. Geographical distributions
579 of spiny pocket mice in South America: insights from predictive models. *Global Ecol.*
580 *Biogeogr.* 11, 131– 141.

581

582 Auster, P.J., 2005. Are deep-water corals important habitats for fishes? in: Freiwald, A.,
583 Roberts, J.M. (Eds.), Cold-Water Corals and Ecosystems. Springer, Berlin, Heidelberg,
584 pp. 747–760.

585

586 Brown, C. J., Cooper, K. M., Meadows, W. J., Limpenny, D. S., Rees, H. L., 2002. Small
587 scale mapping of sea-bed assemblages in the eastern English Channel using sidescan
588 sonar and remote sampling techniques. *Estuar. Coast. Shelf Sci.* 54, 263–278.

589

590 Brown, C. J., Blondel, P., 2009. Developments in the application of multibeam sonar
591 backscatter. *Appl. Acoust.* 70, 1242–1247.

592

593 Bryan, T.L., Metaxas, A., 2007. Predicting suitable habitat for deep-water coral in the
594 families Paragorgiidae and Primnoidae on the Atlantic and Pacific continental margins
595 of North America. *Mar. Ecol. Prog. Ser.* 330, 113–126.

596

597 Buhl-Mortensen, P., Dolan, M., Buhl-Mortensen, L., 2009. Prediction of benthic
598 biotopes on a Norwegian offshore bank using a combination of multivariate analysis
599 and GIS classification. *ICES J. Mar. Sci.* 66, 2026–2032.

600

601 Cantor, S.B., Sun, C.C., Tortolero-Luna, G., Richards-Kortum, R., Follen, M., 1999. A
 602 comparison of C/B ratios from studies using receiver operating characteristic curve
 603 analysis. *Journal of Clinical Epidemiology* 52, 885–892.
 604
 605 Costello, M.J., McCrea, M., Freiwald, A., Lundalv, T., Jonsson, L., Brett, B.J., van
 606 Weering, T.C.E., de Haas, H., Roberts, J.M., Allen, D., 2005. Role of cold-water *Lophelia*
 607 *pertusa* coral reefs as fish habitat in the NE Atlantic, in: Freiwald, A., Roberts, J.M.
 608 (Eds.), *Cold-Water Corals and Ecosystems*. Springer, Berlin, Heidelberg, pp. 771–805.
 609
 610 Dahl, F., 1908. Grundsätze und grundbegriffe der biocoenotischen forschung. *Zool.*
 611 *Anz.* 33, 349-353.
 612
 613 Davies, A.J., Wisshak, M., Orr, J.C., Roberts, J.M., 2008. Predicting suitable habitat for
 614 the cold-water coral *Lophelia pertusa* (Scleractinia). *Deep-Sea Res. Pt. I*, 55, 1048–
 615 1062.
 616
 617 Dolan, M.F.J., Grehan, A.J., Guinan, J.C., Brown, C., 2008. Modelling the distribution of
 618 cold-water corals in relation to bathymetric variables: adding spatial context to deep-
 619 sea video. *Deep-Sea Res. Pt. I*. 55, 1564–1579.
 620

621 Dons, C., 1944. Norges korallrev. K. Norske Vidensk. Selsk. Forh. 16, 37–82.

622

623 Dullo, C.W., Flögel, S., Rüggeberg, A., 2008. Cold-water coral growth in relation to the

624 hydrography of the Celtic and Nordic European continental margin. Mar. Ecol. Prog.

625 Ser. 371, 165– 176.

626

627 Eastwood, P.D., Souissi, S., Rogers, S.I., Coggan, R.A., Brown, C.J., 2006. Mapping

628 seabed assemblages using comparative top-down and bottom-up classification

629 approaches. Can. J. Fish. Aquat. Sci. 63, 1536–1548.

630

631 Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J.,

632 et al., 2006. Novel methods improve prediction of species' distributions from

633 occurrence data. Ecography 29, 129–151.

634

635 Embling, C.B., , Gillibrand, P.A., Gordona, J., Shrimpton, J., Stevick, P.T., Hammond P.S.,

636 2010. Using habitat models to identify suitable sites for marine protected areas for

637 harbour porpoises (*Phocoena phocoena*). Biol. Conserv. 143, 267-279.

638

639 Frederiksen, R., Jensen, A., Westerberg, H., 1992. The distribution of the scleractinian
640 coral *Lophelia pertusa* around the Faeroe Islands and the relation to internal tidal
641 mixing. *Sarsia* 77, 157–171.

642

643 Freeman, E., 2007. PresenceAbsence: An R Package for Presence-Absence Model
644 Evaluation. USDA Forest Service, Rocky Mountain Research Station, 507 25th street,
645 Ogden, UT, USA.

646

647 Freiwald, A., Fosså, J.H., Grehan, A., Koslow, T., Roberts, J.M., 2004. Coldwater Coral
648 Reefs. UNEP-WCMC, Cambridge, UK.

649

650 Guinan, J.C., Grehan, A.J., Wilson, M.F.J., Brown, C., 2009. Quantifying relationships
651 between video observations of cold-water coral and seafloor features in Rockall
652 Trough, west of Ireland. *Mar. Ecol. Prog. Ser.* 375, 125–138.

653

654 Guinan, J., Brown, C., Dolan, M.F.J., Grehan, A.J., 2009. Ecological niche modelling of
655 the distribution of cold-water coral habitat using underwater remote sensing data.
656 *Ecol. Info.* 4, 83–92.

657

658 Hitchen, K., 2004. The geology of the UK Hatton-Rockall margin. *Mar. Petrol. Geol.* 21,
659 993-1012.

660

661 Holmes, K.W., Van Niel, K.P., Kendrick, G.A., Radford, B., 2007. Probabilistic large-area
662 mapping of seagrass species distributions. *Aquat. Conserv.: Mar. Freshwat. Ecosyst.* 17,
663 385-407.

664

665 Hovland, M., 2005. Pockmark-associated coral reefs at the Kristin field off Mid-Norway,
666 in: Freiwald, A., Roberts, J.M. (Eds.), *Cold-Water Corals and Ecosystems*. Springer,
667 Berlin, Heidelberg, pp. 623–632

668

669 Howell, K.L., 2010. A benthic classification system to aid in the implementation of
670 marine protected area networks in the deep/high seas of the NE Atlantic. *Biol.*
671 *Conserv.* 143, 1041-1056.

672

673 Husebø, Å., Nøttestad, L., Fosså, J.H., Furevik, D.M., Jørgensen, S.B., 2002. Distribution
674 and abundance of fish in deep-sea coral habitats. *Hydrobiologia* 471, 91–99.

675

676 Huvenne, V.A.I., Beyer, A., de Haas, H., Dekindt, K., Henriët, J.P., Kozachenko, M., Olu-
677 Le Roy, K., Wheeler, A.J., 2005. The seabed appearance of different coral bank

678 provinces in the Porcupine Seabight, NE Atlantic, results from sidescan sonar and ROV
679 seabed mapping, in: Freiwald, A., Roberts, J.M. (Eds.), Cold-Water Corals and
680 Ecosystems. Springer, Berlin, Heidelberg, pp. 535-569.

681

682 Méléder, V., Populus, J., Guillaumont, B., Perrot, T., Mouquet, P., 2010. Predictive
683 modelling of seabed habitats: case study of subtidal kelp forests on the coast of
684 Brittany, France. Mar. Biol. 157, 1525-1541.

685

686 ICES., 2008. Report of the Working Group on the Biology and Assessment of Deep-Sea
687 Fisheries Resources (WGDEEP), ICES Headquarters, Copenhagen.

688

689 ICES., 2010. Report of the Working Group on Elasmobranch Fishes (WGEF), Horta,
690 Portugal.

691

692 Kelly, N.M., Fonseca, M., Whitfield, P., 2001. Predictive mapping for management and
693 conservation of seagrass beds in North Carolina. Aquat. Conserv.: Mar. Freshwat.
694 Ecosyst. 11, 437–451.

695

696 Le Goff-Vitry, M.C., Pybus, O.G., Rogers, A.D., 2004. Genetic structure of the deep-sea
 697 coral *Lophelia pertusa* in the northeast Atlantic revealed by microsatellites and internal
 698 transcribed spacer sequences. *Mol. Ecol.* 13, 537–549.
 699
 700 Liu, C., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of
 701 occurrence in the prediction of species distributions. *Ecography* 28, 385–393.
 702
 703 Mienis, F., de Stigter, H.C., White, M., Duineveld, G., de Haas, H., vanWeering, T.C.E.,
 704 2007. Hydrodynamic controls on cold-water coral growth and carbonate-mound
 705 development at the SW and SE Rockall Trough Margin, NE Atlantic Ocean, Deep-Sea
 706 Res. Pt. I 54, 1655–1674.
 707
 708 Mortensen, P., Buhl-Mortensen, L., 2004a. Deep-water corals and their habitats in The
 709 Gully, a submarine canyon off Atlantic Canada, in: Freiwald, A., Roberts, J.M. (Eds.),
 710 Cold-Water Corals and Ecosystems. Springer, Berlin, Heidelberg, pp. 247–277.
 711
 712 Mortensen, P.B., Buhl-Mortensen, L., 2004b. Distribution of deep-water gorgonian
 713 corals in relation to benthic features in the Northeast Channel (Atlantic Canada). *Mar.*
 714 *Biol.* 144, 1223–1238.
 715

716 Musgrove, F.W., Mitchener, B., 1996. Analysis of the pre-Tertiary rifting history of the
717 Rockall Trough. *Petrol. Geosci.* 2, 353-360.

718

719 Pearson, R.G., Raxworthy, C., Nakamura, M., Peterson, A.T., 2007. Predicting species'
720 distributions from small numbers of occurrence records: a test case using cryptic
721 geckos in Madagascar. *J. Biogeogr.* 34, 102–117.

722

723 Pearson, R.G., Dawson, T.P., Lin, C., 2004. Modelling species distributions in Britain: a
724 hierarchical integration of climate and land-cover data. *Ecography* 27, 285–298.

725

726 Phillips, S.J., Dudík, M., Schapire, R.E., 2004. A maximum entropy approach to species
727 distribution modelling, in: *Proceedings of the Twenty-First International Conference on*
728 *Machine Learning*, ACM Press, New York, pp. 83.

729

730 Phillips, S.J., Dudík, M., Schapire, R.E., 2006. Maximum entropy approach to species
731 geographic distributions. *Ecol Model.* 190, 231–259.

732

733 Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new
734 extensions and a comprehensive evaluation. *Ecography* 31, 161–175.

735

736 R Development Core Team, 2010. R: A language and environment for statistical
737 computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-
738 0, URL <http://www.R-project.org>.

739

740 Roberts, J., Wheeler, A.J., Freiwald, A., 2006. Reefs of the deep: the biology and
741 geology of cold-water coral ecosystems. *Science* 312, 543–547.

742

743 Thiem, Ø., Ravagnan, E., Fosså, J.H., Bersten, J., 2006. Food supply mechanisms for
744 coldwater corals along a continental shelf edge. *J. Mar. Syst.* 60, 207–219.

745

746 Tittensor, D.P., Baco, A.R., Brewin, P.E., Clark, M.R., Consalvey, M., Hall-Spencer, J.,
747 Rowden, A.A., Schlacher, T., Stocks, K., Rogers, A.D., 2009. Predicting global habitat
748 suitability for stony corals on seamounts. *J. Biogeogr.* 36, 1111–1128.

749

750 Weiss, A.D., 2001. Topographic Positions and Landforms Analysis. ESRI International
751 User Conference, San Diego, USA.

752

753 Wheeler, A. J., Beyer, A., Freiwald, A., de Haas, H., Huvenne, V.A.I., Kozachenko, M.,
 754 Roy, K.O.L., Opderbecke, J., 2007. Morphology and environment of cold-water coral
 755 carbonate mounds on the NW European margin, Int. J. Earth Sci., 96, 37–56.
 756
 757 Wilson, J.B., 1979. 'Patch' development of the deep-water coral *Lophelia pertusa* (L.)
 758 on Rockall Bank. J. Mar. Biol. Assoc. UK. 59, 165–177.
 759
 760 Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., Grehan, A., 2007. Multiscale
 761 terrain analysis of multibeam bathymetry data for habitat mapping on the continental
 762 slope. Mar. Geod. 30, 3–35.
 763
 764 Zibrowius, H., 1980. Les Scleractinaires de la Mediterranee et de l'Atlantique nord-
 765 oriental. Mem. Inst. Oceanogr. (Monaco) 11, p. 227.

766 Table 1: Sea-bed substrate and geomorphology classes identified on Hatton Bank and
 767 George Bligh Bank.

Code	Substrate	Geomorphology
1	Gravelly Sand	Scour
2	Gravel	Relatively Flat Lying Sea Bed
3	Bedrock	Pinnacle / Mound
4	Gravelly Muddy Sand	Escarpment
5	Sand	Iceberg Ploughmarks
6	Mud	Furrow
7	Sandy Gravel	Ridge Crest
8	Muddy Sand	Flank
9	null	Channel
10	Sandy Mud	Ridge
11	null	Depression

768

769 Table 2: Threshold dependent model performance metrics for SD and RD models for three different thresholding approaches.

SD models Threshold	Full model build data			Average training			Average test		
	PCC	sensitivity	specificity	PCC	sensitivity	specificity	PCC	sensitivity	specificity
Sensitivity-specificity equality	0.72	0.72	0.72	0.71 (0.08)	0.71 (0.08)	0.71 (0.08)	0.73 (0.12)	0.73 (0.12)	0.73 (0.12)
Sensitivity-specificity sum maximization	0.67	0.92	0.58	0.70 (0.05)	0.81 (0.17)	0.68 (0.07)	0.80 (0.10)	0.72 (0.16)	0.82 (0.10)
ROC plot-based approach	0.74	0.72	0.75	0.72 (0.03)	0.76 (0.13)	0.71 (0.03)	0.78 (0.10)	0.74 (0.14)	0.80 (0.10)
RD models									
Sensitivity-specificity equality	0.90	0.94	0.90	0.79 (0.05)	0.79 (0.05)	0.79 (0.05)	0.93 (0.08)	0.96 (0.10)	0.93 (0.08)
Sensitivity-specificity sum maximization	0.90	0.94	0.90	0.77 (0.14)	0.92 (0.11)	0.76 (0.14)	0.92 (0.11)	1.00 (0)	0.90 (0.13)
ROC plot-based approach	0.90	0.94	0.90	0.84 (0.11)	0.82 (0.07)	0.84 (0.11)	0.93 (0.08)	0.96 (0.10)	0.93 (0.08)

772 Table 3: Area Under the Curve (AUC) scores for SD and RD models for full models and
 773 all partitions of the occurrence data into training-test datasets.

	SD Model	Training AUC	Test AUC	RD Model	Training AUC	Test AUC
	Full model	0.964	0.808	Full model	0.998	0.940
	Cross validation models					
	1	0.839	0.695	1	0.924	0.792
	2	0.567	0.774	2	0.909	0.870
	3	0.856	0.744	3	0.946	0.912
	4	0.634	0.884	4	0.854	1.000
	5	0.797	0.721	5	0.876	1.000
	6	0.815	0.957	6	0.876	1.000
	7	0.822	0.854			
	8	0.821	0.802			
	9	0.828	0.627			
	10	0.830	0.973			
	Mean	0.781	0.803		0.897	0.929
774	Standard Deviation	0.098	0.113		0.035	0.087
775						

776 **Figure Captions**

777 Figure 1: The study area with sample details shown. Depth contours taken from the
778 GEBCO digital atlas and are in 100m isobaths down to 1000m, thereafter in 500m
779 isobaths.

780

781 Figure 2: Jackknife of regularised training gain for a) *Lophelia pertusa* species and b)
782 *Lophelia pertusa* reef. “Without variable” – each variable is excluded in turn and a
783 model created with the remaining variables; “With only variable” – model constructed
784 using only one variable; “With all variables” – full model build.

785

786 Figure 3: Response curves generated from a model built using only the corresponding
787 variable for a) *Lophelia pertusa* species and b) *Lophelia pertusa* reef. Y axis =
788 probability of presence, X axis label given above each plot, or for substrate codes 1-10
789 and geomorphology codes 1-11, see Table 1.

790

791 Figure 4: Modelled distribution of *Lophelia pertusa* species and *Lophelia pertusa* reef a)
792 on a subsection of Hatton Bank (inset) and b) on a small area identified in a. Maps c-e
793 show the individual environmental layers from the same area as b, and illustrate the
794 relationship between predicted presence areas and the 2 most important
795 environmental variables in each model.

20°0'0"W

18°0'0"W

16°0'0"W

14°0'0"W

58°0'0"N

58°0'0"N

a.

b.

3a.

3b.

18°0'0"W

SD Model	Training AUC	Test AUC	RD Model	Training AUC	Test AUC
Full model	0.964	0.808	Full model	0.998	0.940
Cross validation models					
1	0.839	0.695	1	0.924	0.792
2	0.567	0.774	2	0.909	0.870
3	0.856	0.744	3	0.946	0.912
4	0.634	0.884	4	0.854	1.000
5	0.797	0.721	5	0.876	1.000
6	0.815	0.957	6	0.876	1.000
7	0.822	0.854			
8	0.821	0.802			
9	0.828	0.627			
10	0.830	0.973			
Mean	0.781	0.803		0.897	0.929
Standard Deviation	0.098	0.113		0.035	0.087