Performance of deep-sea habitat suitability models assessed using independent data, and implications for use in area-based management

Howell, Kerry

https://pearl.plymouth.ac.uk/handle/10026.1/21043

10.3354/meps14098
Marine Ecology Progress Series
Inter Research

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.
Performance of deep-sea habitat suitability models assessed using independent data, and implications for use in area-based management.

Kerry L. Howell¹*, Amelia E Bridges¹, Kyran P. Graves¹, Louise Allcock², Giulia la Bianca¹, Carolina Ventura-Costa³, Sophie Donaldson¹, Anna-Leena Downie⁴, Thomas Furey⁵, Fergal McGrath⁵, Rebecca Ross⁶.

¹. School of Biological and Marine Science, Plymouth University, Plymouth. PL4 8AA. UK
². Ryan Institute and School of Natural Sciences, National University of Ireland Galway, Galway, Ireland.
³. University of Aveiro, Aveiro, Portugal
⁴. Centre for Environment, Fisheries and Aquaculture Science, Lowestoft, Suffolk. NR33 0HT. U.K.
⁵. Marine Institute, Rinville, Oranmore, Co. Galway. H91 R673

* Corresponding author email: kerry.howell@plymouth.ac.uk

Abstract

Marine spatial management requires accurate data on species and habitat distributions. In the deep sea, these data are lacking. Habitat suitability modelling offers a robust defensible means to fill data gaps, provided models are sufficiently reliable. We test the performance of published models of two deep-sea habitat-forming taxa at low and high resolutions (~1 km and 200 m grid-cell size), across the extended EEZs of UK and Ireland. We construct new data-rich models and compare new and old estimates of the area of habitat protected, noting changes in the protected area network since 2015. Results of independent validation suggest all published models perform worse than expected considering original cross-validation results, but model performance is still good or fair for Desmophyllum pertusum reef, with poorer performance for Pheronema carpenteri sponge models. High-resolution models using multibeam data out-perform low-resolution GEBCO-based models. Newly constructed models are good to excellent according to cross-validation. New model spatial predictions reflect published models, but with a significant reduction in predicted extent. The current marine protected area network and the European Union ban on bottom trawling below 800m protect 40% and 60% of D. pertusum reef-suitable habitat respectively, and 11% and 100% of P. carpenteri suitable habitat respectively within the model domain. We conclude high-resolution models of D. pertusum reef distribution are a useful tool in spatial management. The poorer performing P. carpenteri model indicate areas for more detailed study. Whilst low-resolution models can provide conservative estimates of percentage area-based conservation targets following the precautionary principle, high-resolution sea-floor mapping supports the development of better-performing models.

Key words: Deep sea; habitat suitability modelling; species distribution modelling; marine conservation; marine spatial planning
1. INTRODUCTION

As we begin the UN Decade of Ocean Science for Sustainable Development, the call for more holistic management of the marine environment is clear. Marine spatial planning (MSP) is an important tool supporting implementation of the ‘ecosystem approach’ to environmental management. An approach outlined in the Convention on Biological Diversity (CBD), and enshrined in the UN Sustainable Development Goals (UN General Assembly 2015). Maps lie at the heart of spatial management including maps of human uses, socio-economics, political and legal arrangements, and critically biophysical conditions and assemblages or communities of marine organisms, such as kelp forests and coral reefs. There is a pressing need to develop reliable accurate maps of the spatial distribution of marine ecosystems to support conservation initiatives.

Efforts to map benthic marine communities have largely focused on shallow water environments (Andrefouet et al. 2006, Traganos et al. 2018). Mapping deep-water communities is much more difficult because there is no direct equivalent to optical remote sensing which provides wide coverage of high-resolution data with direct observation of terrestrial and shallow (<10m) coastal habitats. The majority of deep-water benthic biological mapping has been achieved using modelling approaches. Species distribution modelling, also called habitat suitability modelling, uses data on the presence, absence, abundance or biomass of a species, assemblage or community, and relevant environmental data, to produce a statistical model of the relationship between species and their environmental drivers. The model can be used to make predictions of the distribution of the target species/community based on environmental data alone (Bryan & Metaxas 2007, Rengstorf et al. 2014, Howell et al. 2016). This type of mapping lends itself well to use in the marine environment as the
physical environment is often cheaper and simpler to measure than the biological components. There are a wealth of local, regional and global physical spatial models of the marine environment including oceanographic, bathymetric and productivity models. Benthic biological data are generally available for coastal marine areas and together with physical environmental data, offer great potential to produce relatively data-rich modelled maps. However, availability of benthic biological data decreases as you move away from the coast and into the deep sea (Webb et al. 2010), and this can present challenges in the development of reliable models.

The deep sea is increasingly subject to human use and there is an urgent need to implement more effective, integrated management of deep-sea ecosystems, through use of area-based management tools and marine spatial planning. The last 15 years has seen a growing trend in the use of predictive mapping techniques to generate models of the distribution of key species and assemblages in the deep sea (e.g. Bryan & Metaxas 2007, Guinan et al. 2009, Howell et al. 2011, Rengstorf et al. 2014, Robert et al. 2016, Howell et al. 2016, Pearman et al. 2020). These efforts have focused particularly on those species and assemblages that appear in key marine conservation legislation, and have produced modelled maps from a wide range of regions, of different extents and spatial resolutions. Models that provide large spatial coverage of a scale useful to national and regional MSP efforts, tend to use low-resolution (>1 km²) modelled global environmental datasets in their production (Howell et al. 2016). Higher-resolution environmental datasets, such as multibeam bathymetry, and regional or site-specific oceanographic models, tend to only be employed in the construction of models with a more limited spatial extent (Pearman et al. 2020), rendering them less useful for national and regional scale MSP, but still informative.
While the potential applications of modelled maps in MSP have been demonstrated (Ross & Howell 2013, Howell et al. 2016, Stirling et al. 2016, Rowden et al. 2017), models are not yet widely used despite their obvious potential (Marshall et al. 2014, Reiss et al. 2015). This contrasts with many other fields, for example fisheries and climate science, where models are routinely used to forecast future scenarios, and the results used to make management and policy decisions (Hilborn 2012, IPCC 2014). The reasons for this are not clear. Addison et al. (2013), in their review of common objections to the use of models in environmental decision-making, identify nine key objections that are symptoms of three fundamental issues: (1) misconceptions about the role of models in decision-making, (2) poor modelling practice and (3) a lack of effective communication and/or trust between modellers and decision-makers.

Objections around modelling practice and outputs include issues with model accuracy and uncertainty. Model performance is usually tested using random subsampling from the full model build dataset, so called cross-validation. However, the lack of true independence between testing and training data sets, as well as spatial sorting bias is known to artificially inflate model performances (Veloz 2009, Hijmans 2012) leading to a phenomenon where many models appear to perform well yet provide very different spatial predictions (Piechaud et al. 2015; Howell et al. 2016). This may serve to compound concerns around model accuracy and uncertainty, and highlight the need for independent validation of model performance to help allay these concerns and encourage wider use of model output.

In the North East Atlantic habitat suitability models for scleractinian cold water coral reef and an aggregation forming deep-sea sponge *Pheronema carpenteri* have been developed for the continental shelf claim areas of the UK and Ireland (Ross & Howell 2013, Ross et al. 2015). *Desmophyllum pertusum* reef is widely recognised as a distinct biological community or ‘biotope’, and occurs as thickets, discrete reefs, and giant carbonate mounds up to 300 m
high and several km in diameter. Within this region reefs have been observed on Hatton, George Bligh and Rockall Banks, the Wyville-Thomson Ridge, and in Explorer and Dangaard Canyons (Howell 2010, Howell et al. 2010), the Porcupine Seabight (Foubert et al. 2005, Huvenne et al. 2005), Porcupine Bank (Kenyon et al. 1998), southern Rockall Bank (Mienis et al. 2006, Wienberg et al. 2008) and Outer Hebrides (Roberts et al. 2005) as well as further north and south (Wheeler et al. 2007). Observations occur over depths from ~120 m to ~1000 m, with most reported from 600-800 m. Reef structures are highly biodiverse (Roberts et al. 2006), and have an important role as essential fish habitat (Husebø et al. 2002, Auster 2005). *P. carpenteri* is a small spherical glass sponge that occurs singularly or in dense aggregations, predominantly (but not exclusively) on fine sandy mud and mud substrata. Within UK and Irish waters, aggregations are a recognised biotope, and communities composed of this species have been described from 1250 m in the Porcupine Seabight (Rice et al. 1990), 1100 m in the Hatton-Rockall Basin (Hughes & Gage 2004, Howell et al. 2014), and from 1450 m on Goban Spur (Lavaleye et al. 2002), with historical records of additional aggregations from Ireland to Spain in 1000-2000 m water (Le Danois 1948) and in the Northern Rockall Trough (Wyville-Thomson 1874). Aggregations are associated with an increase in abundance and richness of macrofauna observed within spicule mats and sponge bodies (Rice et al. 1990, Bett & Rice 1992). Recent studies have suggested that known aggregations may be poorly connected (potentially isolated) (Ross et al. 2019) and experience a substantive impact from bottom trawl fishing (Vieira et al. 2020). From a policy perspective both *D. pertusum* reef and *P. carpenteri* aggregations are considered Vulnerable Marine Ecosystems (VME) under United Nations General Assembly Resolution 61/105, and as ‘threatened and/or declining species and habitats’ under the OSPAR Convention for the Protection of the Marine Environment of
the north-east Atlantic 1992. Understanding their distribution is therefore an important component to the development of area-based management of the region.

The Ross & Howell (2013) models were constructed using global scale environmental data layers and are at a resolution of ~1 km². The Ross et al. (2015) models were constructed using high resolution multibeam datasets and are at a resolution of 200x200 m grid cell size. Both models were produced using the same underlying presence / absence biological dataset for each response variable, *D. pertusum* reef habitat and *P. carpenteri* species. All four models performed well when tested using cross-validation methods, and in general high-resolution models performed better than low resolution models according to threshold-dependent evaluation. However, the spatial predictions and resulting maps derived from models of different resolution were notably different. The aim of this study is to undertake independent validation of these published models of VME distribution in the UK and Irish ECS claim areas, in order to assess model performance and inform future use in MSP and conservation. Specifically, we will 1) independently validate model performance using newly collected independent data, 2) construct new relatively data-rich models using the same modelling method as the prior publications, 3) quantify changes in predicted distributions and assessments of percentage protection targets for each VME (VME indicator taxa in the case of *P. carpenteri*) as a result of new models.

### 2. MATERIALS & METHODS

#### 2.1. Site and Model Description
The study considers the full extent of the Irish, and a partial extent of the UK’s extended continental shelf claim area in the N E Atlantic (Fig 1). A network comprising three different types of Marine Protected Area (MPA) exists in this area for the protection of deep-sea habitats (Fig 1). These are Special Areas for Conservation, OSPAR MPAs and North East Atlantic Fisheries Commission (NEAFC) closures to bottom trawling for the protection of VMEs. While the sites do not constitute a coherently designed MPA network, they enable illustration of the potential use of habitat maps in area-based management. In addition, there is a ban on bottom trawling below 800 m in European and UK waters.

Ross & Howell (2013) and Ross et al. (2015) scleractinian reef models were predominantly constructed using Desmophyllum pertusum reef presence / absence data. However, a small number of presence points for Solenosmilia variabilis reef were also included in the models. In our experience S. variabilis appears to occupy the same topographic niche as D. pertusum but occurs in deeper water. Records included in the models were at the shallow end of their distribution only. The resulting models largely predicted the niche of D. pertusum reef but with a slightly deeper reach. This study evaluates and builds upon the original models using new D. pertusum reef data alone with S. variabilis reef data omitted. Ross & Howell’s (2013) models are of resolution 750x750 m grid cell size, and cover the full extent of both Irish and UK continental shelf limit. Ross et al. (2015) models are of resolution 200x200 m grid cell size and cover the full extent of the Irish, and partial extent of the UK’s continental shelf limit in the N E Atlantic. Both studies used their models to assess progress towards percent protection conservation targets, and reported between 20 – 29% of scleractinian reef suitable habitat and 1.9-2.9% of P. carpenteri suitable habitat is within the MPA network.
2.2. Biological Data

New data for both *P. carpenteri* and *D. pertusum* reef were compiled from five research cruises to the northeast Atlantic: i) Eurofleets2 funded DeepMap cruise CE15011 (2015), with ROV Holland I; ii) NERC funded Deep Links JC136 (2016), with ROV ISIS; iii, iv, v) Sea Rovers RH17001 (2017), RH18002 (2018) and CE19015 (2019), jointly funded by the Irish Government and EU, with ROV Holland I. These research cruises were not conducted for the sole purpose of model validation, but this was a consideration in transect line planning for all cruises. Transect lines ranged from approximately 100m to 3.1 km, with an average length of 1.3 km. Collectively these research cruises provide a dataset consisting of 195 high definition ROV video transects spread across the study area (Fig 1). This collective dataset is referred to throughout as the new dataset.

For the original datasets presence of target habitat was determined from both quantitative and qualitative analysis of stills image data taken at 1 minute intervals along transects as described in Howell et al. (2010). *P. carpenteri* presence was determined from species lists from analysed sample data. *D. pertusum* reef habitat description follows that provided in Howell (2010), and subsequently adopted for use in the UK Deep Sea Habitat Classification (Parry et al. 2015). For the new independent dataset presence of the target habitat / species was determined by expert evaluation of image-based data alone. Habitat identification was undertaken by two annotators and designated when the habitat extent satisfied the OSPAR minimum biotope area threshold (25 m$^2$). For quality assurance, 5% of transects were independently analysed by Howell following inter-observer agreement standards used in published evidence (MacLeod et al., 2010).
2.3. Original Model Validation

For each of the four published models, the new biological dataset was plotted in ArcGIS on raster grids of published model output, in their respective output projections, and ROV point-based position data were reduced to one point per cell to avoid over-/under-weighting the importance of specific environmental conditions. Where cells contained any ROV position points interpreted as presence points, the one point per cell was denoted as a presence, all other points were denoted as absence. As the original models were masked for novel climates, new data points that did not sit on old model predictions were removed from the dataset as they were considered out of the original model domain. The final independent validation datasets for the 200 x 200m model included 2018 data points for *D. pertusum* reef and 1937 data points for *P. carpenteri* aggregations; for the 750 x 750 m model, the independent validation datasets included 646 data points for *D. pertusum* reef and 597 data points for *P. carpenteri* aggregations (Table 1). To assess the potential effect of spatial autocorrelation in inflating model performance, independent validation was also undertaken by reducing the datasets to one point per ROV transect. For each response variable a single presence point was randomly selected within each transect, and a single absence point from absence transects. This provided 173 and 163 validation points for the *D. pertusum* and *P. carpenteri* 200x200m models respectively and 186 and 182 validation points for the *D. pertusum* and *P. carpenteri* 750x750 models respectively.

The probability values from published model layers (coglog Maxent output) were extracted for each data point. Threshold independent metrics of model performance (Area Under the Receiver Operator Curve, AUC) for each model were calculated and compared to the original published models. Threshold dependent metrics of model performance (specificity,
sensitivity, and percent correctly classified) were also calculated by converting extracted probability values to binary presence-absence using 1) the thresholds defined in the original publications, and 2) new thresholds that maximised model performance against the new dataset.

2.4. Construction of new models

Newly collected high-resolution multibeam bathymetry data (Supplementary Material 1.0) were added to that described in Ross et al. (2015) and used to create grids of cell size 200x200 m that were re-projected from their original projection (WGS84) into Goode Homolosine Ocean (GHO) equal-area projection in order to allow for correct calculation of derived topographic layers and area.

2.4.1 Variable selection

Seven topographic variables were derived from the bathymetric data using the ArcGIS Benthic Terrain Modeller add-in (Walbridge et al. 2018): terrain ruggedness, curvature, plan curvature, profile curvature, slope, broad-scale bathymetric positions index (BBPI) and fine-scale bathymetric position index (FBPI). Information on the calculation and use of each of these variables can be found in the existing literature (Guinan et al. 2009, Ross & Howell 2013). The inner and outer radii for BBPI were 5 and 50 raster cells, respectively, facilitating identification of topographic features at 10 km scale such as canyons and hills. For FBPI, the inner and outer radiuses were 1 and 5 raster cells, respectively, allowing for the identification of features within the <1 km scale such as gullies. Generalised Additive Models (GAMs) were
used to build bottom temperature and salinity layers using *in-situ* CTD data from ROV and drop camera transects, as well as archived CTD casts from the British Oceanographic Data Centre (BODC) database. GAMs were implemented in R (R Core Team 2020) using the ‘mcgv’ package (Wood 2011) with depth, latitude and longitude used as explanatory variables. A detailed description is given in Supplementary Material 2.0. New and original biological datasets for each of *D. pertusum* reef and *P. carpenteri* presence / absence were combined, reprojected into GHO and plotted in ArcGIS on raster grids of environmental data. ROV/drop camera point-based position data were reduced to one point per cell, where cells containing any presence observations were denoted as a presence, all other points were denoted as absence. Environmental data were extracted for each data point.

Maximum Entropy (MaxEnt) modelling (Phillips et al. 2006, Elith et al. 2011) is a presence-background modelling technique that has a successful performance record (Elith et al. 2006), particularly in studies with low prevalence (low number of presence records). Although MaxEnt was designed to account for covariation in datasets and can perform well with correlated variables (Feng et al., 2019), previous studies have found that pre-selection of variables leads to better-performing models (Ross and Howell, 2013). Environmental variables were therefore first assessed for covariance using correlation matrices and Variance Inflation Factors in R. Strong correlations and VIFs between variables (≥ ±0.7 and ≥ 3, respectively) were addressed by removing one variable from each correlated pair based on the jackknife procedure. Jackknifing calculates the individual contribution of variables to a model and produces model performance statistics (termed ‘gain’ in MaxEnt) for each. Once correlates were removed, a model with all remaining variables was built. Following principles of model parsimony, final sets of variables were selected by systematically removing the variable contributing the least to the model (based on model gain with and without that
variable) until the drop in overall performance was deemed unacceptable. This process is described in detail in Supplementary Material 3.0 and 4.0.

2.4.2 Modelling

If used with presence-only data, MaxEnt randomly selects a specified number of ‘background’ points that are considered to represent locations with an equal likelihood of having been sampled that act as the absence points to inform the model (Elith et al. 2011). Whilst ‘absence’ points are presented in this study for each target taxa/habitat, it is not possible to be certain that they are not present somewhere within a 200 m grid cell due to the limited field-of-view of camera equipment compared to the size of grid cells, and therefore the data within this study represent ‘pseudo-absences’. Having pseudo-absence data allowed for the MaxEnt samples-with-data (SWD) approach to be used whereby environmental values are provided in a spreadsheet for both the presence and pseudo-absence points, instead of allowing MaxEnt to randomly select background points to act as absences. The benefit of the SWD approach is that as both the presence and pseudo-absence points come from the same sampling campaigns, it allows for the control of some bias in sampling locations and experimental design that can facilitate improved predictive performance (Phillips & Dudík 2008).

Preliminary models with different parameters were systematically trialed, including the changing of feature classes (linear, quadratic, product, hinge and threshold) and the regularisation parameter (0.1, 0.5, 1, 3, 5, 10) to avoid over-fitting/-smoothing (Phillips & Dudík 2008). The final feature classes selected for both target habitat models were linear, quadratic and product features. Through trialing, hinge and threshold features were removed
due to lack of ecological applicability in this study; with these features turned on, the response curves produced did not make biological sense. The *D. pertusum reef* model used a regularisation parameter of 1, whilst the *P. carpenteri* model used 0.5. These parameters were chosen because they struck a balance between the model overfitting and over-generalising - this was apparent from the shape of the response curves and AUC scores. The final MaxEnt models were projected onto the study area in a raster format and constrained to sampled conditions using the MaxEnt novel climates output (i.e. areas where environmental values fall within those on which the model was trained). Environmental data layers used in the final models are plotted in Supplementary Material 5 and final model details are provided in Supplementary Material 6.

2.5. Evaluation of New Models

Both presence and pseudo-absence records were used to evaluate the MaxEnt models’ performance by partitioning the data using a 70/30 split 10 times to create 10 sets of training and test data. These datasets were compiled manually rather than using the automated MaxEnt splitting tool to reduce spatial autocorrelation in the data. To achieve this, datasets were split such that whole transects fell into either a training or testing dataset. This avoided a scenario where a single transect would be split into training and testing points, leading to a within-transect testing point validating the same transect (Howell et al. 2011). The prevalence within each test and training dataset was compared to the prevalence of the full dataset and any datasets identified as having >+/−1% change in the amount of presence data were discarded and another random partition made until all test and train datasets satisfied the criteria. Using the partitioned data, 10 new models were built for each habitat and evaluated
using the ‘PresenceAbsence’ package (Freeman & Moisen 2008) in R, employing both threshold-independent (AUC) and threshold-dependent metrics.

Three thresholding techniques were used to assess model performance, as suggested in Liu et al. (2009), and recognising that different thresholding methods seek to achieve different ends. Chosen thresholds were sensitivity-specificity equality (Sens=Spec), sensitivity-specificity sum maximisation (MaxSens+Spec) and minimum distance to the top left corner in the receiver operating characteristic curve plot (MinROCdist). Using the presence.absence.accuracy() function, the thresholding techniques and resulting model performances were assessed using three widely used indices: sensitivity, specificity and percent correctly classified (PCC). True skill statistic (TSS) can be calculated from sensitivity and specificity and is used in place of Cohen’s kappa as it corrects the overall accuracy of the model predictions using the accuracy expected to occur by chance (Allouche et al., 2006). For both AUC and threshold-dependent metrics the mean and standard deviation for each metric was calculated for the 10 partitioned datasets and for the full model.

2.6. Quantification of Habitat Distribution and Marine Protected Area Analysis

The thresholding technique that gave the highest average of performance across the three chosen indices was selected for use in the final models. A binary raster of predicted presence and absence was produced as well as a raster of probability of predicted presence. Model fit was visualized by plotting the match-mismatch of binary predictions (Supplementary Material 7). In addition, the relative probability maps from all ten partitioned test/training models were used to produce standard deviation rasters to convey spatial uncertainty in the model predictions (Supplementary Material 8). The number of predicted presence raster cells
within different MPA polygons and below 800 m were calculated and then expressed as
percentages of total presences in the whole study area, UK waters, and Irish waters. Values
derived from published and new models were compared.

3. RESULTS

3.1. Original Model Validation

Results of the independent validation suggest that all published models perform worse than
expected based on cross-validation results for both threshold dependent and independent
metrics (Table 2 and 3). Model performance is still considered good (0.8–0.9) or fair (0.7–0.8)
for scleractinian cold-water coral reef habitat models, with poorer performance for the
*Pheronema carpenteri* models, particularly at low resolution. Independent validation using
the thinned dataset of one point per ROV transect (removing effects of spatial
autocorrelation) gave similar results. The extremely low prevalence of the *P. carpenteri*
dataset (Table 1) mean that model performance as measured by PCC is very much influenced
by correct prediction of absences (specificity), this also means threshold selection will be
strongly influenced by specificity and might explain why the new thresholds are all very low.
High-resolution models out-performed low-resolution models for both taxa.

3.2. New Models

Results of variable correlation analysis and step-by-step documentation of the variable pre-
selection procedure are provided in Supplementary Material 3.0 and 4.0.
3.3. New Model Evaluation

Consideration of common performance indices (Table 4) allowed for selection of final thresholding methods. For both models, Sens=Spec was selected as the chosen thresholding method, providing thresholds for *D. pertusum* reef and *P. carpenteri* aggregations of 0.44 and 0.37, respectively. For *D. pertusum reef*, the AUC value for the full internally validated model and all cross validation models was deemed excellent (0.9+). The 0.44 threshold determined by Sens=Spec generated good (0.8+) results for PCC, sensitivity and specificity for all models.

For *P. carpenteri*, the AUC value for the full and all cross validation models was deemed excellent. When thresholded at 0.26, all threshold-dependent metrics (PCC, sensitivity and specificity) for the full and training *P. carpenteri* models were classified as excellent (0.9+ full model and training sensitivity) or good (0.8-0.9 for training PCC and specificity) when internally validated. All cross-validation models were classified as good (0.8-0.9).

3.4. New Model Variable Importance

When variables are considered in isolation for *D. pertusum reef*, model gain is highest for temperature (70.5% contribution), followed by rugosity (23.3%) and FBPI (6.2%) as depicted in the jackknife plot (Supplementary Material 6.0). Temperature also decreased the model gain the most when removed as a variable, further illustrating its importance as the major variable on which predictions are reliant. For the *P. carpenteri model*, when variables are considered in isolation, model gain is highest for depth (41%) followed closely by temperature (35.9%), then BBPI (20.1%) and profile curvature (3%). When omitted from the complete
model the variable that decreased model gain the greatest was depth, closely followed by temperature.

3.5. Old (data poor) vs. New (data rich) high resolution models

Model performance determined by cross-validation suggests new models (Table 4) are comparable but of lower performance than old models (Table 3). New model spatial predictions in general follow those of the Ross et al. (2015) models, however, there are some notable differences (Fig. 2). Cold-water coral reef is predicted present on all banks, seamounts and the continental slope in the region, but the distribution is more restricted than that predicted by Ross et al. (2015). As with the previous model, *P. carpenteri* is predicted present on the continental slope, Porcupine Seabight, Rosemary Bank Seamount, around the Hatton-Rockall Plateau, and particularly in the Hatton-Rockall Basin. Presence is also predicted near the Wyville-Thomson Ridge where historical records refer to “the Holtenia grounds” (Wyville Thomson, 1874). The most noticeable difference is in the change in predicted distribution in the south-west section of the Hatton-Rockall Basin (circled in Fig. 2 c & d). Presence is predicted for both taxa inside the existing MPA network but, following the overall trend, the predicted distribution for *D. pertusum* reef is a contracted version of the 2015 predictions (Fig. 3 a and b). Predictions for *P. carpenteri* presence inside MPAs has changed little from the 2015 model.

3.6. Comparison of percentage area protected by 2015 MPA network

For both taxa there is a significant reduction in predicted extent of suitable habitat in km² in the new models when compared to the Ross & Howell (2013) and Ross et al. (2015) models.
The difference is most striking for *D. pertusum* reef where the low-resolution 2013 model predicts an extent 39 times larger, and the 2015 model 6 times larger, than the new model for the whole study area. Some of this reduction will be due to the removal of all *S. variabilis* data points from the model data, which will have led to a slight contraction in predicted depth range, however it is clear from Fig. 3 that there is a general contraction in predicted distribution between new and old models. As *D. pertusum* reefs are only found shallower than 1200m in this region, consideration of only those areas shallower than this depth reveal the same over-all trend. However, there is an increase in the estimates of the percentage of suitable habitat contained within the 2015 MPA network when calculated from the new model as compared to old models.

3.7. Assessment of percentage area protected by the MPA network present in 2020

Assessment of the proportion of suitable habitat included within the present day MPA network (Table 6) found that *D. pertusum* reef suitable environments are the most well protected within the study area (~40% contained within MPAs) with protection at national levels varying from 84% in UK to 24% in Irish waters. This is a significant increase from the 12-32% protection under the 2015 network assessed using all models (Table 5). *P. carpenteri* suitable habitat is the least well-protected of the two habitats assessed, with ~11% of predicted suitable environments included within a current MPA, with protection at national levels varying from ~49% in UK to ~4% in Irish waters. However, this again is a significant increase on the 2015 MPA network, where the new model suggests only 7% of suitable habitat was protected by the 2015 MPA network. The addition of new MPAs in UK waters between 2015 and 2020 have taken the UK from around 59% to 84% protection for *D.*
pertusum reef and from 25% to 49% protection for *P. carpenteri*. It should be noted however that the Ross et al. (2015) and new model only cover a partial extent of the UK’s continental shelf limit and data are biased to those areas that have been designated as MPAs. Thus, estimates of percentage protection are likely substantial overestimates. The EU ban on bottom trawling below 800m is estimated to protect 100% of the habitat suitable for *P. carpenteri*, and 42% of *D. pertusum* reef suitable habitat. Measured against IUCN targets both habitats are within or above the 20-30% protection level recommended.

4. DISCUSSION

4.1. Original Model Validation

Habitat suitability modelling (HSM) is a potentially valuable tool in the field of marine environmental management, but there remain questions around the true accuracy and reliability of modelled maps that may serve as a barrier to growth in use. In this study we have tested the performance of four published models at two different resolutions, 750x750 m (Ross & Howell 2013) and 200x200 m (Ross et al. 2015). Two for scleractinian cold water coral reef habitat and two for the sponge species *Pheronema carpenteri*. In the original published papers, all models performed well when tested using cross-validation methods, and performance was mixed when comparing low and high resolution models, according to threshold-dependent evaluation. While high-resolution *D. pertusum* reef models out performed low-resolution models, low-resolution models for *P. carpenteri* performed as well as high-resolution models according to threshold-dependent evaluation, and better than high resolution models according to threshold-independent evaluation (AUC). Our study has shown that when tested using independent data all models perform worse than
expected based on published cross-validation results for both threshold-dependent and independent metrics. Although models perform worse than under cross-validation, model performance is still considered good (0.9–0.8) or fair (0.8–0.7) for scleractinian cold-water coral reef habitat models, with poorer performance for the *P. carpenteri* sponge models, particularly at low resolution and when measured by sensitivity. High-resolution models out-performed low-resolution models for both taxa when assessed using independent data.

Our findings are in broad agreement with the very small number of comparable studies that have independently validated deep-sea sponge and coral HSM published models, with some notable differences. Rooper et al. (2016, 2018) independently validated HSM for corals and sponges in the eastern Bering Sea slope, outer shelf in Alaska and Aleutian Islands. These models were developed based on data from bottom trawl surveys at a resolution of 100x100 m grid cell size and validated using camera-based surveys. These studies found that while model performance decreased when comparing cross-validation to independent AUC scores, performance was still acceptable for coral models. This taken with our own findings suggest that high resolution models (<200x200 m grid cell size) of deep-sea coral distributions can be accurate and can provide useful information for spatial management of these vulnerable taxa.

However, low-resolution models may not perform well. Bowden et al. (2021) recently evaluated 47 HSM from eight published studies, all focused on the area around New Zealand, using independent data. All models were at 1km or 30 arc-seconds grid cell size, and in all cases model performance was lower than in published cross-validation values. Anderson et al. (2016) found that their models of the distribution of four scleratinian species
(not *Desmophyllum pertusum*) across the South Pacific Regional Fisheries Management Organisation area and adjoining EEZs were not successful in accurately predicting suitable habitat for reef-forming deep-sea corals when independently validated. These models were also constructed on a 30 arc-second grid (~1 km²) and data resolution was given as a possible explanation for model failure in the face of independent testing. Specifically, these authors cited the limitations of the bathymetry dataset used, which in turn affected the precision of each of the environmental predictor variables. Both studies report on models of comparable resolution to the low-resolution Ross & Howell (2013) model tested here.

Interestingly the Ross & Howell (2013) model appears to have performed better than the Anderson et al. (2016) models in the face of independent data. Anderson et al. (2016) cite missing critical predictor variables, particularly substrate type, lack of true absence data, spatial bias in distribution of presence records, and aspects of the topography in the study area, as possible reasons for their model’s poor performance. Ross & Howell’s (2013) model did make use of background data to account for spatial bias in the dataset, which may have resulted in better performance when subjected to independent testing. However, a principal difference between the Anderson et al. (2016) models and the Ross & Howell (2013) model is the focus of the models. Ross & Howell (2013) modelled scleractinian reef habitat where Anderson et al. (2016) modelled scleractinian species presence. The difference is important as the former occupies a restricted subset of the environmental niche of the latter (Howell et al. 2011), and a narrower niche width can result in a better performing model (Kadmon et al. 2003, Tsoar et al. 2007). This concept is used to explain the possible poor performance of Rooper et al.’s (2016, 2018) sponge models. These authors suggested that the difference they observed in their high-resolution (100x100 m grid cell size) coral and sponge model performance may be a result of lumping species
together into a large taxonomic group called ‘sponge’. This essentially merged species with very different habitat preferences, ultimately giving the group a wide environmental niche. The coral group in their study was dominated by a single family (Primnoidae) and thus was less affected by this pooling action.

Niche width is unlikely to explain the poor performance of the *P. carpenteri* model. This hexactinellid (glass sponge) is found predominantly on fine sediments where it loosely anchors to the substrate using long spicules at the base of the organism. Aggregations in the NE Atlantic are found over a very narrow depth range from 1000 to 1300 m (Rice et al. 1990) and appear to occupy a very specific niche. Cross-validation of HSMs created for this species suggested model performance was excellent (Ross & Howell 2013, Ross et al. 2015). However, independent validation suggests that while the models have fair to good PCC and specificity, they have poor sensitivity, meaning that the resulting maps may be indicating an absence where there is in fact a presence. Examination of the spatial distribution of false negatives suggests most (25 of 28 data points) are found on offshore seamounts and banks. These habitat types, and therefore this particular aspect of *P. carpenteri*’s environmental niche, was not represented in the dataset used to build the published models and could help explain why the models partially fail. However, aspects of the ecology of *P. carpenteri* may also explain the poor model performance.

*P. carpenteri*, in common with other deep-sea sponge species that form aggregations, are thought to be associated with regions of enhanced bottom currents related to the interaction of internal waves with sloping boundaries (Rice et al. 1990, Klittgaard et al. 1997, Davison et al. 2019) and raised features like the Mid-Atlantic Ridge (van Haren et al. 2017). The causal link is suggested to be an increase in the supply of food as a result of the
resuspension of organic matter (Rice et al. 1990). Oceanographic variables (and variability) may therefore be of critical importance in determining the distribution of *P. carpenteri*. The omission of such predictor variables from the Ross & Howell (2013) and Ross et al. (2015) models may also explain why both models partially fail when tested with independent data. The inclusion of oceanographic variables in deep-sea marine SDM has been found to improve model performance when tested with cross-validation (Rengstorf et al. 2014, Pearman et al. 2020) further supporting their inclusion in any future model development.

Our results suggest that for both scleractinian reef and *P. carpenteri*, the high-resolution models out-perform the low-resolution models when tested with independent data. This is an important finding as it suggests our ability to produce useful models of deep-sea benthic species and habitat distribution is dependent on availability of high-resolution environmental data including bathymetry data. Current maps of the seafloor are derived using satellite altimetry, which gives an average achievable resolution in the order of 8 km (Mayer et al. 2018). The percentage of the seafloor that has been measured by echo-sounders is considerably less than 18% and only about 9% of the seafloor is covered by high-resolution multibeam sonar data (Mayer et al. 2018). Recently an international effort has begun with the objective of facilitating the complete multibeam mapping of the world ocean by 2030. The Nippon Foundation GEBCO Seabed 2030 Project has the potential to improve significantly the quality of HSM it is possible to produce for deep-sea taxa by providing high-resolution bathymetry data. However, access to high-resolution oceanographic model output, as well as un-biased datasets of the distribution of target species and assemblages, and a good understanding of the biology and ecology of those species and assemblages, are also necessary to improve the quality of models. Targeted
efforts to collect these data over the next decade (Howell et al. 2020a,b) will be important in the further development of this field.

The good performance of the high-resolution scleractinian reef habitat model suggests that it may be a useful tool in the spatial management of cold-water coral reef in this region. Cold-water coral reef is considered a Vulnerable Marine Ecosystem (VME) under United Nations General Assembly Resolution 61/105, and, in the North East Atlantic is also classed as ‘threatened and/or declining habitat’ under the OSPAR Convention. Within European waters it is also recognised as an Annex I habitat under the EU Habitats and Species Directive (92/43/EEC). Collectively these policies require relevant management authorities to take actions to protect cold-water coral reef habitat. Specifically, UNGA 61/105 states “In respect of areas where vulnerable marine ecosystems, including seamounts, hydrothermal vents and cold water corals, are known to occur or are likely to occur based on the best available scientific information, to close such areas to bottom fishing and ensure that such activities do not proceed unless conservation and management measures have been established to prevent significant adverse impacts on vulnerable marine ecosystems”. Actions have so far been limited to those areas where cold water coral reef has been observed either through visual or physical sampling means. However, the high-resolution model provides best available scientific information on where cold-water coral reef is likely to occur in this region, and thus could be used to support decisions around further measures.

Similarly, the good performance of the high-resolution *P. carpenteri* model in terms of PCC and specificity, and fair AUC score, suggests it also may be a useful tool in the spatial management of the region. However, it must be noted that this is a presence / absence
HSM, and therefore it only indicates likely presence / absence of suitable habitat for the species (a VME indicator taxa), not the aggregation (a VME). This, together with the notable deficiencies in the model outlined above, suggests it is less useful than the scleractinian reef habitat model, but may still have value in indicating areas for further consideration given the precautionary principle.

4.2. New Model performance and interpretation

The newly constructed high-resolution models for *D. pertusum reef* and *P. carpenteri* have been developed using more than twice the input data used in the original Ross et al. (2015) models (Table 1), and consideration of oceanographic predictor variables (temperature and salinity) as terms in the models. Cross-validation suggests good performance for both models. In general, model performance increases with increasing sample size, however the nature of this relationship is variable and can depend on modelling method, prevalence, and species range size (Stockwell & Peterson 2002, Wisz et al. 2008, van Proosdij et al. 2016). The inclusion of oceanographic variables in deep-sea HSM has also been found to improve model performance (Rengstorf et al. 2014, Pearman et al. 2020). This suggests that the new models should perform better than the original 2015 models, although this can only be assessed using new independent data.

In this study we have used a presence–background approach rather than a presence-absence approach since our model input data are drawn from multiple surveys using multiple gear types and spanning more than 30 years. In our opinion, absences cannot be inferred from our dataset with certainty, and so we opted to be cautious in our use of absence data. However, it should be noted that evidence suggests presence-absence models perform better than
presence-only models, particularly where species/assemblages occupy all suitable habitat, making absence data reliable (Brotons et al. 2004), although Maxent has been found to perform equally as well as presence-absence models (González-Irusta et al. 2014). Future modelling efforts may wish to consider use of presence-absence approaches where authors feel absence data are reliable.

Temperature was a significant term in both new models and is a fundamental variable that controls species distributions. *D. pertusum* has been observed living under a wide range of temperatures (4–13 °C) (Freiwald et al. 2004), with an upper thermal tolerance of 15 °C (Brooke et al. 2013). Response curves for Maxent models for *D. pertusum reef* (Supplementary Material 6.0) suggest the highest likelihood of occurrence of reef habitat in the study area is at temperatures of ~8°C, which is almost the center of the species thermal niche. There are no data available on the thermal niche of *P. carpenteri*. Howell et al. (2016) reported this species to occur over a temperature range of 2.73–20.9 °C (mean 5.17 °C, standard deviation 2.03) in the northern North Atlantic. Response curves for Maxent models for *P. carpenteri* (Supplementary Material 6.0) suggest this species occupies a narrow thermal niche, with peak likelihood of occurrence at between 6-8°C, falling sharply to no occurrences below approximately 3°C or above 10°C. The wide range reported in Howell et al. (2016) is likely a result of poor position data from the older records used in that model in order to provide whole North Atlantic data coverage.

New model spatial predictions in general follow those of the Ross et al. (2015) models. However, there are some notable differences, particularly in the spatial prediction for *P. carpenteri* in the southern region of the Hatton-Rockall Basin (Fig. 2 c, d). In this region available CTD data suggest the temperature is cooler than that at equivalent depths in the
Rockall Trough and on the European continental slope, making this region less suitable for *P. carpenteri* than predicted by the 2015 model, which did not include temperature. Interestingly the Howell et al. (2016) model, which did include temperature, also predicted this area as suitable habitat, however the thermal niche of *P. carpenteri* was likely incorrectly defined in that model as previously noted. The principal difference in the spatial predictions for the *D. pertusum reef* model is a general contraction of the 2015 predictions in the current model. This is well illustrated in Fig. 3 a and b where current model predictions are much more focused than those of the 2015 model.

4.3. Re-assessment of current area closures and percentage protection targets for these VMEs

For both taxa there is a significant reduction in predicted extent in the new models when compared to the Ross & Howell (2013) and Ross et al. (2015) models (Table 5). The 2013 low-resolution models predicted 39 times and 4 times greater extent for *D. pertusum reef* and *P. carpenteri* respectively. This difference has important implications for onward use of models in decision-making. For example, calculations of ecosystem services such as carbon sequestration (Barnes et al. 2019; Barnes et al., 2021) or nutrient cycling (Hoffman et al. 2009) based on modelled extent may be grossly overestimated if based on low-resolution models. Similarly, the 2015 models predicted a greater extent of suitable habitat than the new model by 6 times and 1.4 times for *D. pertusum reef* and *P. carpenteri* respectively, suggesting that estimates of extent based on model predictions should be used with caution and considered likely overestimates.

In contrast, estimates of percentages of predicted suitable environments protected by the regional MPA network increased when calculated using the new model compared to the
2013 and 2015 models. The Convention on Biological Diversity originally set out a target of 10% of marine areas to be protected by 2010 (UNEP/CBD/COP/DEC/VII/5) (later moved to 2020 (UNEP/CBD/COP/10/27)), and that is now being followed up with calls for 30% by 2030 (CBD, 2020). While these percentage area targets are not habitat specific, Aichi Target 11 makes specific reference to ‘ecologically representative and well-connected systems of protected areas’ (UNEP/CBD/COP/10/27/Annex), which implies that different marine habitat types should be protected at that level. The independently validated 2015 models suggest that for the area modelled and the 2015 MPA network, both the UK and Ireland have surpassed the original 10% protection target for *D. pertusum* reef, while the UK have also surpassed this for *P. carpenteri* suitable habitat. In addition, the UK have surpassed the 30% target for *D. pertusum* reef habitat in the modelled area. The picture is the same for the new model. However, in both the 2015 and new model, Ireland protects <10% of suitable habitat for *P. carpenteri*, implying that further MPAs may be required. Ireland has committed to protecting 30% of its habitat by 2030 (Marine Protected Area Advisory Group, 2020) and data such as these can help guide that process.

The situation is broadly similar when considering the 2020 MPA network although the estimates of percentage of habitat protected in UK waters are much higher. It must, however, be noted that the current MPA network is not ‘strictly protected’ in line with IUCN specifications and in some cases management measures have yet to be drawn up. It is also important to remember that the modelled area in UK waters is much more limited than that modelled in Irish waters due to the limited availability of multibeam mapping in UK waters. The areas that have been mapped (and thus modelled onto) in UK waters tend to be associated with protected status, thus the UK figures are likely gross overestimates. Estimates of percentage of suitable habitat made from the low resolution 2013 model are
lower than all other estimates, and, reiterating the findings of Ross et al. (2015), suggests that low-resolution models result in conservative estimates in this context, which is in line with the precautionary principle and suggests low-resolution models may have a use in this area.

An interesting finding is that the ban on bottom trawling below 800 m in EU waters (UK is currently following) protects >30% of both habitats estimated from the new model with 100% of *P. carpenteri* suitable habitat protected. While a significant achievement, it is important to again consider the issue of representativeness in Aichi Target 11. Cold-water coral reefs occurring at different depths support different assemblages of associated species in line with the well-documented turnover of species along the depth gradient (Rowe & Menzies 1969, Howell et al. 2002, Carney 2005). In order to be representative, protection for cold-water coral reef sites must span its known depth range (thermal niche) necessitating protection of sites shallower than 800 m. In addition, the twin threats of ocean acidification and global warming mean that shallower areas of predicted suitable habitat in this region may be key refuge sites for cold-water coral reef (Jackson et al. 2014). Ocean acidification is causing the aragonite saturation horizon (ASH) to shoal exposing deep-water coral reefs to waters that are corrosive to coral skeletons (Guinotte et al. 2006). In parallel, seawater temperatures are increasingly exposing reefs to novel conditions. While live *D. pertusum* can tolerate long-term exposure to combined end-of-the-century temperature and pCO2 scenarios (Hennige et al. 2015, Büscher et al. 2017), the dead coral skeletons that make up the reef framework are weakened by acidified conditions and become more susceptible to bioerosion and mechanical damage (Hennige et al. 2015). This ultimately leads to crumbling, collapse, and loss of complexity of the larger habitat, and resulting ecosystem services (Hennige et al. 2020). In this region, the East Mingulay Special Area of Conservation (SAC), Wyville Thomson Ridge SAC,
and North West Rockall Bank SAC represent important strongholds for reef habitat (Jackson et al. 2014) and therefore the 800 m bottom-trawling ban alone will not meet the qualitative aims of Aichi Target 11.

5. CONCLUSION

Independent testing of four published models has shown that for the taxa considered, high-resolution models (<200x200 m grid cell size) can be accurate and can provide useful information for spatial management of these vulnerable taxa. With respect to UNGA Resolution 61/105, the high-resolution cold-water coral reef model provides best available scientific information on where this VME is likely to occur in this region, and thus could be used to support decisions around further measures. Our ability to produce useful models of deep-sea benthic species and habitat distribution is highly dependent on the availability of high-resolution environmental data including bathymetry data. To improve model performance significant research effort is needed to map the seafloor, oceanographic environment, and distribution of species and assemblages (presence, absences, density) in order to provide more, better quality, model input data. In addition, further research effort is needed to provide a more complete understanding of the importance of environmental variables to target taxa, and their interactions at a variety of scales. For well performing high-resolution models (200x200m), estimates of extent based on model predictions should be used with caution and considered likely overestimates. Low-resolution models (750x750m) may be useful in providing conservative estimates in progress towards percentage protection targets but are not recommended for use in estimates of extent. For D. pertusum reef and P. carpenteri the UK and Ireland have made good progress towards
the 10% CBD target for conserving habitats and species within MPAs. This together with the EU ban on bottom trawling below 800 m, provide a level of protection for both, however representativity needs to be considered in these assessments. Assessment of UK progress is limited by a lack of available multibeam data.

6. ACKNOWLEDGEMENTS

We would like to thank the scientists, officers and crew of all research cruises that have contributed to the collection of data for this study. The 2015 data collection on RV Celtic Explorer was funded under the European Union’s FP7 Research Infrastructures Programme under the grant agreement 312762 (EUROFLEETS2). The 2016 data collection was funded by the UK Natural Environment Research Council, grant number NE/K011855/1 - DeepLinks project. SeaRover reef habitat data acquired offshore Ireland during 2017, 2018, and 2019 have kindly been made available by the Government of Ireland in support of this research. The Sensitive Ecosystem Assessment and ROV Exploration of Reef (SeaRover) was commissioned by the Marine Institute in partnership with National Parks and Wildlife Service (NPWS), and funded by the European Maritime and Fisheries Fund (EMFF), Department of Agriculture, Food and the Marine (DAFM) & NPWS. The project was coordinated by the Department of Environment, Climate & Communications funded INFOMAR programme team, with research support from National University Ireland Galway, Plymouth University, and Institute of Marine Research Norway. INFOMAR is jointly managed by Marine Institute & Geological Survey Ireland. This article is delivered under the MISSION ATLANTIC project funded by the European Union’s Horizon 2020 Research and Innovation Program under grant agreement No. 639 862428. The following persons offered support in
data collection and advice throughout: David O’Sullivan, Yvonne Leahy, Janine Guinan, Nils Piechaud. This study uses CTD data provided by the British Oceanographic Data Centre.

7. REFERENCES


Barnes DK, Sands CJ, Richardson A, Smith N (2019) Blue carbon natural capital shallower than 1000 m in isolated, small and young Ascension Island’s EEZ. Front Mar Sci 6:663


Büscher JV, Form AU, Riebesell U (2017) Interactive effects of ocean acidification and warming on growth, fitness and survival of the cold-water coral *Lophelia pertusa* under different food availabilities. Front Mar Sci 4:101


Howell KL (2010) A benthic classification system to aid in the implementation of marine protected area networks in the deep/high seas of the NE Atlantic. Biol Conserv 143:1041-1056


Howell KL, Holt R, Endrino IP, Stewart H (2011) When the species is also a habitat: comparing the predictively modelled distributions of Lophelia pertusa and the reef habitat it forms. Biol Conserv 144:2656-2665


Hughes DJ, Gage JD (2004) Benthic metazoan biomass, community structure and bioturbation at three contrasting deep-water sites on the northwest European continental margin. Prog Oceanogr 63:29-55


Table 1: Breakdown, total (presence/absence), of biological datasets used to build habitat suitability models and independently validate Ross & Howell (2013) and Ross et al. (2015).

<table>
<thead>
<tr>
<th></th>
<th>Ross &amp; Howell 2013 (GEBCO)</th>
<th>Ross et al. 2015 (200m)</th>
<th>Howell et al. 2021</th>
<th>Ross &amp; Howell 2013 (GEBCO) Validation</th>
<th>Ross et al. 2015 (200m) Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D. pertusum</strong></td>
<td>864 (75/789)</td>
<td>1,284 (116/1,168)</td>
<td>3,291 (227/3,064)</td>
<td>646 (64/582)</td>
<td>2,018 (122/1896)</td>
</tr>
<tr>
<td><strong>P. carpenteri</strong></td>
<td>864 (53/811)</td>
<td>1,284 (74/1,210)</td>
<td>3,196 (139/3,057)</td>
<td>597 (32/565)</td>
<td>1,937 (66/1871)</td>
</tr>
</tbody>
</table>
Table 2: Performance statistics for the published Ross & Howell (2013) models according to original cross validation and new independent validation. Threshold values are predicted probabilities of presence. SD = Standard deviation.

<table>
<thead>
<tr>
<th>Method</th>
<th>Method</th>
<th>PCC (SD)</th>
<th>Sens. (SD)</th>
<th>Spec. (SD)</th>
<th>TSS (Sens+ Spec-1)</th>
<th>AUC (SD)</th>
<th>Threshold values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D. pertusum reef</strong></td>
<td><strong>MinROCdist</strong></td>
<td>0.82</td>
<td>0.75</td>
<td>0.82</td>
<td>0.57</td>
<td>0.86</td>
<td>0.48</td>
</tr>
<tr>
<td>Original cross validation with original threshold (Ross &amp; Howell 2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent validation with original threshold.</td>
<td></td>
<td>0.68</td>
<td>0.78</td>
<td>0.67</td>
<td>0.45</td>
<td>0.74</td>
<td>0.48</td>
</tr>
<tr>
<td>Independent validation but tuned to maximize model performance</td>
<td>Sens=Spec</td>
<td>0.70</td>
<td>0.70</td>
<td>0.70</td>
<td>0.40</td>
<td>0.74</td>
<td>0.50</td>
</tr>
<tr>
<td>(new threshold selected)</td>
<td></td>
<td>(0.02)</td>
<td>(0.05)</td>
<td>(0.02)</td>
<td></td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>Independent validation with original threshold and thinned dataset.</td>
<td><strong>MinROCdist</strong></td>
<td>0.77</td>
<td>0.71</td>
<td>0.78</td>
<td>0.49</td>
<td>0.79</td>
<td>0.48</td>
</tr>
<tr>
<td>Independent validation but tuned to maximize model performance</td>
<td>Sens=Spec</td>
<td>0.73</td>
<td>0.74</td>
<td>0.73</td>
<td>0.47</td>
<td>0.79</td>
<td>0.44</td>
</tr>
<tr>
<td>(new threshold selected) using thinned dataset</td>
<td></td>
<td>(0.03)</td>
<td>(0.07)</td>
<td>(0.04)</td>
<td></td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td><strong>P. carpenteri</strong></td>
<td><strong>MinROCdist</strong></td>
<td>0.95</td>
<td>0.96</td>
<td>0.95</td>
<td>0.91</td>
<td>0.99</td>
<td>0.19</td>
</tr>
<tr>
<td>Original cross validation with original threshold (Ross &amp; Howell 2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent validation with original threshold.</td>
<td></td>
<td>0.91</td>
<td>0.34</td>
<td>0.95</td>
<td>0.29</td>
<td>0.65</td>
<td>0.19</td>
</tr>
<tr>
<td>Independent validation with original threshold.</td>
<td></td>
<td>(0.01)</td>
<td>(0.09)</td>
<td>(0.01)</td>
<td></td>
<td>(0.05)</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>PCC (SD)</td>
<td>Sens. (SD)</td>
<td>Spec. (SD)</td>
<td>TSS (Sens+ Spec-1)</td>
<td>AUC (SD)</td>
<td>Threshold values</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>----------</td>
<td>------------</td>
<td>------------</td>
<td>--------------------</td>
<td>----------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td><strong>D. pertusum reef</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original cross validation with original threshold (Ross et al. 2015)</td>
<td>MinROCdist</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.70</td>
<td>0.91</td>
<td>0.43</td>
</tr>
<tr>
<td>Independent validation with original threshold.</td>
<td></td>
<td>0.72 (0.01)</td>
<td>0.88 (0.03)</td>
<td>0.70 (0.01)</td>
<td>0.58</td>
<td>0.87</td>
<td>0.43</td>
</tr>
<tr>
<td>Independent validation but tuned to maximize model performance (new threshold selected)</td>
<td>Sens=Spec</td>
<td>0.77 (0.01)</td>
<td>0.75 (0.04)</td>
<td>0.77 (0.01)</td>
<td>0.52</td>
<td>0.87 (0.01)</td>
<td>0.48</td>
</tr>
<tr>
<td>Independent validation with original threshold and thinned dataset.</td>
<td></td>
<td>0.82 (0.03)</td>
<td>0.82 (0.07)</td>
<td>0.82 (0.03)</td>
<td>0.64</td>
<td>0.90 (0.03)</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table 3: Performance of the published Ross et al. (2015) models according to original cross validation and new independent validation. Threshold values are predicted probabilities of presence. SD = Standard deviation.
<table>
<thead>
<tr>
<th></th>
<th>Sens=Spec</th>
<th>MinROCdist</th>
<th>MaxSens+Spec</th>
<th>MaxSens+Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent validation</td>
<td>0.82</td>
<td>0.96</td>
<td>0.84</td>
<td>0.74</td>
</tr>
<tr>
<td>but tuned to maximize model</td>
<td>(0.03)</td>
<td>(0.96)</td>
<td>(0.01)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>performance (new threshold selected)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>using thinned dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**P. carpenteri**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Original cross validation</td>
<td>MinROCdist</td>
<td>0.96</td>
<td>0.96</td>
<td>0.75</td>
</tr>
<tr>
<td>with original threshold (Ross et al.</td>
<td></td>
<td>(0.96)</td>
<td>(0.96)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent validation</td>
<td>0.90</td>
<td>0.47</td>
<td>0.90</td>
<td>0.75</td>
</tr>
<tr>
<td>with original threshold.</td>
<td>(0.01)</td>
<td>(0.06)</td>
<td>(0.02)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Independent validation</td>
<td>0.86</td>
<td>0.47</td>
<td>0.90</td>
<td>0.75</td>
</tr>
<tr>
<td>but tuned to maximize model</td>
<td>(0.03)</td>
<td>(0.13)</td>
<td>(0.02)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>performance (new threshold selected)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>using thinned dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent validation</td>
<td>0.84</td>
<td>0.67</td>
<td>0.86</td>
<td>0.75</td>
</tr>
<tr>
<td>with original threshold and thinned</td>
<td>(0.03)</td>
<td>(0.06)</td>
<td>(0.03)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>dataset.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent validation</td>
<td>0.84</td>
<td>0.6</td>
<td>0.86</td>
<td>0.75</td>
</tr>
<tr>
<td>but tuned to maximize model</td>
<td>(0.03)</td>
<td>(0.13)</td>
<td>(0.03)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>performance (new threshold selected)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>using thinned dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Threshold-dependent evaluation indices for training, test, and full models. Final thresholds and associated evaluation metrics shaded.

<table>
<thead>
<tr>
<th>Thresholding approach</th>
<th>Average Training – Internal validation</th>
<th>Average Test – Cross validation</th>
<th>Full Model – Internal validation</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCC (SD)</td>
<td>Sens. (SD)</td>
<td>Spec. (SD)</td>
<td>PCC (SD)</td>
</tr>
<tr>
<td><strong>D. pertusum reef</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sens=Spec</td>
<td>0.83 (0.01)</td>
<td>0.82 (0.03)</td>
<td>0.83 (0.01)</td>
<td>0.83 (0.01)</td>
</tr>
<tr>
<td>MaxSens+ Spec</td>
<td>0.81 (0.01)</td>
<td>0.89 (0.02)</td>
<td>0.80 (0.01)</td>
<td>0.80 (0.01)</td>
</tr>
<tr>
<td>MinROCdist</td>
<td>0.82 (0.01)</td>
<td>0.87 (0.03)</td>
<td>0.81 (0.01)</td>
<td>0.81 (0.01)</td>
</tr>
<tr>
<td><strong>P. carpenteri</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sens=Spec</td>
<td>0.88 (0.01)</td>
<td>0.88 (0.03)</td>
<td>0.88 (0.01)</td>
<td>0.89 (0.01)</td>
</tr>
<tr>
<td>MaxSens+ Spec</td>
<td>0.85 (0.01)</td>
<td>0.96 (0.02)</td>
<td>0.85 (0.01)</td>
<td>0.87 (0.01)</td>
</tr>
<tr>
<td>MinROCdist</td>
<td>0.87 (0.01)</td>
<td>0.94 (0.02)</td>
<td>0.86 (0.01)</td>
<td>0.89 (0.01)</td>
</tr>
</tbody>
</table>
Table 5: Area of predicted suitable habitat broken down into entire model extent, and model extent in UK and Irish jurisdictions. Percentage of predicted suitable habitat protected by the MPA and NEAFC Closure network used by Ross et al. (2015) for the purpose of comparison, also broken down by national MPAs and NEAFC Closures.

<table>
<thead>
<tr>
<th></th>
<th><strong>D. pertusum reef</strong></th>
<th></th>
<th><strong>P. carpenteri</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ross &amp; Howell 2013 GEBCO Model</td>
<td>Ross et al. 2015 200m² Model</td>
<td>New Model</td>
<td>Ross &amp; Howell 2013 GEBCO Model</td>
</tr>
<tr>
<td>Entire Model Extent</td>
<td>Predicted Suitable Habitat</td>
<td>185,240.25 km²</td>
<td>30,106.10 km²</td>
<td>4748.32 km²</td>
</tr>
<tr>
<td></td>
<td>Predicted Suitable Habitat within MPAs/NEAFC Closures</td>
<td>12.81%</td>
<td>20.00%</td>
<td>31.61%</td>
</tr>
<tr>
<td>Model extent within UK Shelf Claim</td>
<td>Predicted Suitable Habitat</td>
<td>57,425.06 km²</td>
<td>8,281.48 km²</td>
<td>1,244.00 km²</td>
</tr>
<tr>
<td></td>
<td>Predicted</td>
<td>29.84%</td>
<td>56.00%</td>
<td>58.95%</td>
</tr>
<tr>
<td>Model extent within Irish Shelf Claim</td>
<td>Predicted Suitable Habitat</td>
<td>48,139.31 km²</td>
<td>21,665.48 km²</td>
<td>3,412.36 km²</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Predicted Suitable Habitat within MPAs/NEAFC Closures</td>
<td>13.67%</td>
<td>12.60%</td>
<td>21.65%</td>
<td>1.10%</td>
</tr>
</tbody>
</table>
Table 6: Area of predicted suitable habitat for *D. pertusum* reef and *P. carpenteri* broken down into entire model extent, and model extent in UK and Irish jurisdictions. Percentage of predicted suitable habitat protected by the most up to date MPA/NEAFC Closure network and the EU / UK 800m trawl ban, also broken down by nation.

<table>
<thead>
<tr>
<th></th>
<th><em>D. pertusum</em> reef</th>
<th><em>P. carpenteri</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Entire Model Extent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted Suitable Habitat</td>
<td>4,748.32 km²</td>
<td>54,289.48 km²</td>
</tr>
<tr>
<td>800m Trawl Ban</td>
<td>60.11%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2020 MPA/NEAFC Network</td>
<td>40.26%</td>
<td>11.45%</td>
</tr>
<tr>
<td><strong>Model extent within UK Shelf Claim</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted Suitable Habitat</td>
<td>1,244.00 km²</td>
<td>8,886.76 km²</td>
</tr>
<tr>
<td>800m Trawl Ban</td>
<td>48.74%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2020 MPA/NEAFC Network</td>
<td>83.64%</td>
<td>49.30%</td>
</tr>
<tr>
<td><strong>Model extent within Ireland Shelf Claim</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted Suitable Habitat</td>
<td>3,412.36 km²</td>
<td>43,936.44 km²</td>
</tr>
<tr>
<td>800m Trawl Ban</td>
<td>64.25%</td>
<td>100.00%</td>
</tr>
<tr>
<td>2020 MPA/NEAFC Network</td>
<td>24.45%</td>
<td>3.79%</td>
</tr>
</tbody>
</table>
Figure 1: Map of the UK and Ireland’s Continental Shelf Limits (black line) showing the original dataset from Ross & Howell (2013), and Ross et al. (2015) together with the new dataset (compiled from five different surveys over 5 years) used to independently validate the models and subsequently build new models. The current network of deep-sea Marine Protected Areas is shown, together with the 800m isobath, below which bottom trawling is prohibited. Bathymetry shown is the 200 x 200 m gridded multibeam dataset (see text below for detail) shaded for depth with contours of 200m, 500m, 1000m and intervals of 1000m thereafter shown in grey. Map projected in British National Grid for aesthetic reasons.
Figure 2: Above threshold full model prediction maps for: (a) scleractinian cold-water coral reef distribution from Ross *et al.* (2015); (b) *D. pertusum* reef distribution with the new dataset; (c) *P. carpenteri* aggregation distribution from Ross *et al.* (2015); (d) *P. carpenteri* aggregation distribution with the new dataset. The Hatton-Rockall Basin is circled in red in c and d. White background indicates the prediction has been masked for novel climates. Maps projected in British National Grid for aesthetic reasons.
Figure 3: Examples of changes to protected area model predictions. (a) scleractinian cold-water coral reef distribution within the North-West Porcupine Bank MPA from Ross et al. (2015); (b) *D. pertusum* reef distribution within the North-West Porcupine Bank MPA with the new dataset. ROV transects plotted showing presences as yellow stars and absence as white circles. Maps projected in British National Grid for aesthetic reasons.